Aspects of semantics and their influence on word production in language impaired and unimpaired individuals

Leonie Flora Charlotte Lampe

BSc (Patholinguistics), MSc (Experimental and Clinical Linguistics)

Department of Cognitive Science

Faculty of Medicine, Health and Human Sciences

Macquarie University, Sydney, Australia

A thesis submitted in fulfilment of the requirements for a degree of

Doctor of Philosophy (PhD)

February 2021

The work presented in this thesis was carried out as part of the joint

International Doctorate for Experimental Approaches to Language and Brain (IDEALAB)

Macquarie University (Australia)

University of Newcastle (United Kingdom)

Rijksuniversiteit Groningen (the Netherlands)

Universität Potsdam (Germany)

The work presented in this thesis was funded by the

international Macquarie University Research Excellence Scholarship (iMQRES)

I dedicate this thesis to my curious and inquisitive grandmother, Eleonore Frieda Wendler.

Thank you for always supporting and encouraging me to pursue my education and to follow my dreams even if that meant that we were on different continents!

17/09/1930 – 18/09/2020

Contents

Thesis Abstract

Spoken word production is semantically mediated, but debates remain regarding how the structure and complexity of semantic representations influence the spread of activation at the semantic level and co-activation of other items at the lexical level and how this affects the speed and accuracy of, and brain activity during, word production. This thesis focused on six feature-based semantic variables that capture aspects of semantics (number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness) and investigated which of these variables affect picture naming performance. The underlying mechanisms of these variables were explored using a rich methodological approach focusing on different populations (participants with and without aphasia), types of data (behavioural and electrophysiological), and tasks (standard and speeded picture naming).

The experimental study reported in **Chapter 2** investigated effects of the semantic variables on picture naming in a large group of participants with aphasia. There were effects of number of semantic features, semantic similarity, and typicality on error types, some of which depended on the integrity of the participant's semantic system. A more homogeneous subgroup showed an effect of number of semantic features on naming accuracy. The results were interpreted in the context of current theories of semantics and word production and highlighted that these theories are underspecified regarding the mechanisms by which item-inherent semantic variables might operate.

Chapter 3 explored effects of the same semantic variables on picture naming in neurotypical participants. Number of semantic features facilitated performance, while intercorrelational density and distinctiveness had inhibitory effects. These findings were interpreted as being due to spreading activation at the semantic level and competition at the lexical level.

In **Chapter 4**, electrophysiological data collected during overt picture naming was analysed using waveform and microstate analyses. Number of semantic features was significant in the waveform analysis and in the microstate analysis number of semantic features, intercorrelational density, number of near semantic neighbours, and semantic similarity were found to influence activity in the semantic and lexical network involved in word production. This activity is suggested to be either related to the target word itself or distributed across a cohort of co-activated representations.

Chapter 5 reports a comparison of effects of semantic variables in speeded deadline and standard picture naming to test whether their effects are systematically stronger in speeded naming. There was a stronger effect of distinctiveness in speeded naming and a stronger effect of number of semantic features in standard naming. These differences could not be explained by greater responsiveness to input in the speeded naming task.

Overall, this thesis has resulted in a better understanding of the effects of semantic variables and underlying mechanisms in picture naming. To explain the effects, theories of word production require mechanisms of semantic facilitation and interference, which could be implemented as spreading activation at the semantic level and competition at the lexical level. However, most current theories of word production need further specification to explain these effects.

Declaration

I, Leonie Lampe, certify that the word in this thesis entitled "Aspects of semantics and their influence on word production in language impaired and unimpaired individuals" has not been previously submitted for a degree to any other institution or university. The work reported in this thesis was undertaken during the time I was enrolled as PhD candidate at Macquarie University under the supervision of Professor Lyndsey Nickels, Dr Solène Hameau, and Associate Professor Paul F. Sowman.

I also certify that this thesis is an original piece of research, which does not contain materials previously published and it has been written by me. Any help and assistance that I have received in my work, and the preparation of the thesis has been appropriately acknowledged. Furthermore, all information sources and literature that have contributed to the thesis have been acknowledged throughout the thesis.

The research presented in this thesis was approved by the Macquarie University Human Ethics Review Committee (reference numbers: 5201200905, *Understanding language processing, its breakdown and treatment*, and 3531, *Understanding language processing*).

Leonie Lampe, 5th of February 2021

Authorship contribution statement

Leonie Lampe: Conceptualisation of Papers 1–4, Methodology for Papers 1–4, Software for Papers 2–4, Formal analysis for Papers 1–4, Investigation for Papers 2–4, Data curation for Papers 1–4, Writing – original draft for Papers 1–4, Writing – Review & Editing for Papers 1–4, Visualisation for Papers 1–4, Project administration, Funding acquisition

Lyndsey Nickels: Conceptualisation of Papers 1–4, Methodology for Papers 1–4, Writing – Review & Editing for Papers 1–4, Supervision, Project administration

Solène Hameau: Conceptualisation of Papers 1–4, Methodology for Papers 1–4, Writing – Review & Editing for Papers 1–4, Supervision, Project administration

Paul F. Sowman: Methodology for Paper 3, Software for Paper 3, Formal analysis for Paper 3, Writing – Review & Editing for Paper 3, Supervision

Audrey Bürki: Conceptualisation of Paper 3, Methodology for Paper 3, Software for Paper 3, Formal analysis for Paper 3, Writing – Review & Editing for Paper 3

Nora Fieder: Conceptualisation of Paper 1, Methodology for Paper 1, Writing – Review & Editing for Paper 1

Acknowledgements

One of the first scientific papers I ever read, almost 10 years ago for a 1st year bachelor unit on aphasia diagnosis and treatment, was a paper by Professor Lyndsey Nickels. From the very beginning of my academic career and long before I met you, Lyndsey, your work inspired me and your fascination with the mechanisms of language and its breakdown was (and continues to be) infectious. Doing a PhD with you was literally a dream come true and I could not have wished for a better primary supervisor for my candidature! I am deeply grateful for your kindness, dedication, guidance, patience, advice, expertise, and your generous and thorough feedback, which all helped to shape my PhD journey into a fantastic experience. I cannot thank you enough!

I would also like to thank my other supervisors for their invaluable support. Most importantly, Dr Solène Hameau, who so generously shared her knowledge and time with me. Thank you for your encouragement throughout my PhD and for being such an enthusiastic supporter of my project!

I am also very thankful to Associate Professor Paul Sowman, Dr Nora Fieder, and Dr Audrey Bürki, who supported me at various stages of my PhD. Thank you for always being there to help, for your patience, and the valuable input you gave me—it was always much appreciated!

I was lucky to be part of the wonderful international PhD program IDEALAB and would like to extend my appreciation to those who have been working tirelessly to make this program possible: its directors Professor David Howard, Professor Lyndsey Nickels, Professor Barbara Höhle, Professor Roelien Bastiaanse, and Professor Gabriele Miceli, but also Lesley McKnight, Anja Pakpe, Alice Pomstra, and Leah Mercanti. Thank you for all your hard work, the academic and organisational support you gave along the way, and, for providing me and my colleagues with this special PhD experience!

I also consider myself fortunate for having been a PhD candidate in the Department of Cognitive Science. I thank Professor Genevieve McArthur, Associate Professor Paul Sowman, Lesley McKnight, and Katie Webb for creating an exceptionally welcoming and supportive environment and a sense of community that made even tough times, from getting stuck at international airports to having to work from home during COVID-19, manageable. Thanks also to Marcus Ockenden and Craig Richardson for technical support and Serje Robidoux and Peter Humburg for statistical advice.

Big thanks also to the members of the Aphasia and the Cognitive Neuropsychology groups at Macquarie University and the Cognitive Science: Language and Methods group at University of Potsdam—thank you for creating a stimulating environment, listening to countless practice talks, and for your sincere feedback throughout the years!

The research in this thesis would not have been possible without the participants sharing their time and efforts with me. Thank you!

A very special thank you goes out to all the wonderful friends I made during my PhD. Thank you for our precious little circle of niceness, your positivity, advice, the laughter, and, most importantly, for our friendship! Thank you, Ana Murteira, Leanne Ruggero, Inga Mehrani, Hannah Rapaport, Andrea Salins, Jidde Jacobi (and Myra Luinge!), Carola Ruiz Hornblas, Valentina Pescuma, Prerana Sabnis, Ella Creet, Ann-Kathrin Ohlert, and Aida Salcic. I would also like to thank my other IDEALAB colleagues who accompanied me on every step of this PhD journey: Irene Graafsma, Kaimook Siriboonpipattana, Suzan Dilara Tokac, and Frank Tsiwah. Special thanks also to my wonderful colleagues in the Department and beyond: Diane Mézière, Gemma Perry, Omid Ghasemi, Bruna Tessaro, Maria Korochkina, Ryssa Moffat, Luke Mills, Sara Hjortborg, Nick Cooney, Chris Hewitson, Caroline Candebat, Julianna Kadar, Catarina Vila Pouca, and to all my friends in Europe. I am so grateful that I shared this experience with all of you!

I would like to thank my family for all their support along the way. Mama, Papa, Jasper, Jens, Karl, Steffi, Elke, Helmut, Anne, Laura, Norbert, Martin, Greta, Martha, Nele und Frieda, ich danke euch für eure Unterstützung, euren Enthusiasmus für dieses Abenteuer, eure Besuche down under, und für jedes *Guten Morgen* am Abend und *Gute Nacht* zum Frühstück!

And lastly, I want to thank you, Matti, for your love and patience throughout the years! Du warst meine größte Stütze auf diesem Weg, hast mich aufgeheitert, abgelenkt und mir zugehört. Ich danke dir für die zahllosen großen und kleinen Abenteuer der letzten Jahre und freue mich auf viele weitere mit dir!

I could not have done it without you all!

CHAPTER 1

General Introduction

Verbal communication is a fundamental human skill. We often use spoken language to connect with other people in our environment as it allows us to share thoughts and to express feelings with relative ease. Verbal communication enables us to share messages with other humans, be it in a professional conversation about the latest business figures, in an emotional argument of a couple, or when cracking a joke with friends. Transmitting a message to a communication partner is usually an effortless task, which does not involve much conscious planning and engagement. However, despite this seemingly easy process, verbal communication is actually a highly complex construct, which we heavily rely on to function flawlessly. It involves formulating a message and translating it into a word form, which has to be selected amongst many thousands of words in the mental lexicon. Subsequently, this abstract word has to be articulated, which requires the selection of the appropriate sounds and ultimately the synchronised activation of oral and facial muscles is needed for production. If the message needs to be expressed in multiple words or sentences, these have to be assembled and connected to one another in a grammatically correct and coherent way. We get an insight into the full magnitude of this complexity when verbal communication breaks down, as can be the case in people with language impairments, which, for example, often occur after stroke.

The seemingly most simple elements of verbal communication are single words. However, even for this apparently basic level of verbal communication, psycho- and neurolinguistic research is just beginning to understand the processes involved in successful word production and to grasp how and why the production of words might be disordered in people with language impairments. In this thesis, I focus on how the structure of, and relationship between, the meanings of words influence word production in people with and without language impairments. In this General Introduction I lay out the significance of the topic for the broader research context, present relevant theoretical frameworks, and introduce the most important concepts for the experimental chapters of this thesis.

Theoretical accounts of word production

Research into word production has yielded numerous attempts to formally describe the processes that are likely involved (e.g., Abdel Rahman & Melinger, 2009; Dell, 1986; Howard et al., 2006; Levelt et al., 1999). These theories are intended to describe processes during word production, for example the type of information processed at each level of the model, the number and characteristics of the levels involved in word production, and the dynamics within and between levels. To formalise the structures and dynamics required for word production, authors of these theories often drew on data from impaired word production or experimental paradigms in which participants name pictures in specifically designed contexts.

The different models of spoken word production agree that it is, broadly speaking, a two-stage process with a first meaning-based and a second phonology-based stage (Butterworth, 1989; Garrett, 1980; Levelt, 1989). These broad stages are concerned with different kinds of information: semantic, lexical/syntactic, and phonological. More specifically, processing of word meaning (semantics) occurs at a conceptual level, selection of an abstract word form at a lexical level, and access to phonemes at a phonological level. However, this broad layout quickly gains complexity with model-specific details on information representation and processing dynamics, most of which are not of direct importance for this thesis. Different models of word production disagree about the direction of activation flow (feedforward vs bi-directional, allowing for feedback), timing of activation spread within or between levels of the model (sequential vs cascading vs interactive links), and the type and organisation of information at the different levels of the model. For example, different models disagree about the number and function of sub-processes necessary for lexical-syntactic processing: While some models propose only one level of abstract word form (e.g., Caramazza, 1997) others suggest a two-stage process differentiating between *lemmas* (semantic-syntactic information) and *lexemes* (morphological-phonological information) (e.g., Levelt et al., 1999). In this thesis, to remain neutral in this debate, I use the term *lexical representations* and refer to processing at the *lexical level,* following Rapp and Goldrick (2000).

Two broad distinctions can be drawn to group models of word production. The first distinction relates to the representation of word meaning proposed by the different models (i.e., decomposed into semantic features vs non-decomposed holistic representations) and the second distinction concerns the process underlying lexical selection (i.e., competition vs no competition). I review both of these concepts below before introducing the most prominent theories of word production, with a focus on their proposed architecture of semantic representation and processing and the lexical selection process, the two elements of word production that I am most concerned with in this thesis.

Semantic representation

In this section, I describe how word knowledge is represented in current theories of word production, but do not aim to give a full account of the various theoretical debates around (computational) models of semantics. As argued by Vigliocco and Vinson (2007) and Vinson et al. (2013), current word production theories are mostly underspecified concerning a clear standpoint on any differentiation between non-verbal conceptual world knowledge and the part of conceptual knowledge that can be verbalised (i.e., semantic memory), which contains the lexical-semantic representations of individual words (McRae & Jones, 2013). Most theories of word production do not go into detail on distinctions between these concepts and start the formalisation of the word production process with lexical semantics (e.g., Dell, 1986; Levelt et al., 1999). As the distinction between conceptual and lexical-semantic representations is not a focus for most word production theories, I do not attempt to differentiate between them in this thesis. I generally refer to a generic semantic level, where processing of semantic representations (e.g., lexical concepts) takes place.

Moreover, current theories of word production make simplified assumptions regarding the characteristics of semantic representations. Outside of word production research, elaborate models of semantic representations and semantic processing have been proposed and are heatedly discussed, such as connectionist attractor networks (e.g., Cree et al., 1999; Farah & McClelland, 1991; Masson, 1995; Plaut & Shallice, 1993). However, these advances have usually not been incorporated in theories of word production, presumably because most experimental findings in word production research can be explained in the context of highly simplified semantic representations (Vinson et al., 2013). However, despite being underspecified regarding many of the more recent points of discussion in, and developments of, semantic research, current production theories do differ with respect to their proposed semantic representations: While some theories consider word meaning to be holistically represented, others assume word meaning to be decomposed into semantic features.

Holistic representations

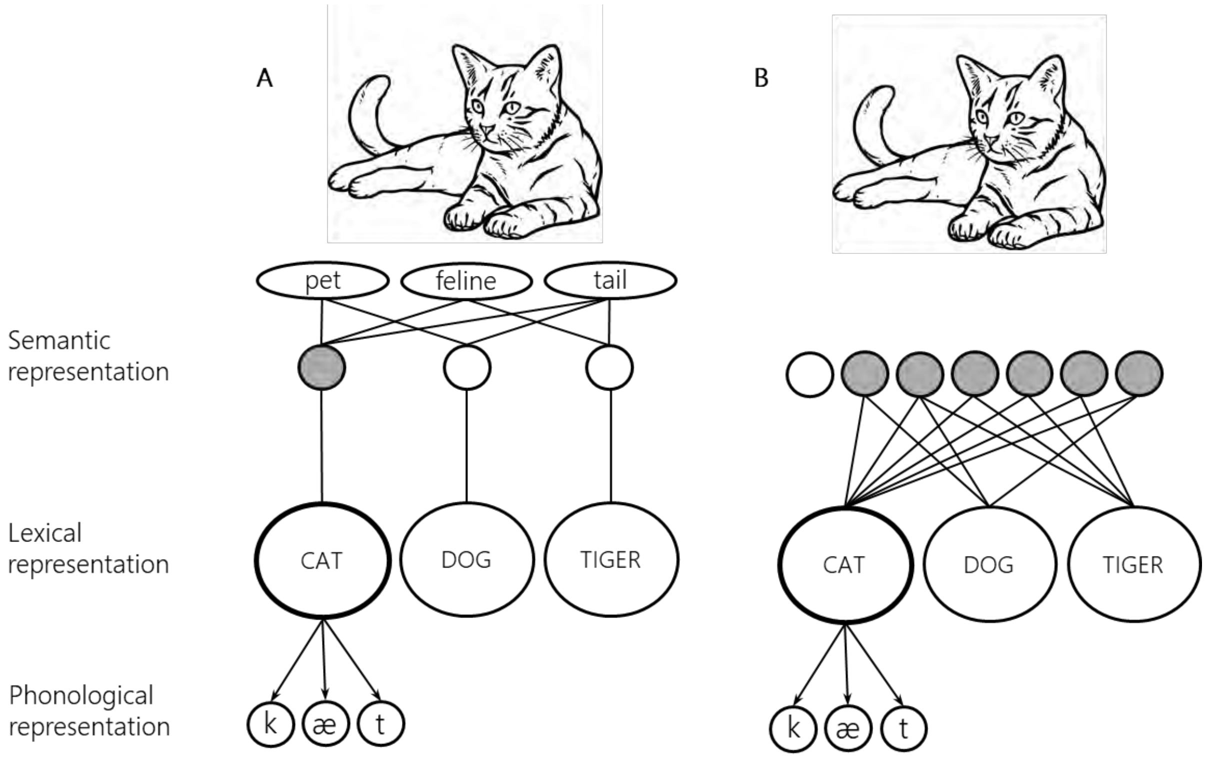
Holistic theories of semantic representations assume that word meaning is characterised as *lexical concepts*, which are abstract, unitary representations of the words of a language (e.g., 'cat'). Hence, lexical concepts are a verbalisable subset of all conceptual knowledge. Relationships between these holistic representations are expressed via labelled links between concepts that represent various types of relations (e.g., hierarchical *is a* or part-whole relations *has a*: 'cat' *is a* 'pet' and *has a* 'tail'; e.g., Collins & Loftus, 1975; Roelofs, 1992). Although each lexical concept maps onto a single lexical representation, activation that spreads between concepts via these concept-concept links can cause multiple lexical concepts and their lexical representations to be co-activated (e.g., 'cat' activates 'pet' and 'tail', which each activate 'dog'; schematically represented in Figure 1, Panel A).

Decomposed representations

In contrast, word production theories that propose semantic representations to be decomposed assume that word meaning is broken down into smaller units, *semantic features* (e.g., *purrs, has fur, has whiskers, has a tail*, etc. for the concept 'cat'), which, in combination, express the complex meaning of a word. According to theories assuming semantics to be decomposed into semantic features, relationships between words are driven by featural properties, such as feature overlap. Moreover, activation is often forwarded from each semantic feature to multiple lexical representations for which this given semantic feature is characteristic, causing multiple lexical representations to be simultaneously co-activated (e.g., Dell, 1986; schematically represented in Figure 1, Panel B). If bidirectional excitatory connections between levels of the model are assumed, this may cause the semantic features of semantically related words (e.g., 'dog') to become activated through activation of the lexical form of these words via the shared semantic features with the target (e.g., Dell et al., 1997).

Figure 1

Simplified schematic representation of the spread of activation when naming the picture 'cat' assuming a decomposed (Panel A) or a holistic semantic architecture (Panel B)



Note. Semantic representation activated by the picture is represented by grey circles (lexical concept or semantic features, respectively). The selected lexical representation 'cat' is represented by a stronger outline.
Picture source 'cat' drawing: https://bit.ly/2Smkm8n.

Lexical selection

Theories of word production generally agree that during word planning multiple lexical representations receive activation at the lexical level, via activation spreading between lexical concepts (e.g., Levelt et al., 1999) or to multiple lexical representations based on shared semantic information (e.g., Dell et al., 1997; Figure 1). However, there is disagreement about whether these co-activated representations influence the production of a target word, with some theories suggesting lexical selection to be a competitive process while others do not assume selection of a target word to be influenced by other co-activated representations.

Competitive lexical selection

A prominent position is the *lexical selection by competition* account (e.g., Abdel Rahman & Melinger, 2009, 2019; Howard et al., 2006; Levelt et al., 1999; Roelofs, 2018), which proposes negative influences of co-activated lexical representations on the selection of the target word's lexical representation, slowing this process down and making it more susceptible to error.

However, in the class of models proposing lexical selection to be a competitive process, there is disagreement about the exact mechanism underlying this interference. One commonly proposed mechanism is the *Luce ratio*, a mathematical rule according to which the highest activated lexical representation is selected, with the selection mechanism taking both the level of activation of the target and the co-activated representations into account (Levelt et al., 1999; Luce, 1959). Alternatively, lateral inhibitory links between co-activated lexical representations have been suggested, where the greater the activation of another lexical representation the slower the rise in activation, and ultimately selection, of the target representation (e.g., Harley, 1993; McClelland & Rumelhart, 1981; Stemberger, 1985). Yet, regardless of the exact mechanism, interference from co-activated representations should increase the greater the number of co-activated lexical representations and the higher their activation levels.

Non-competitive lexical selection

In contrast to theories incorporating lexical selection by competition, other theories of word production do not see the necessity for lexical selection to be competitive (e.g., Dell, 1986; Mahon et al., 2007; Oppenheim et al., 2010). In these models, a lexical representation is selected once it surpasses a certain threshold level of activation (e.g., Oppenheim et al., 2010). Alternatively, the highest active representation is selected after a fixed number of time steps (e.g., Dell, 1986). Importantly, under neither of these suggestions is the selected lexical representation actively affected by the activation levels of co-activated lexical representations.

In models with non-competitive lexical selection, inhibitory effects observed during word production are linked to different mechanisms, such as processes at pre-lexical (in the form of a learning mechanism; e.g., Navarrete et al., 2014; Oppenheim et al., 2010) or post-lexical levels (Mahon et al., 2007) (more detail in the section on the most important models of word production).

Below, I review the most prevalent theories of word production with a special emphasis on the semantic representation adopted by each model and their views on lexical processing. Some theories of word production were designed based on or to explain data examining effects of semantic context in relatively complex experimental word production tasks, which, arguably, do not directly resemble natural word production. These tasks include the Picture-Word Interference paradigm, where participants name a picture while ignoring a (usually written) distractor word and Blocked Cyclic Naming, where participants name pictures in small blocks of semantically homogeneous (i.e., from the same semantic category or context) or heterogeneous (i.e., unrelated) items. Yet, if any of the theories reviewed below can truly claim to provide an accurate representation of the process of word production, it also has to be able to account for data from simple tasks, such as 'simple' picture naming, as was used in this thesis.

The main focus of the aforementioned tasks and of theories of word production is to understand and describe lexical processing. From behavioural data we know that processing speed and naming accuracy can vary drastically, which suggests that semantic and lexical processing can be influenced by external (i.e., experimental paradigms) or internal (item-inherent word characteristics; more on this in the section "Semantic variables") factors. This *facilitation* or *interference* of performance can cause responses to be faster or slower as well as more or less accurate. Experimental findings of *semantic facilitation* have mostly been associated with processes at the semantic level and *semantic interference* with processes at the lexical level (e.g., Abdel Rahman & Melinger, 2009; but this also depends on the specific model of word production). In contrast to semantic interference (e.g., Rose & Abdel Rahman, 2017; Vieth et al., 2014a), semantic facilitation in word production has rarely been the focus of experimental investigations (but see e.g., Python et al., 2018a). Similarly, accounting for semantic interference is the focus of word production theories, while semantic facilitation is often neglected.

Theories with competitive lexical selection with a focus on semantic and lexical processing

Word-form Encoding by Activation and VERification (WEAVER++; Levelt et al., 1999)

WEAVER++ proposes three strata (lexical concepts, lemmas, and word-forms) with activation spreading in a feed-forward manner between them. Importantly, the model does not assume inhibitory links within or between strata; however, there is competition at each level of the model to select only one candidate for further processing.

When naming a picture, first the target's lexical concept node is activated. In WEAVER++, word meaning is non-decomposed and representations consist of undivided wholes with links to related concepts. When accessing a target lexical concept ('cat'), activation spreads to concepts related in meaning ('dog', 'tiger') within the semantic level, via labelled links that reflect different relationships between lexical concept nodes. All activated lexical concepts send activation to their corresponding lemmas. The connection between concepts and lemmas is bidirectional, allowing for activation to spread back to the lexical concepts. Importantly, selection between co-activated lemma nodes is competitive and achieved via the Luce choice rule. Subsequently, word-form encoding takes place, which involves the morpho-phonological representation of the selected lexical representation being retrieved and phonetic encoding taking place. Hence, while semantic interference is located at the lexical level, the locus and mechanism of semantic facilitation is unspecified in WEAVER++.

Swinging Lexical Network (Abdel Rahman & Melinger, 2009, 2019)

The Swinging Lexical Network Account was developed to explain both facilitatory and inhibitory effects observed in semantic context manipulation paradigms (Picture-Word Interference, Blocked Cyclic Naming). This theory proposes holistic lexical concept nodes, and that activation spreads bidirectionally between them and between the semantic and the lexical level. The general effect of spreading activation between lexical concepts is facilitation (*conceptual priming*): Mutually related lexical concepts enhance each other's activation through bidirectional links. Importantly, this spread of activation is thought to be flexible and to depend on the specific context; for example, it can be based on associative relations and allows the formulation of ad hoc categories (e.g., fishing trip). However, if this spread of activation results in activation at the lexical level of a lexical cohort of sufficient size and activation strength, competition between co-activated representations can result in a net inhibitory effect of the semantic context, which outweighs facilitation at the conceptual level. Similar to WEAVER++, competition in the Swinging Lexical Network Account is implemented based on the Luce ratio.

Howard, Nickels, Coltheart, and Cole-Virtue, 2006

Howard et al. (2006) developed a set of principles to explain effects of cumulative semantic interference, where word production gets increasingly harder the more words from the target's semantic category were previously named in the experiment. This account proposes that semantic representations related to the target get co-activated via spreading activation between the concepts or shared semantic features (*shared activation;* importantly, Howard et al. (2006) modelled both a decomposed and a non-decomposed semantic system with comparable conclusions regarding the necessary characteristics of the model). The co-activated concepts in turn activate their respective lexical representations, which delay word production (*competition*) due to lateral inhibition between the co-activated lexical representations. Hence, responses are slower, the higher the activation of the target word's competitors. Finally, previously selected (named) representations are more easily activated subsequently, due to strengthened conceptual-lexical connections (*priming*).

Theories with non-competitive lexical selection with a focus on semantic and lexical processing

Interactive Activation Model (Dell, 1986, 1988; Dell et al., 1997)

In contrast to the other theories reviewed here, which were primarily developed based on chronometric data, Dell's model was built on speech errors observed in naming. It is a two-step model with semantic activation spreading from activated semantic features to corresponding lexical representations and then to phonemes. All connections between levels of the model are bi-directional, with information spreading in a feedforward (cascading) and feedback manner. The strength of connections is thought to be a product of recent experience and learning (although this is not implemented) and they are all excitatory, with no inhibitory connections. Importantly, this model conceives of word meaning as being decomposed into semantic features, which, upon activation, activate their associated lexical representations. Feedback from the lexical level to the semantic level can cause semantic features of semantically related non-target words to be co-activated. Lexical selection is achieved by activation converging from both semantic and phonological processing and is thought to be non-competitive, with the most active lexical candidate being selected after a certain number of activation cycles (time steps). Hence, bi-directional links between the semantic and lexical levels in this model architecture can explain effects of semantic facilitation, while it does not implement mechanisms for semantic competition.

Incremental Learning Model (Oppenheim et al., 2010)

Similar to Howard et al. (2006), Oppenheim et al.'s (2010) incremental learning model was created to explain effects of cumulative semantic interference. Semantic representations are decomposed into semantic features and activate the lexical representations that contain these semantic features, which results in multiple lexical representations getting activated at the lexical level. However, selection at the lexical level is non-competitive (note that Oppenheim et al. also modelled a variant of this account with competitive lexical selection) but rather aided by a booster mechanism, which amplifies the activation levels of all active lexical representations to increase activation differences between the co-activated representations. This continues until activation of a lexical representation reaches an absolute threshold and is selected for production. Consequently, any co-activated competitors do not influence selection times. Activation boosting continues until a certain number of boosts have been reached. If, at that point, no lexical representation exceeds the activation threshold, no lexical selection takes place, resulting in an omission. Specific to this theoretical account is, moreover, that after selection, the semantic-to-lexical connections between the semantic representation and the co-activated lexical representations are adjusted by an implicit error-based learning mechanism: The connection between the selected lexical representation and its semantic representation is strengthened, which facilitates future retrieval of the same word. In contrast, connections between co-activated but unselected representations and their semantic representations are weakened, which inhibits subsequent selection of these lexical representations. Hence, under this theoretical account, non-competitive semantic interference takes place at a pre-lexical level and semantic facilitation is argued to be associated with conscious planning processes that go beyond a model of word production.

Ballistic Model of Lexical Access (Mahon & Navarrete, 2016; Navarrete et al., 2014)

The ballistic model assumes that holistic lexical concepts are activated when the system is in a communicative intentional state. The selection of a lexical concept entails the lexicalisation of its semantic representation. Hence, the lexical level is a direct reflection of what happened at the semantic level. The most highly activated lexical representation will be retrieved, without competition. Consequently, the lexical level is not a 'decision point', but decisions are taken at the pre- and post-lexical levels. Therefore, lexical processing itself cannot be slowed, however, it can be speeded by semantically related contexts via semantic priming. In the presence of an external non-target distractor, like in the Picture-Word Interference paradigm, inhibition from the distractor is thought to arise from a conflict between the target word and the distractor during post-lexical processing at the output buffer *(Response Exclusion Hypothesis*, Mahon et al., 2007). This process of response exclusion was described as being sensitive to coarse semantic characteristics of the target word (such as its semantic category), which facilitates the removal of non-target representations from consideration.

If, however, no non-target distractor is present in the paradigm, as in the Blocked-Cyclic Naming paradigm, the authors place competitive processes in semantic-to-lexical connections (similar to Oppenheim et al., 2010). Both incremental strengthening of connections between the semantic representation and the selected lexical representation as well as incremental weakening of these connections to non-selected lexical representations are proposed to take place after lexical retrieval. In addition, within one experimental block in the Blocked-Cyclic Naming paradigm, the target's lexical concept is primed from semantically related items, resulting in facilitation via a spread of activation between semantically related lexical concepts.

Taken together, the models of word production reviewed above assume that picture naming is semantically mediated and therefore include a level that is concerned with representation and processing of word meaning. In addition, all models (with the exception of the Interactive Activation Model; e.g., Dell, 1986) predict some sort of influence from co-activated representations, albeit at different levels and time-points of processing. Hence, semantic activation and spread of activation to other semantic representations (via links between concepts, through feedback from the lexical level, or even connections between semantic features; McRae et al., 1997, 1999) determines the landscape of further processing, especially of lexical processing. The level at which negative interference from co-activated representations takes place is often, but not always, described as the lexical level (lexical selection by competition). In contrast, few models explicitly locate semantic facilitation in the model, however, where discussed, they are associated with spreading activation and feedback processes at the semantic level (e.g., Abdel Rahman & Melinger, 2009).

Importantly, all accounts of inhibition in current word production theories are related to contextual influences either from distractors that are presented simultaneous to the target word (e.g., Levelt et al., 1999) or previously named words (e.g., Howard et al., 2006; Oppenheim et al., 2010). A notable exception is the updated theory by Abdel Rahman and Melinger (2019), that also considers mechanisms of semantic co-activation without context manipulations and also accounts for effects of interference and facilitation that are limited to processing of the target word itself—influences from endogenous word characteristics. However, it is still largely unclear which word characteristics may be influential during processing and what the mechanisms underlying their effects may be. Hence, more fine-grained analyses of aspects of semantics and their influences on word production are necessary, which is the aim of this thesis.

Semantic variables

Standard, or simple, picture naming (I call this 'simple' as it only involves presenting one picture at a time without showing potentially distracting information along with the target picture or explicitly manipulating the preceding context) allows one to study processing of words without the need to incorporate influences from other target or non-target (distractor) words in the experiment. This is made possible when exploring effects of item-inherent word properties (psycholinguistic variables) on word production. Importantly, the investigation of effects of item-inherent word property variations does not require any manipulation of the context a word is presented in and is consequently unaffected by confounding influences associated with the different context manipulation paradigms (e.g., working memory, executive control functions, response strategies, e.g., Fieder et al., 2019).

There has been considerable research indicating that the different levels of word production are influenced by several item-inherent psycholinguistic variables, such as, for example, target familiarity or frequency of occurrence. This influence is reflected in a relationship between the accuracy of a response and/or the latency of correct naming and the respective variable (e.g., Alario et al., 2004; Nickels, 1997; Nickels & Howard, 1994; Strijkers et al., 2010; see Perret & Bonin, 2018, for a meta-analysis and review). However, models of word production do not always specify their predictions regarding effects of item-inherent properties (but see Levelt et al., 1999, who explicitly localise the effect of frequency at the level of accessing word forms). This may be the case because it has not fully been established which variables influence performance and where in the word production process these effects may arise. Therefore, these item-inherent variables, their mechanisms, and functional loci in the word production process need to be more thoroughly investigated to determine if they have reliable effects on word production. If they do, models of word production need to explicitly account for any such effects.

What can semantic variables tell us?

Computationally oriented research into semantic representations has led to the proposal of a number of semantic word characteristics that describe aspects of the semantic representation of the target word (e.g., McRae et al., 1997). I call these item-inherent word characteristics *semantic variables*[1]. Semantic variables formalise different aspects of the distribution and activation of semantic information of a target word and its relationship to other words using mathematical and statistical methods. These variables have so far mostly been tested in the context of input tasks, such as semantic categorisation task or feature verification (e.g., Fujihara et al., 1998; McRae et al., 1997; Mirman & Magnuson, 2008; Randall et al., 2004), where the focus of the investigation is semantic processing itself.

---

[1] Other groups of researchers (e.g., Professor Lorraine Tyler's group at University of Cambridge) have termed them *conceptual structure statistics*.

In contrast, and despite the central role of semantic information in word production, so far, relatively little research has targeted how semantic variables may affect and modulate processes of word production. However, studying effects of semantic variables in word production is particularly interesting as they may affect both semantic and lexical processing: While effects may originate at the semantic level, they may have dramatic consequences for the target word's activation at the lexical level as well as for the size and strength of activation of a co-activated lexical cohort. Hence, even without manipulating the context a target picture appears in, thoroughly investigating effects of semantic variables can help us to better understand processes of word production and the activation environment in which word production takes place. Ultimately, this may allow to ask theoretical questions about semantic and lexical processing, which can help to adjudicate between different theories of word production.

Semantic variables that have received considerable previous attention are imageability, concreteness, and animacy. Imageability is the ease of creating a mental image of a concept and concreteness the accessibility of sensory experience related to a target concept. The two measures are highly correlated and are usually obtained by participant ratings. Participants with and without aphasia have been found to respond slower and/or less accurately for words of lower imageability and/or concreteness in several tasks and modalities (e.g., participants with aphasia: oral reading, Berndt et al., 2005; picture naming, Nickels and Howard, 1995; unimpaired participants: picture naming, Alario et al., 2004; Ellis & Morrison, 1998; lexical decision, Balota et al., 2004; semantic classification, Yap et al., 2012; word reading: Strain et al., 1995). In participants with aphasia, imageability has also been found to predict the occurrence of semantic errors (e.g., Nickels, 1995). Moreover, it has been shown that event-related brain potentials (ERPs) differed when processing words depending on their imageability/ concreteness (e.g., Amsel & Cree, 2013; Binder et al., 2005; Pexman et al., 2007b; Swaab et al., 2002).

In addition, animate and inanimate objects have been suggested to differ in their semantic structure (i.e., suggestions that animate and inanimate concepts differ in the composition of their semantic features, intercorrelations among their features, and their distinctiveness; e.g., Gonnerman et al., 1997; McRae et al., 1997). Effects of animacy have been found, for example, in picture naming of neurotypical speeded participants (slower and less accurate responses to living things, Hodgson & Lambon Ralph, 2008), and there is a substantial body of literature investigating category-specific deficits in participants with semantic impairments (e.g., Best et al., 2006; Devlin et al., 1998; Garrard et al., 2001; Rico Duarte et al., 2009; Tyler & Moss, 2001).

More recently, the interest in investigating effects of a range of semantic variables on word production has increased and the topic has received more attention (e.g., Fieder et al., 2019; Hameau et al., 2019; Rabovsky et al., 2016). Yet, much of this new knowledge has so far not explicitly been integrated into theories of word production (Vinson et al., 2013). Current models of word production are, as outlined earlier, generally underspecified regarding semantic organisation and particularly regarding effects of item-inherent semantic variables. While the characteristics (e.g., lexical competition) of some of the models allow us to tentatively interpret effects of semantic variables in the context of these models, others may require previously unspecified post-hoc assumptions about processing to be able to account for effects of semantic variables, while other effects may be incompatible with certain theoretical frameworks altogether. Hence, investigating semantic variables and their effects on word production may, on the long run, allow us to adjudicate between models of word production.

Operationalising semantic knowledge and semantic variables

The representation of a specific word in semantic memory comprises different types of information, including facts about the concept, perceptual information, and general world knowledge. This allows us to connect otherwise meaningless lexical units (e.g., words) with complex meaning representations, to recognise objects around us, "and to interact with the world in a knowledge-based manner" (Jones et al., 2015, p. 233). Yet, despite its central role in human behaviour, the exact way semantic information is learned and stored, and the type of information activated and retrieved in semantic memory during cognitive tasks has not yet been fully established. However, the hypothesised type of information stored about a word and the way it is presumed to be connected to other words has direct consequences when attempting to operationalise semantic representations or their relations as semantic variables.

One way to capture semantic knowledge is to explicitly ask participants to rate stimuli on certain aspects of the target's meaning, most commonly, on Likert scales. Ratings have, for example, been used to estimate the number of semantic neighbours (competitors) a word has (e.g., Bormann, 2011; Bormann et al., 2008), or its typicality within its semantic category (e.g., Rossiter & Best, 2013). Research using this rating-based approach has resulted in important insights, however, it is impossible to know on what basis participants rate the stimuli. For example, it could be that when rating the ease of conjuring an image for a word (imageability), other variables, like word frequency or concept familiarity, influence this rating process.

More recent developments, particularly in computational linguistics, try to determine aspects of meaning more objectively with the help of large-scale text corpora or databases (context-based and association-based approaches below). These go hand in hand with theoretical proposals of the representation and processing of semantic information in semantic memory. In the next sections, I outline the proposals regarding semantics that have been utilised to operationalise semantic variables for word production research (e.g., Hameau et al., 2019; Kittredge et al., 2007a; Rose & Abdel Rahman, 2016) and introduce how semantic variables are derived in the context of the different proposals. Finally, I explain the rationale behind selecting only feature-based semantic variables for this thesis.

Importantly, these different proposals, with exception of the feature-based account, are at odds with the simplified implementation of semantics employed by current word production models (i.e., holistic or feature-based semantic representations). Moreover, they entail drastic differences in the distribution of semantic activation and in the composition of the cohort assumed to be co-activated during processing (see Table 2 in Hameau et al., 2019).

Context-based approaches

Like classical holistic theories of semantic representation, context-based accounts propose that the meaning of a word is represented through its relationship to other words (Vigliocco & Vinson, 2007). What I refer to as context-based models of semantic representations are models of the nature and structure of semantic knowledge (see Günther et al., 2019, for review) also known as distributional, corpus-based, semantic space, or co-occurrence models (Jones et al., 2015). The general idea underlying this class of approaches is that word knowledge is learned from the linguistic environment, with the lexical context a word occurs in being used for the computation of its meaning. Hence, semantic information in context-based approaches is derived from the statistical distribution of words, however, the particular learning mechanisms used to build semantic representations vary depending on the specific model.

Two popular context-based accounts that assume different cognitive mechanisms are Latent Semantic Analysis (LSA, a latent inference model; e.g., Landauer, 2006; Landauer et al., 1998; Landauer & Dumais, 1997) and Hyperspace Analogue to Language (HAL, a passive co-occurrence model; Lund & Burgess, 1996) (for in-depth reviews see e.g., Bullinaria & Levy, 2007; Jones et al., 2015; Riordan & Jones, 2011). They use vectors of word production frequencies in large corpora of natural language (LSA) or in a moving window (HAL) resulting in a similarity matrix that allows the calculation of various measures, for example a measure of semantic neighbourhood density (e.g., Hameau et al., 2019; Kittredge et al., 2007a) or semantic similarity (e.g., Vigliocco et al., 2004). A drawback of such context-based approaches is that their semantic relations emphasise associative and thematic relationships (Hameau et al., 2019; Reilly & Desai, 2017) as word meaning that is acquired through experience with the world via our senses is neglected (Vinson et al., 2013).

Association-based approaches

A different approach to semantic cognition is in human word association data. For example, De Deyne et al. (2017) proposed word knowledge to be represented by a network of word nodes that are connected via links reflecting associative relations between two nodes. These large-scale semantic networks are based on word association data, which is usually collected in free association tasks.

Participants in experiments generating free association norms produce free associations in response to given target words, which allows a characterisation of the concepts that are related to a given target word (Mirman & Magnuson, 2008). Different groups have collected association data on a large-scale (e.g., University of South Florida free association norms: Nelson et al., 2004; Small Word of Words; De Deyne et al., 2019) to describe semantic knowledge and relationships between concepts. These have allowed the calculation of several measures such as number of associates (Hameau et al., 2019; Pexman et al., 2007a; Rabovsky et al., 2012) or the strength of the first associate (Griffiths & Steyvers, 2003). By now, a body of literature (e.g., Melinger & Abdel Rahman, 2013; Rose & Abdel Rahman, 2016) has proposed that semantic influences on word production are not restricted to categorical relationships but that associatively related representations are also important for word production.

Feature-based approaches

Among different accounts of the structure of semantic memory (see Jones et al., 2015; McRae & Jones, 2013, for reviews), the featural view is the most prevalent (Murphy, 2002). Individual features are understood to be verbalised proxies for the knowledge actually underlying word meaning (McRae, 2004; Vinson et al., 2013). In other words, features are held to be the basic components of the meaning of a word and each feature represents a certain semantic property of the given concept (such as in the example 'cat' given earlier: *purrs, has fur, has whiskers, has a tail*, etc.). The idea that complex concepts can be decomposed into more basic elements has been criticised in the past. Important points of discussion in this debate are, for example, the impossibility of identifying defining features for all meanings, whether word meaning is analysable as smaller components at all, as well as the *hyponym problem* (i.e., if all features of 'animal' are part of 'cat', why do we not usually erroneously say "animal" instead of "cat") (e.g., J. A. Fodor et al., 1980; J. D. Fodor et al., 1975; J. A. Fodor, 1976). However, going into more detail on this debate is beyond the scope of this General Introduction and thesis (but see e.g., Levelt, 1989; Roelofs, 1997, for reviews and discussions in the context of theories of word production).

According to decomposed, feature-based theories of conceptual knowledge (Cree et al., 1999; Farah & McClelland, 1991; Masson, 1995; McRae et al., 1997; Tyler et al., 2000), a concept's meaning is represented as the pattern of activation across such semantic features. McRae and colleagues (e.g., McRae et al., 1997, 1999) advanced the idea that semantic memory includes implicit knowledge about feature relationships and suggested feature-feature and feature-concept statistics to be embedded in the semantic representation of concepts and to play important roles during word processing (some of these statistics are the focus of this thesis and are further described below).

Case for focusing on feature-based semantic variables

In this thesis, I focus on semantic variables that can be operationalised based on conceptual knowledge that is represented as semantic features. Various semantic feature norm databases using participant-generated features have been developed by multiple groups (e.g., Devereux et al., 2014; McRae et al., 2005; Rosch & Mervis, 1975; Vinson & Vigliocco, 2008) to capture word meaning and operationalise conceptual knowledge as semantic features. In feature-generation tasks, participants are usually asked to list characteristic features that describe concepts. Here, I used McRae et al.'s (2005) feature norm database, which is a corpus of 541 basic-level concepts. For each of the concepts, 30 participants were asked to generate semantic features to define and describe the concept. Participants were provided with 10 empty lines and were instructed to generate different types of features (e.g., perceptual, functional, encyclopaedic). This approach is thought to make participants systematically analyse their semantic knowledge about a concept and to reveal the dimensions of meaning they considered psychologically salient (Vinson et al., 2013). Semantic feature norms are therefore thought to provide a window into important aspects of word meaning and to be a proxy of conceptual knowledge (e.g., Vigliocco & Vinson, 2007)—they provide us with a model of semantic representation of single words. Hence, any effects of feature-based semantic variables on word production processes may reflect underlying principles of semantic organisation and its consequences for lexical processing and therefore have to be accounted for by theories of word production. Moreover, as noted by Clarke and Tyler (2015), feature norms and variables derived from them are thought to share properties with the neural underpinnings of conceptual representation and processing in the brain, "although semantic features are not claimed to be the neural units of meaning" (Clarke & Tyler, 2015, p. 678).

Critics of the use of feature norms have raised concerns that they are not direct representations of the true properties likely underlying word meaning. For example, they comprise only features that can be verbalised and have been argued to suffer from a systematic underrepresentation of highly frequent, shared features (e.g., *breathes*; e.g., Vinson et al., 2013). However, despite their limitations I believe that feature norms are the best currently available method for specifying conceptual content. Measures based on semantic features are reproducible, transparent, allow for adaptation, extension, and replication by other researchers. Moreover, feature-based semantic variables represent a relatively objective way to operationalise conceptual knowledge: In contrast to rating approaches, participants are not asked to directly estimate a certain aspect of the semantic representation (e.g., concept typicality within its semantic category, e.g., Rossiter and Best, 2013; number of semantic competitors, e.g., Bormann, 2011) and are hence naïve to the specific aspect of meaning to be measured. Rather, the feature norms can be used as a source of information from which different measures can be more objectively determined using mathematical and statistical methods. Moreover, current theories of word production mostly do not assume association-based or context-based relations (but see Abdel Rahman & Melinger, 2019, who also propose associative relationships between lexical concepts), but specify feature-based relations (i.e., feature-based semantics or connections between lexical concepts via shared properties), which facilitates interpretation of any effects of feature-based semantic variables in the context of these theories.

In the absence of clear pointers from the literature regarding which semantic dimensions and relationships are encoded in our language systems, I aim to provide a thorough investigation of one approach to operationalising semantic knowledge and focus on semantic variables that can be operationalised based on semantic features. This choice was also reinforced by a study by Hameau and colleagues, who compared different approaches to operationalising semantic neighbourhood density (i.e., feature-based, association-based, context-based) and their effects on picture naming and found that the feature-based measure[2] best captured variability in the data (Hameau et al., 2019). Moreover, I attempted to avoid the use of multiple databases (e.g., McRae et al.'s (2005) feature database and Nelson et al.'s (2004) association database), as they each contain different words, which would have decreased the number of items with full item information. Subsequent research may focus on effects of semantic variables derived from context- and association-based approaches and conduct similarly systematic and thorough explorations of these domains before assessing effects of variables across methods of operationalisation.

---

[2] Note that Hameau et al.'s (2019) feature-based measure combined the number of feature-based near semantic neighbours and the number of rated competitors.

Importantly, by utilising a semantic feature norm database to calculate semantic variables I do not intend to take a position in the debate regarding whether semantic representations in word production theories are decomposed or holistic. One could argue that the type of semantic representation adopted by a word production theory has direct consequences for the proposed flow of information and for which semantic relationships are predicted to influence word production. However, I do not wish to over-interpret any given effects of feature-based variables as incompatible with holistic architectures, as they could also reflect semantic relationships as captured in non-decomposed theories of semantics (e.g., Collins & Loftus, 1975; Levelt et al., 1999; Roelofs, 1992) (see e.g., Paper 1 in Chapter 2, for a discussion). In my interpretations of significant effects, I therefore instead focus on explaining how the respective semantic variables might operate at the semantic and particularly the lexical level of word production and discuss the theoretical characteristics necessary to account for any effects.

To further narrow down the scope of this thesis, I focus on feature-based semantic variables that describe some aspect of the meaning of the entire concept. In contrast, I do not touch on the contribution of specific types of features (e.g., sensory or functional) to the knowledge of words and the discussions about selective impairments for certain types of knowledge (e.g., debate about domain specific (living/nonliving) impairments).

Throughout this thesis, I have used McRae et al.'s (2005) feature-norm database to calculate semantic variables. More specifically, the semantic variables that this work focusses on are number of semantic features, intercorrelational density, number of near semantic neighbours, typicality, semantic similarity, and distinctiveness.

Previous investigations of effects of feature-based semantic variables on word processing

The aforementioned semantic variables selected for investigation in this thesis were retrieved from previous work into effects of semantic variables on word production. A thorough review of the literature investigating effects of the six feature-based semantic variables in word production and their theoretical interpretations is given in Chapter 2 (Paper 1; participants with aphasia) and Chapter 3

(Paper 2; neurotypical participants). Below I briefly introduce each of the variables and provide an overview with examples of previous research into their effects outside word production.

Number of semantic features

Number of semantic features can be thought to indicate the richness of the semantic representation of the target word. This measure is almost always retrieved from feature norm databases (e.g., Devereux et al., 2014; McRae et al., 2005). For example, in the McRae et al. database, the concept 'cat' is represented by the features *has fur, an animal, a pet, eats, has whiskers, meows, purrs, has four legs, has legs, has a tail, has claws, is domestic, a baby is a kitten, a feline, eats mice, has paws, is independent, a mammal, kills,* and *has eyes*.

The effect of number of semantic features generally described in the literature, which holds across modalities, is that concepts consisting of more semantic features are responded to faster than those consisting of fewer features (e.g., Grondin et al., 2009; Pexman et al., 2002, 2003, 2008; Yap et al., 2011). Effects of number of semantic features have also been studied across receptive semantic and lexical tasks using electroencephalography (EEG). For example, Amsel (2011), Rabovsky et al. (2012), and Kounios et al. (2009) reported differences in ERPs when comparing processing of words with high or low numbers of semantic features in a variety of tasks (i.e., delayed imagery judgement, lexical decision, and semantical relatedness judgement, respectively). Differences in the number of semantic features associated with a word have also been suggested to underlie effects of concreteness (e.g., Plaut & Shallice, 1993), where concrete words are responded to more quickly than abstract words (Binder et al., 2005; Strain et al., 1995). However, this account has been disputed by for example Amsel and Cree (2013) and Kounios et al. (2009) who presented data that suggested that concreteness and number of semantic features are separate measures.

Related measures capturing numbers of particular types of features (e.g., perceptual or functional features; e.g., Rico Duarte & Robert, 2014) or the number of distinguishing or shared features (see also section "Distinctiveness" below; e.g., Grondin et al., 2009; Miozzo et al., 2015) have also been shown to influence performance, however, in this thesis, I focus on the overall number of semantic features measure (previous behavioural findings from word production studies are presented in detail in Chapters 2 and 3 and ERP findings in Chapter 4).

Intercorrelational density

The idea behind a measure of featural intercorrelation is that different semantic features tend to occur together. For example, the features *has fur* and *has four legs* tend to occur together across concepts (e.g., 'cat', 'dog', 'wolf', 'caribou', 'cougar'). Stronger correlations between features are argued to allow for greater mutual co-activation in clusters of intercorrelated features, due to, for example, bidirectional feature-feature connections (e.g., Cree et al., 1999; McRae et al., 1997, 1999). Hence, the degree to which a specific feature (e.g., *has fur*) is correlated with the other features of a concept (e.g., 'cat') (termed *intercorrelational strength* by McRae et al., 1997, 1999) determines the level of activation of that particular feature but also of the other features of the concept (e.g., *has four legs, has whiskers, has a tail,* etc.) as features in a cluster of intercorrelated features boost each other's activity. This affects processing, with more strongly correlated features speeding up activation and thus decreasing processing times. Strength of the correlation between the feature and a concept has been shown to predict response times in feature verification tasks where participants were asked to verify features as true or false of a concept (McRae et al., 1997, 1999). For example, *is hunted* is more strongly correlated with the other features of 'deer' than of 'duck' and was therefore faster verified than for the concept 'deer' than the concept 'duck' (for similar findings see also e.g., Garrard et al., 2005; Randall et al., 2004; Taylor et al., 2004). Differences in feature correlations in different semantic categories have also been suggested as one of the reason for domain-specific impairments in participants with language disorders (e.g., Devlin et al., 1998; Gonnerman et al., 1997).

In this thesis, however, I was interested in semantic variables that capture an aspect of semantic information of the whole concept (i.e., concept-specific variables) and not of single features in relation to concepts. McRae et al. (1997) suggested the measure *intercorrelational density* as a means of capturing the relationship between all the semantic features of a concept. It indexes the degree of feature-feature correlations among the features of a concept. Higher intercorrelational density has been shown to lead to faster responses in a domain decision task (Taylor et al., 2012; note that they called their measure *correlational strength*) and faster convergence of a computational model (McRae et al., 1997). Importantly though, for words with high intercorrelational density, McRae et al. also reported that some features were co-activated that are not really part of the concept but are strongly correlated with the actual features of the concept (e.g., bird-features like *beak* for the concept 'jet' via *wings* and *flies*). Intercorrelational density has been previously investigated in a few word production studies (Clarke et al., 2013; Rabovsky et al., 2016, 2021; Taylor et al., 2012), some of which also collected evoked responses, and which are reviewed in Chapters 3 and 4.

Number of near semantic neighbours

Semantic neighbours of a target word are other words that share part of their semantic information with the target. Different ways of capturing the semantic neighbourhood of words have been implemented using the different approaches to operationalising semantic knowledge described above: feature-based (e.g., Mirman, 2011), association-based (e.g., Hameau et al., 2019), and context-based (e.g., Kittredge et al., 2007b) neighbours, as well as ratings of the number of category coordinates (e.g., Bormann, 2011). Number of semantic neighbours has been reported to affect processing in a variety of tasks, such as visual lexical decision in unimpaired speakers (e.g., Buchanan et al., 2001; Pexman et al., 2008) and word production in people with aphasia (e.g., Blanken et al., 2002; Bormann, 2011; Bormann et al., 2008), while in unimpaired participants word production seems to mostly be unaffected by number of semantic neighbours (Bormann, 2011; Hameau et al., 2019).

Looking more closely at the feature-based semantic neighbourhood variable used in the past, it was actually a measure of number of *near* semantic neighbours, which are words that share a relatively large proportion of their semantic features with the target word (feature vector cosine similarity > 0.5; Fieder et al., 2019; or > 0.4, Mirman, 2011, respectively). Mirman and Magnuson (2008; Experiment 1) reported that number of near semantic neighbours affected performance in lexical selection and category discrimination tasks. They also found longer response latencies for words with many near semantic neighbours in contrast to words with few near neighbours in a concreteness judgement task (Experiment 2). Moreover, feature-based number of near semantic neighbours has been the focus of a number of studies investigating word production in participants with and without aphasia, which are reviewed in Chapters 2 and 3.

In addition to number of near semantic neighbours, number of *distant* semantic neighbours has been used as another semantic variable in some studies. In contrast to near semantic neighbours, distant semantic neighbours are words that share very little semantic information with the target word (e.g., cosine of < .25 and > 0 in Mirman and Magnuson, 2008), but are much more abundant. In a semantic categorisation task, Mirman and Magnuson reported that the presence of many distant neighbours tended to have facilitatory effects on processing. This measure has also been used in some word production studies (e.g., Fieder et al., 2019; Mirman, 2011). In this thesis, number of distant semantic features was not included as one of the semantic variables of interest. Justification for this decision is provided in Chapter 2.

Typicality

Typicality is one of the most researched semantic variables. It indexes the items' representativeness of its semantic category: 'robin' is a representative, a typical, bird, while 'ostrich' is an atypical bird, due to 'robin' having many prototypical features of the category 'bird' (*it is small, flies, has wings, a beak*, etc., e.g., Rossiter & Best, 2013). To operationalise typicality, researchers have most often used ratings in which participants rated the target's typicality in its semantic category (e.g., Rossiter & Best, 2013) or other measures that have been argued to capture typicality (e.g., frequency of instantiation, e.g., Barsalou, 1985; category potency, e.g., Battig & Montague, 1969; dominance of the category superordinate, e.g., Ashcraft, 1978). However, there are also more objective approaches to capture typicality based on semantic feature norms (e.g., family resemblance score, Rosch & Mervis, 1975).

Barsalou (1985) claimed that no other variable is as important for performance on a wide range of tasks in participants with and without language impairments. Typicality generally facilitates performance such that responses are faster and/or more accurate for items with higher typicality in contrast to more atypical items. The typicality effect has been found for unimpaired participants in semantic classification and category-verification tasks. For instance, participants have been reported to be faster in deciding that a 'robin' is a bird versus that an 'ostrich' is a bird (e.g., Fujihara et al., 1998; Holmes & Ellis, 2006; Larochelle et al., 2000; McCloskey & Glucksberg, 1979; Morrison & Gibbons, 2006; Rips et al., 1973; Rosch & Mervis, 1975; Sandberg et al., 2012) and typicality has been shown to affect performance in animacy decision tasks (e.g., Morrison & Gibbons, 2006; Räling et al., 2016), category fluency (e.g., Hernández-Muñoz et al., 2006), and in a feature verification task (e.g., Ashcraft, 1978). In addition to offline tasks, typicality has also been found to affect online semantic processing: Increased N400 amplitudes, which are associated with semantic processing, have been reported for atypical words in category-verification or semantic categorisation tasks (e.g., Fujihara et al., 1998; Heinze et al., 1998; Monetta et al., 2003; Räling et al., 2015). Typicality has also been found to affect picture naming accuracy and speed in unimpaired participants (e.g., Dell'Acqua et al., 2000; Holmes & Ellis, 2006) and these findings are reviewed in Chapter 3.

Effects of typicality have also been studied in participants with language impairments. In participants with aphasia, typicality has been most commonly examined using categorisation or semantic decision tasks (e.g., Grober et al., 1980; Kiran et al., 2007; Kiran & Thompson, 2003; Riley & Thompson, 2010; Sandberg et al., 2012; Stanczak et al., 2006). Effects of typicality in aphasia have also been revealed in tasks that require verbal output, such as picture naming (Laiacona et al., 2001; Rossiter & Best, 2013; see Chapter 2, for a comprehensive review) and category-exemplar generation (e.g., Grossman, 1981; Hough & Pierce, 1988). However, there is disagreement about whether effects of typicality in people with aphasia differ depending on their type of aphasia (e.g., Grober et al., 1980; Grossman, 1981; Hough & Pierce, 1988; Kiran & Thompson, 2003). For individuals with Semantic Dementia, a neurodegenerative disorder of semantic knowledge, also known as the semantic variant of primary progressive aphasia (svPPA), that results in fluent progressive aphasia (Gorno-Tempini et al., 2011), better performance for high typicality items over low typicality items have been found in a category selection task (Mayberry et al., 2011; "matching-to-sample") and picture naming (Woollams et al., 2008; Woollams, 2012), which are discussed in more detail in Chapter 2.

Semantic similarity

Semantic similarity indexes the closeness in meaning of the target and other words in an experimental task or the mental lexicon more generally. The semantic similarity of two concepts can be expressed as the cosine similarity of their feature vectors as provided by, for example, the feature database by McRae et al. (2005). Cosine similarity ranges from 0 (no shared semantic features) to 1 (identical feature vectors) and indexes the distance between the feature vectors of two concepts. Higher mean cosine overlap has been found to speed processing in semantic categorisation but not in lexical decision (Mirman & Magnuson, 2008). Moreover, Cree et al. (1999) found graded priming effects as a function of featural overlap between prime and target. Similarly, Mirman and Magnuson (2009) reported a graded semantic competition effect due to differences in semantic similarity between near neighbours, distant neighbours, and non-neighbours using a visual world eye tracking paradigm.

In word production, semantic similarity has mostly been examined in context manipulation tasks that contrast the influence of semantically more similar and semantically less similar distractors on target word processing. Semantically more similar distractors have mostly been found to interfere more strongly with target processing (e.g., Rose et al., 2019; Vieth et al., 2014b; Vigliocco et al., 2004; but see Mahon et al., 2007, for contradictory findings). Moreover, Rose and Abdel Rahman (2017) found that semantic similarity influenced the cumulative semantic interference effect in continuous naming and associated ERP effects. Fieder et al. (2019) report the only previous investigation of item-inherent semantic similarity (i.e., how similar the featural representation of the target is to other representations in the mental lexicon) on picture naming, and this study is described in more detail in Chapters 2 and 3.

Distinctiveness

Featural distinctiveness indexes the degree to which a feature is shared with many other concepts (e.g., *has fur*) or is more uniquely associated with a single or only few concepts (up to 3 in McRae et al., 2005) (e.g., *moos*). Semantic features that are relatively unique to a concept (i.e., distinguishing features) have been found to influence processing in various paradigms, indicating a privileged role in the computation of word meaning. For example, participants verified distinguishing features as being appropriate for a concept faster than shared features in feature verification experiments (e.g., Cree et al., 2006; Randall et al., 2004) and rated them as supporting their ability to name from descriptions more than shared features (Marques, 2005). In line with this, for individuals with aphasia, Mason-Baughman and Wallace (2013a, 2013b) found that knowledge of distinguishing features, as identified in a feature-word sorting task, determined participants' success in distinguishing target nouns from semantically related foils in a spoken to written word matching task. Moreover, it has been suggested that distinguishing and shared features might be lost at different rates during the progression of degenerative conditions like Alzheimer's disease and Semantic Dementia, which has also been discussed as a reason underlying category specific impairments (e.g., Catricalà et al., 2015; Garrard et al., 2005; Gonnerman et al., 1997; Laisney et al., 2011; Moss et al., 1998; Rico Duarte et al., 2009; Rogers et al., 2004; Tyler et al., 2000; Tyler & Moss, 2001; see also Caramazza & Shelton, 1998, for a participant with stroke aphasia). Finally, there is discussion around the role of the distinctiveness of part-whole distractors in the Picture Word Interference paradigm (is the part distractor a distinctive or shared feature of the target?) and its effect on the polarity of semantic effects (e.g., Vieth et al., 2014a).

However, as for intercorrelational density, previous research has also considered a concept-based measure of distinctiveness that captures information regarding the whole concept and not on single semantic features. In contrast to feature distinctiveness, concept distinctiveness measures how special or informative the features of an item are on average, with higher distinctiveness indicating a higher proportion of distinguishing semantic features. In a megastudy analysis, Siew (2020) found that words with greater distinctiveness were acquired earlier in life. Moreover, greater distinctiveness inhibited visual lexical decision performance and facilitated semantic decision for concrete concepts while it inhibited semantic decision for abstract concepts. Concept distinctiveness has previously also been investigated using behavioural data and evoked responses in word production (Clarke et al., 2013; Rabovsky et al., 2016; Taylor et al., 2012), which are reviewed more closely in Chapters 3 and 4.

Taking different perspectives to better understand effects of semantic variables

Overcoming previous shortcomings

The abundance of literature that was briefly discussed in the previous sections indicates that there is a strong interest in using statistical regularities of the semantic structure of concepts (i.e., semantic variables) to better understand semantic and word processing. However, the focus of these investigations has mostly been semantic representations themselves, entailing a concentration on semantic input tasks (e.g., semantic categorisation). In comparison, so far, relatively few studies have tested the effects of semantic variables on output tasks like spoken word production. However, given the strong interconnectivity of semantic and lexical processing in word production, word production is a particularly interesting area in which to study effects of semantic variables. This is because these variables may have significant consequences for the activation environment in which lexical selection, a much debated level of word production, takes place. Hence, knowledge about which semantic variables influence our word production behaviour may be useful to advance and test theories of word production. However, the critical first step is to establish which of the previously suggested semantic variables reliably influence performance. That was the aim of this thesis.

Given the previous studies in this field (described in more detail in Chapters 2 and 3), one might ask why further investigations into *which* semantic variables influence word production are necessary. Previous research into effects of item-inherent variables, such as semantic variables, has often focused on one (or few) variables. However, the findings generated following such an approach may be misleading due to meaning being a multidimensional construct with simultaneously occurring effects of different dimensions of meaning, each of which may explain unique variance (Pexman et al., 2013; Taylor et al., 2012). Therefore, here, in an attempt to identify those variables that reliably influence our behaviour, I follow a more comprehensive approach in which I simultaneously consider several different semantic variables that previous research has identified as potentially influential in word production.

In addition, previous investigations did not always control sufficiently for the psycholinguistic variables that influence word production, nor for variation between participants and items (see

Chapters 2 and 3, for more in-depth discussion). However, these practices might be particularly problematic when trying to establish reliable effects of semantic variables as effects of semantic and other psycholinguistic variables vary between participants. This may cause effects of semantic variables to not reach significance, which is particularly likely when considering effects of naturally occurring semantic variations without maximising differences between items (orthogonal item sets) or using context manipulations.

Studying word production in aphasia

Most of the previous work on semantic variables has investigated word production in unimpaired participants with the aim to better understand word planning and the time course of processes involved in word production in the unimpaired system. However, studying effects of semantic variables in people with aphasia allows us to test how impairments to different components of the word production model may affect processing (Nickels, 1995). Previous work has located effects of semantic variables in semantic and lexical processing. This suggests that people with aphasia with impairments in these locations may be especially affected by the semantic variables (Shallice, 1988). Studying how different semantic variables predict difficulties during semantic and lexical processing in word production in language impaired individuals can facilitate our understanding of semantic representations and processing as well as their influence on lexical retrieval of words in the mental lexicon. Hence, studying effects of semantic variables in people with aphasia can help to better understand the mechanisms underlying these effects on word production.

Previously, few semantic variables have been studied in people with aphasia. While earlier research targeted some of the aforementioned semantic variables (e.g., number of (near) semantic neighbours, typicality) in picture naming in participants with aphasia, other semantic variables have not been studied (e.g., number of semantic features, intercorrelational density). Hence, the experimental investigation reported in Chapter 2 of this thesis is the first to test effects of some semantic variables in people with aphasia.

Speed of processing

As mentioned above, most studies investigating processes during word production have applied context manipulation paradigms where the experimental context a target word occurs in is thought to affect the processing of the target (Picture-Word Interference, Blocked Cyclic Naming) or where the effect of interest builds up across the course of the experiment (Cumulative Semantic Interference). In contrast, in this thesis I was interested in effects of item-inherent variables, which do not require any context manipulations to arise. The idea is that item-inherent variables may affect processing even under the simplest processing conditions and are thus relevant for 'real life' communication situations. Hence, all experimental chapters presented in this thesis used standard picture naming to study effects of the semantic variables.

Effects of some semantic variables have previously been found in standard picture naming in participants with and without aphasia (evidence reviewed in Chapters 2 and 3, respectively). However, some effects have only been found to be significant in neurotypical participants under limited processing conditions in a speeded picture naming paradigm. In speeded picture naming, participants are asked to prioritise naming speed over the accuracy of their response and are required to name the pictures at a time point that is likely before their language system is ready to respond (i.e., most commonly around 600ms with average response latencies in standard picture naming tasks often being much longer, e.g., around 800ms in Valente et al. (2014) or around 900ms in the study presented in Chapter 3). The specific instructions used in the past have varied slightly with participants requested to respond at a certain point using an acoustic countdown (tempo naming; e.g., Fieder et al., 2019; Hodgson & Lambon Ralph, 2008; Kello, 2004; Kello & Plaut, 2000; Mirman, 2011) or before a signal (to 'beat the beep'; deadline naming; e.g., Damian & Dumay, 2007; Gerhand & Barry, 1999; Hodgson & Lambon Ralph, 2008; Kello et al., 2000; Lloyd-Jones & Nettlemill, 2007; Moses et al., 2004; Vitkovitch et al., 1993; Vitkovitch & Humphreys, 1991). Previous work has suggested that the time pressure engages a cognitive control mechanism to shorten the time course of processing (i.e., input gain account, e.g., Kello & Plaut, 2000; Mirman, 2011), which may modulate the sensitivity of processing units to excitatory and inhibitory inputs and consequently the strength of effects of semantic variables (Mirman, 2011).

And indeed, previous studies have found significant effects of number of near semantic neighbours in the context of a speeded picture naming task in unimpaired participants (Fieder et al., 2019; Mirman, 2011), while studies using standard picture naming have failed to find significant effects of that variable (see Hameau et al., 2019; Lampe et al., 2017; and Bormann, 2011, for a similar measure). Hence, in Chapter 5, I investigated this observation more systematically and tested whether naming speed affected the impact of semantic variables on word planning by comparing effects of semantic variables in speeded and standard picture naming.

Brain activity during word planning

Most word production models and interpretations of effects of item-inherent variables are based on behavioural data from language impaired or unimpaired participants. Behavioural measures of naming allow us to hypothesise regarding what might be happening during word planning based on the *final outcome* of word production (i.e., the actual word uttered and the time it took the participant to do so). In contrast, brain imaging techniques like electroencephalography or magnetoencephalography (EEG or MEG, respectively) allow us to access information on brain activity *during* word production—they provide us with a window into online processing. Evoked responses such as event related potentials (ERPs) as measured with EEG or event related fields (ERFs) as measured with MEG hence complement behavioural data, adding an additional dimension of information on the word production process: insight into brain activity during word production and possible differences caused by experimental manipulations. In addition, the high temporal precision of EEG and MEG has been used as a source of information on the time course of experimental effects.

A substantial body of work has examined evoked responses in context manipulation word production paradigms attempting to uncover the effects of semantic interference and facilitation and their time course (Picture-Word Interference, e.g., Aristei et al., 2011; Blackford et al., 2012; Dell'Acqua et al., 2010; Hirschfeld et al., 2008; Piai et al., 2012; Python et al., 2018b; Rose et al., 2019; Blocked Cyclic Naming, e.g., Aristei et al., 2011; Janssen et al., 2011, 2015; Maess et al., 2002; Python et al., 2018a;

Cumulative semantic interference, e.g., Costa et al., 2009; Rose & Abdel Rahman, 2017; see Nozari & Pinet, 2020, for a recent review and discussion). However, evoked responses also allow testing of effects and their time course of item-inherent variables on word production (e.g., word frequency, Laganaro, 2014; Levelt et al., 1998; Piai et al., 2012; Strijkers et al., 2010; name agreement, Cheng et al., 2010; age of acquisition, Laganaro, 2014; Laganaro et al., 2012; Laganaro & Perret, 2011; several variables, Valente et al., 2014).

The high temporal precision of ERP and ERF data has also been used in an attempt to formulate time course estimates of the processes of word production (in addition to spatial information on their neural correlates, Indefrey, 2011; Indefrey & Levelt, 2004). According to these time course estimates, semantic processing spans the first 200ms of word planning and lexical processing occurs between 200 and 275ms post picture onset (see also Strijkers & Costa, 2011, for a confirmation of the initiation of lexical access within 200ms). Lemma retrieval is followed by phonological encoding and articulation commences around 600ms (but see Chapter 4, for discussion of possible issues with this time course).

So far, to the best of my knowledge, only three studies have been conducted that have aimed to understand the neural correlates and time course of effects of item-inherent semantic variables. Clarke et al. (2013) and Miozzo et al. (2015) studied effects of semantic variables using MEG; however, both studies used orthogonal components derived from Principal Component Analyses, which contained various measures that were grouped together. While they reported significant effects of (some of the) semantic components, their approach of combining variables in principal components unfortunately makes it impossible to comment on effects of individual semantic variables. In contrast, Rabovsky et al. (2021) studied effects of number of semantic features and intercorrelational density using EEG and reported that both variables affected semantic and lexical processing for word production.

While these previous studies each investigated effects of at least two semantic variables, previous behavioural research has suggested other variables that also affect performance in word production, whose influence had been untested (and uncontrolled) in these two studies. Consequently, in Chapter 4, I extend this evidence base by using EEG to explore the brain dynamics associated with the effects of six semantic variables in word production.

Thesis overview

Previous research has shown that various word characteristics, item-inherent variables, can influence the ease of word production, which may cause differences in response latency, naming accuracy, types of error produced, and in brain activity while planning a word, as uncovered with EEG and MEG data. However, the influence of such item-inherent variables relating to semantic information about the target word (i.e., semantic variables) on processes of word production are still under-researched. Importantly, previous studies on effects of semantic variables often focused on individual measures but disregarded the fact that if they indeed influence our word production performance, different semantic and other item-inherent psycholinguistic variables would operate simultaneously.

It is the overall aim of this thesis to contribute well-controlled studies to the evidence base relating to effects of semantic variables on word production by thoroughly investigating effects of six feature-based semantic variables: number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness. Using multiple experimental approaches, the studies in this thesis aim to determine *which* feature-based semantic variables reliably influence behaviour and to better understand *how* these variables affect processing in the context of word production models.

In the Discussions of the experimental papers of this thesis, as well as in the General Discussion, I speak to how assumptions around semantic and lexical processing in current models of word production can account for effects of the semantic variables, as these models have to be able to explain experimental findings in order to be considered comprehensive and valid theories of word production. Therefore, this thesis intends to bridge the gap between theoretical and computational approaches to semantic representation and processing and current theories of word production. By exploring which aspects of semantic representation influence word production, I attempt to merge these two schools of thought. This thesis therefore presents exploratory work, which lays the foundations for subsequent confirmatory research, that may utilise the insights of the work presented here to more directly test models of word production or characteristics of its sub-processes (e.g., Is lexical selection a competitive process?). Where current theories are unable to account for effects of semantic variables, the findings of this thesis may be taken as a starting point to modulate and improve models of word production in order for them to be able to explain the presented effects; however, it is beyond the scope of this work to propose a new and improved theory of word production.

This thesis presents four experimental chapters written in journal article format, which are followed by a General Discussion of the findings. In **Chapter 2** (Paper 1) I explore effects of the feature-based semantic variables on word production in a large group of 175 participants with stroke-induced aphasia with varied naming impairments. In a subgroup analysis with 60 participants, I focus on a more homogeneous group of participants with lexical and/or semantic impairments.

Chapter 3 (Paper 2) investigates effects of the same feature-based semantic variables on behavioural measures of word production in 87 unimpaired participants. The same group of participants is again presented in **Chapter 4** (Paper 3), which represents an electrophysiological investigation of the processes of word production. Here, two popular analysis approaches are utilised to test for effects of the semantic variables on the brain activity during word processing: a traditional wave-from analysis as well as a microstate analysis. Chapters 3 and 4 each present multiple statistical analyses that increase in complexity to replicate and extend previous work by Rabovsky et al. (2016, 2021).

In **Chapter 5** (Paper 4), I test whether there are differences in the effects of the semantic variables depending on the processing requirements employed by the experimental paradigm. Specifically, I explore in 80 participants if the effects of semantic variables differ between speeded deadline and standard picture naming, an idea that was inspired by the patterns of findings of previous research into effects of number of near semantic neighbours and semantic similarity on word production, which found significant effects in speeded naming paradigms, but non-significant effects in standard naming.

Finally, **Chapter 6**, the General Discussion, summarises the main findings of each paper and ties together the findings of the experimental chapters. The discussion highlights consistencies and differences between the studies to build a new understanding of effects of feature-based semantic variables on word production. I discuss the overall contribution of the findings of this thesis to the literature and outline both limitations and future research directions.

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*(5), 713–734. https://doi.org/10.1080/01690960802597250

Abdel Rahman, R., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience*, *34*(9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Amsel, B. D. (2011). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia*, *49*(5), 970–983. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2011.01.003

Amsel, B. D., & Cree, G. S. (2013). Semantic richness, concreteness, and object domain: An electrophysiological study. *Canadian Journal of Experimental Psychology*, *67*(2), 117–129. https://doi.org/10.1037/a0029807

Aristei, S., Melinger, A., & Rahman, R. A. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience*, *23*(7), 1567–1586. https://doi.org/10.1162/jocn.2010.21474

Ashcraft, M. H. (1978). Property dominance and typicality effects in property statement verification. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 155–164. https://doi.org/10.1016/S0022-5371(78)90119-6

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, *11*(4), 629–654. https://doi.org/10.1037/0278-7393.11.1-4.629

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80*(3 Pt. 2), 1–46. https://doi.org/10.1037/h0027577

Berndt, R. S., Haendiges, A. N., & Mitchum, C. C. (2005). Orthographic effects in the word substitutions of aphasic patients: An epidemic of right neglect dyslexia? *Brain and Language*, *93*(1), 55–63. https://doi.org/10.1016/j.bandl.2004.06.110

Best, W., Schröder, A., & Herbert, R. (2006). An investigation of a relative impairment in naming non-living items: Theoretical and methodological implications. *Journal of Neurolinguistics*, *19*(2), 96–123. https://doi.org/10.1016/j.jneuroling.2005.09.001

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*(6), 905–917. https://doi.org/10.1162/0898929054021102

Blackford, T., Holcomb, P. J., Grainger, J., & Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition*, *123*(1), 84–99. https://doi.org/10.1016/j.cognition.2011.12.007

Blanken, G., Dittmann, J., & Wallesch, C.-W. (2002). Parallel or serial activation of word forms in speech production? Neurolinguistic evidence from an aphasic patient. *Neuroscience Letters*, *325*(1), 72–74. https://doi.org/10.1016/S0304-3940(02)00227-6

Bormann, T. (2011). The role of lexical-semantic neighborhood in object naming: Implications for models of lexical access. *Frontiers in Psychology*, *2*, 1–11. https://doi.org/10.3389/fpsyg.2011.00127

Bormann, T., Kulke, F., Wallesch, C.-W., & Blanken, G. (2008). Omissions and semantic errors in aphasic naming: Is there a link? *Brain and Language*, *104*(1), 24–32. https://doi.org/10.1016/J.BANDL.2007.02.004

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, *8*(3), 531–544.

https://doi.org/10.3758/BF03196189

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, *39*(3), 510–526. https://doi.org/10.3758/BF03193020

Butterworth, B. (1989). Lexical access in speech production. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108–135). The MIT Press.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*(1), 1–34. https://doi.org/10.1162/089892998563752

Catricalà, E., Della Rosa, P. A., Plebani, V., Perani, D., Garrard, P., & Cappa, S. F. (2015). Semantic feature degradation and naming performance. Evidence from neurodegenerative disorders. *Brain and Language*, *147*, 58–65. https://doi.org/10.1016/j.bandl.2015.05.007

Cheng, X., Schafer, G., & Akyürek, E. G. (2010). Name agreement in picture naming: An ERP study. *International Journal of Psychophysiology*, *76*(3), 130–141. https://doi.org/10.1016/j.ijpsycho.2010.03.003

Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2013). From perception to conception: How meaningful objects are processed over time. *Cerebral Cortex*, *23*(1), 187–197. https://doi.org/10.1093/cercor/bhs002

Clarke, A., & Tyler, L. K. (2015). Understanding what we see: How we derive meaning from vision. *Trends in Cognitive Sciences*, *19*(11), 677–687. https://doi.org/10.1016/j.tics.2015.08.008

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

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 of the United States of America*, *106*(50), 21442–21446.

https://doi.org/10.1073/pnas.0908921106

Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a priviledged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychological Learning, Memory and Cognition*, *32*(4), 643–658. https://doi.org/10.1037/0278-7393.32.4.643

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*(3), 371–414. https://doi.org/10.1207/s15516709cog2303_4

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, *57*(2), 195–209. https://doi.org/10.1016/j.jml.2006.11.001

De Deyne, S., Kenett, Y. N., Anaki, D., Faust, M., & Navarro, D. (2017). Large-scale network representations of semantics in the mental lexicon. In M. N. Jones (Ed.), *Big data in cognitive science - Frontiers of cognitive psychology* (pp. 174–202). Routledge/Taylor & Francis Group.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The “Small World of Words” English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

Dell’Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, *32*(4), 588–615. https://doi.org/10.3758/BF03200832

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*, 1–10. https://doi.org/10.3389/fpsyg.2010.00177

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124–142. https://doi.org/10.1016/0749-596X(88)90070-8

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*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Devereux, B. J., Tyler, L. K., Geertzen, J., & Randall, B. (2014). The Centre for Speech, Language and the Brain (CSLB) concept property norms. *Behavior Research Methods*, *46*(4), 1119–1127. https://doi.org/10.3758/s13428-013-0420-4

Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience*, *10*(1), 77–94. https://doi.org/10.1162/089892998563798

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 515–523. https://doi.org/10.1037/0278-7393.24.2.515

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*(4), 339–357. https://doi.org/10.1037/0096-3445.120.4.339

Fieder, N., Wartenburger, I., & Abdel Rahman, R. (2019). A close call: Interference from semantic neighbourhood density and similarity in language production. *Memory & Cognition*, *47*(1), 145–168. https://doi.org/10.3758/s13421-018-0856-y

Fodor, J. A. (1976). *The language of thought*. Harvard University Press.

Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, *8*(3), 263–367. https://doi.org/10.1016/0010-0277(80)90008-6

Fodor, J. D., Fodor, J. A., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry*, *6*(4), 515–535.

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International Journal of Psychophysiology*, *29*(1), 65–75. https://doi.org/10.1016/S0167-8760(97)00099-8

Garrard, P., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2001). Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cognitive Neuropsychology*, *18*(2), 125–174. https://doi.org/10.1080/02643290125857

Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language*, *93*(1), 79–94. https://doi.org/10.1016/j.bandl.2004.08.003

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Vol. 1: Speech and talk* (pp. 176–220). Academic Press.

Gerhand, S., & Barry, C. (1999). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, *73*(2), B27–B36. https://doi.org/10.1016/S0010-0277(99)00052-9

Gonnerman, L. M., Andersen, E. S., Devlin, J. T., Kempler, D., & Seidenberg, M. S. (1997). Double dissociation of semantic categories in Alzheimer's disease. *Brain and Language*, *57*(2), 254–279. https://doi.org/10.1006/brln.1997.1752

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Griffiths, T. L., & Steyvers, M. (2003). Prediction and semantic association. *Advances in Neural Information Processing Systems 15 - Proceedings of the 2002 Conference, NIPS 2002*, *15*.

Grober, E., Perecman, E., Kellar, L., & Brown, J. (1980). Lexical knowledge in anterior and posterior aphasics. *Topics in Catalysis*, *10*(2), 318–330. https://doi.org/10.1016/0093-934X(80)90059-0

Grondin, R., Lupker, S. J., & McRae, K. (2009). Shared features dominate semantic richness effects for concrete concepts. *Journal of Memory and Language*, *60*(1), 1–19. https://doi.org/10.1016/j.jml.2008.09.001

Grossman, M. (1981). A bird is a bird is a bird: Making reference within and without superordinate categories. *Brain and Language*, *12*(2), 313–331. https://doi.org/10.1016/0093-934X(81)90022-5

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006–1033. https://doi.org/10.1177/1745691619861372

Hameau, S., Nickels, L., & Biedermann, B. (2019). Effects of semantic neighbourhood density on spoken word production. *Quarterly Journal of Experimental Psychology*, *72*(12), 2752–2775. https://doi.org/10.1177/1747021819859850

Harley, T. A. (1993). Connectionist approaches to language disorders. *Aphasiology*, *7*(3), 221–249. https://doi.org/10.1080/02687039308249508

Heinze, H. J., Muente, T. F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology*, *47*(2), 121–135. https://doi.org/10.1016/S0301-0511(97)00024-0

Hernández-Muñoz, N., Izura, C., & Ellis, A. W. (2006). Cognitive aspects of lexical availability. *European Journal of Cognitive Psychology*, *18*(5), 730–755. https://doi.org/10.1080/09541440500339119

Hirschfeld, G., Jansma, B., Bölte, J., & Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *NeuroReport*, *19*(12), 1227–1230. https://doi.org/10.1097/WNR.0b013e328309ecd1

Hodgson, C., & Lambon Ralph, M. A. (2008). Mimicking aphasic semantic errors in normal speech production: Evidence from a novel experimental paradigm. *Brain and Language*, *104*(1), 89–101. https://doi.org/10.1016/j.bandl.2007.03.007

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition*, *13*(7–8), 884–910. https://doi.org/10.1080/13506280544000093

Hough, M. S., & Pierce, R. S. (1988). Contextual influences on category concept generation in aphasia. *Clinical Aphasiology Paper*, *18*, 507–520. https://doi.org/10.1080/02687038908249022

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 1–16. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Janssen, N., Carreiras, M., & Barber, H. A. (2011). Electrophysiological effects of semantic context in picture and word naming. *NeuroImage*, *57*(3), 1243–1250. https://doi.org/10.1016/j.neuroimage.2011.05.015

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*(9), 2960–2969. https://doi.org/10.1093/cercor/bhu092

Jones, M. N., Willits, J., & Dennis, S. (2015). Models of semantic memory. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford handbook of computational and mathematical psychology* (1st ed., pp. 232–254). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199957996.013.11

Kello, C. T. (2004). Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 942–955. https://doi.org/10.1037/0096-1523.30.5.942

Kello, C. T., & Plaut, D. C. (2000). Strategic control in word reading: Evidence from speeded responding in the tempo-naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 719–750. https://doi.org/10.1037/0278-7393.26.3.719

Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of stroop interference in speech production. *Journal of Experimental Psychology: General*, *129*(3), 340–360. https://doi.org/10.1037/0096-3445.129.3.340

Kiran, S., Ntourou, K., & Eubank, M. (2007). The effect of typicality on online category verification of inanimate category exemplars in aphasia. *Aphasiology*, *21*(9), 844–866. https://doi.org/10.1080/02687030600743564

Kiran, S., & Thompson, C. K. (2003). The role of semantic complexity in treatment of naming deficits. *Journal of Speech, Language, and Hearing Research*, *46*(4), 773–787. https://doi.org/10.1044/1092-4388(2003/061)

Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2007a). Omissions in aphasic picture naming: Late age-of-acquisition is the culprit, not low semantic density. *Brain and Language*, *103*(1–2), 132–133. https://doi.org/10.1016/J.BANDL.2007.07.081

Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2007b). Omissions in aphasic picture naming: Late age-of-acquisition is the culprit, as well as low semantic density. *Poster Presented at the 45th Meeting of the Academy of Aphasia, Washington, DC*, October 21–23.

Kounios, J., Green, D. L., Payne, L., Fleck, J. I., Grondin, R., & McRae, K. (2009). Semantic richness and the activation of concepts in semantic memory: Evidence from event-related potentials. *Brain Research*, *1282*, 95–102. https://doi.org/10.1016/j.brainres.2009.05.092

Laganaro, M. (2014). ERP topographic analyses from concept to articulation in word production studies. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.00493

Laganaro, M., & Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography*, *24*(1), 19–29. https://doi.org/10.1007/s10548-010-0162-x

Laganaro, M., Valente, A., & Perret, C. (2012). Time course of word production in fast and slow speakers: A high density ERP topographic study. *NeuroImage*, *59*(4), 3881–3888. https://doi.org/10.1016/j.neuroimage.2011.10.082

Laiacona, M., Luzzatti, C., Zonca, G., Guarnaschelli, C., & Capitani, E. (2001). Lexical and semantic factors influencing picture naming in aphasia. *Brain and Cognition*, *46*(1–2), 184–187. https://doi.org/10.1016/S0278-2626(01)80061-0

Laisney, M., Giffard, B., Belliard, S., de la Sayette, V., Desgranges, B., & Eustache, F. (2011). When the zebra loses its stripes: Semantic priming in early Alzheimer's disease and semantic dementia. *Cortex*, *47*(1), 35–46. https://doi.org/10.1016/j.cortex.2009.11.001

Lampe, L. F., Fieder, N., Krajenbrink, T., & Nickels, L. (2017). Semantische Nachbarschaft in der

Wortproduktion bei Aphasie [Semantic neighbourhood in word production in aphasia]. In A. Adelt, Ö. Yetim, C. Otto, & T. Fritzsche (Eds.), *Spektrum Patholinguistik - Panorama Patholinguistik: Sprachwissenschaft trifft Sprachtherapie* (Vol. 10, pp. 103–114). Universitätsverlag Potsdam. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/39701

Landauer, T. K. (2006). Latent semantic analysis. In L. Nadel (Ed.), *Encyclopedia of cognitive science*. Wiley. https://doi.org/10.1002/0470018860.s00561

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. https://doi.org/10.1037/0033-295X.104.2.211

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2–3), 259–284. https://doi.org/10.1080/01638539809545028

Larochelle, S., Richard, S., & Soulières, I. (2000). What some effects might not be: The time to verify membership in "well-defined" categories. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *53*(4), 929–961. https://doi.org/10.1080/713755940

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, *10*(5), 553–567. https://doi.org/10.1162/089892998562960

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Lloyd-Jones, T. J., & Nettlemill, M. (2007). Sources of error in picture naming under time pressure. *Memory & Cognition*, *35*(4), 816–836. https://doi.org/10.3758/BF03193317

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203–208. https://doi.org/10.3758/BF03204766

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience*, *14*(3), 455–462. https://doi.org/10.1162/089892902317361967

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*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Mahon, B. Z., & Navarrete, E. (2016). Modelling lexical access in speech production as a ballistic process. *Language, Cognition and Neuroscience*, *31*(4), 521–523. https://doi.org/10.1080/23273798.2015.1129060

Marques, J. F. (2005). Naming from definition: The role of feature type and feature distinctiveness. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(4), 603–611. https://doi.org/10.1080/02724980443000106

Mason-Baughman, M. B., & Wallace, S. E. (2013a). Semantic feature knowledge in persons with aphasia: The role of commonality, distinctiveness, and importance. *Aphasiology*, *27*(3), 364–380. https://doi.org/10.1080/02687038.2012.730602

Mason-Baughman, M. B., & Wallace, S. E. (2013b). The role of commonality, distinctiveness and importance of semantic features in persons with aphasia. *Brain Injury*, *27*(4), 399–407. https://doi.org/10.3109/02699052.2012.750748

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 3–23. https://doi.org/10.1037//0278-7393.21.1.3

Mayberry, E. J., Sage, K., & Lambon Ralph, M. A. (2011). At the edge of semantic space: The breakdown of coherent concepts in semantic dementia is constrained by typicality and severity but not modality. *Journal of Cognitive Neuroscience*, *23*(9), 2240–2251. https://doi.org/10.1162/jocn.2010.21582

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, *88*(5), 375–407.

https://doi.org/10.1037/0033-295X.88.5.375

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, *11*(1), 1–37. https://doi.org/10.1016/0010-0285(79)90002-1

McRae, K. (2004). Semantic memory: Some insights from feature-based connectionist attractor networks. In B. H. Ross (Ed.), *The psychology of learning and motivation - Advances in research and theory* (pp. 41–86). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(03)45002-0

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

McRae, K., Cree, G. S., Westmacott, R., & Sa, V. R. De. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology*, *53*(4), 360–373. https://doi.org/10.1037/h0087323

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99–130. https://doi.org/10.1080/10430719008404646

McRae, K., & Jones, M. (2013). Semantic memory. In D. Reisberg (Ed.), *Oxford library of psychology. The Oxford handbook of cognitive psychology* (pp. 206–219). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195376746.013.0014

Melinger, A., & Abdel Rahman, R. (2013). Lexical selection is competitive: Evidence from indirectly activated semantic associates during picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 348–364. https://doi.org/10.1037/a0028941

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, *25*(10), 3343–3355. https://doi.org/10.1093/cercor/bhu137

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective and Behavioral Neuroscience*, *11*(1), 32–43. https://doi.org/10.3758/s13415-010-0009-7

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 65–79. https://doi.org/10.1037/0278-7393.34.1.65

Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory and Cognition*, *37*(7), 1026–1039. https://doi.org/10.3758/MC.37.7.1026

Monetta, L., Tremblay, T., & Joanette, Y. (2003). Semantic processing of words, cognitive resources and N400: An event-related potentials study. *Brain and Cognition*, *53*(2), 327–330. https://doi.org/10.1016/S0278-2626(03)00136-2

Morrison, C. M., & Gibbons, Z. C. (2006). Lexical determinants of semantic processing speed. *Visual Cognition*, *13*(7–8), 949–967. https://doi.org/10.1080/13506280544000129

Moses, M. S., Nickels, L., & Sheard, C. (2004). "I'm sitting here feeling aphasic!" A study of recurrent perseverative errors elicited in unimpaired speakers. *Brain and Language*, *89*(1), 157–173. https://doi.org/10.1016/S0093-934X(03)00364-X

Moss, H. E., Tyler, L. K., Durrant-Peatfield, M., & Bunn, E. M. (1998). 'Two eyes of a see-through': Impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase*, *4*(4–5), 291–310. https://doi.org/10.1080/13554799808410629

Murphy, G. L. (2002). *The big book of concepts*. The MIT Press.

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. https://doi.org/10.1016/j.jml.2014.05.003

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Nickels, L. (1995). Getting it right? Using aphasic naming errors to evaluate theoretical models of spoken word recognition. *Language and Cognitive Processes*, *10*(1), 13–45.

https://doi.org/10.1080/01690969508407086

Nickels, L. (1997). *Spoken word production and its breakdown in aphasia*. Psychology Press.

Nickels, L., & Howard, D. (1994). A frequent occurrence? Factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology*, *11*(3), 289–320. https://doi.org/10.1080/02643299408251977

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, *33*(10), 1281–1303. https://doi.org/10.1016/0028-3932(95)00102-9

Nozari, N., & Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics*, *53*, 1–19. https://doi.org/10.1016/j.jneuroling.2019.100875

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*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007a). The neural consequences of semantic richness: When more comes to mind, less activation is observed. *Psychological Science*, *18*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2007.01913.x

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007b). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience*, *19*(8), 1407–1419. https://doi.org/10.1162/jocn.2007.19.8.1407

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin and Review*, *15*(1), 161–167. https://doi.org/10.3758/PBR.15.1.161

Pexman, P. M., Holyk, G. G., & Monfils, M.-H. (2003). Number-of-features effects and semantic processing. *Memory & Cognition*, *31*(6), 842–855. https://doi.org/10.3758/BF03196439

Pexman, P. M., Lupker, S., & Hino, Y. (2002). The impact of feedback semantics in visual word recongnition. *Psychonomic Bulletin & Review*, *9*(3), 542–549. https://doi.org/10.3758/BF03196311

Pexman, P. M., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience*, *7*, 1–3. https://doi.org/10.3389/fnhum.2013.00723

Piai, V., Roelofs, A., & van der Meij, R. (2012). Event-related potentials and oscillatory brain responses associated with semantic and Stroop-like interference effects in overt naming. *Brain Research*, *1450*, 87–101. https://doi.org/10.1016/j.brainres.2012.02.050

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*(5), 377–500. https://doi.org/10.1080/02643299308253469

Python, G., Fargier, R., & Laganaro, M. (2018a). ERP evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex*, *99*, 1–12. https://doi.org/10.1016/j.cortex.2017.09.008

Python, G., Fargier, R., & Laganaro, M. (2018b). When wine and apple both help the production of grapes: ERP evidence for post-lexical semantic facilitation in picture naming. *Frontiers in Human Neuroscience*, *12*, 1–13. https://doi.org/10.3389/fnhum.2018.00136

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, *146*, 240–244. https://doi.org/10.1016/j.cognition.2015.09.016

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2021). Semantic richness and density effects on language production: Electrophysiological and behavioral evidence. Journal of Experimental Psychology: Learning Memory and Cognition, 47(3), 508–517. https://doi.org/10.1037/xlm0000940

Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00011

Räling, R., Holzgrefe-Lang, J., Schröder, A., & Wartenburger, I. (2015). On the influence of typicality and age of acquisition on semantic processing: Diverging evidence from behavioural and ERP responses. *Neuropsychologia*, *75*, 186–200. https://doi.org/10.1016/j.neuropsychologia.2015.05.031

Räling, R., Schröder, A., & Wartenburger, I. (2016). The origins of age of acquisition and typicality effects: Semantic processing in aphasia and the ageing brain. *Neuropsychologia*, *86*, 80–92. https://doi.org/10.1016/j.neuropsychologia.2016.04.019

Randall, B., Moss, H. E., Rodd, J. M., Greer, M., & Tyler, L. K. (2004). Distinctiveness and correlation in conceptual structure: Behavioral and computational studies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 393–406. https://doi.org/10.1037/0278-7393.30.2.393

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460–499. https://doi.org/10.1037/0033-295X.107.3.460

Reilly, M., & Desai, R. H. (2017). Effects of semantic neighborhood density in abstract and concrete words. *Cognition*, *169*, 46–53. https://doi.org/10.1016/j.cognition.2017.08.004

Rico Duarte, L., Marquié, L., Marquié, J. C., Terrier, P., & Ousset, P. J. (2009). Analyzing feature distinctiveness in the processing of living and non-living concepts in Alzheimer's disease. *Brain and Cognition*, *71*(2), 108–117. https://doi.org/10.1016/j.bandc.2009.04.007

Rico Duarte, L., & Robert, C. (2014). Semantic richness influences naming pictures in old age and in Alzheimer's disease. *Journal of Neurolinguistics*, *31*, 55–68. https://doi.org/10.1016/j.jneuroling.2014.06.004

Riley, E. A., & Thompson, C. K. (2010). Semantic typicality effects in acquired dyslexia: Evidence for semantic impairment in deep dyslexia. *Aphasiology*, *24*(6–8), 802–813. https://doi.org/10.1080/02687030903422486

Riordan, B., & Jones, M. N. (2011). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, *3*(2), 303–345. https://doi.org/10.1111/j.1756-8765.2010.01111.x

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic organization and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior 12*, *1*(1), 1–20.

https://doi.org/10.1016/S0022-5371(73)80056-8

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (1997). A case for nondecomposition in conceptually driven word retrieval. *Journal of Psycholinguistic Research*, *26*(1), 33–67. https://doi.org/10.1023/A:1025060104569

Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, *172*, 59–72. https://doi.org/10.1016/j.cognition.2017.12.007

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*(1), 205–235. https://doi.org/10.1037/0033-295X.111.1.205

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rose, S. B., & Abdel Rahman, R. (2016). Cumulative semantic interference for associative relations in language production. *Cognition*, *152*, 20–31. https://doi.org/10.1016/J.COGNITION.2016.03.013

Rose, S. B., & Abdel Rahman, R. (2017). Semantic similarity promotes interference in the continuous naming paradigm: Behavioural and electrophysiological evidence. *Language, Cognition and Neuroscience*, *32*(1), 55–68. https://doi.org/10.1080/23273798.2016.1212081

Rose, S. B., Aristei, S., Melinger, A., & Abdel Rahman, R. (2019). The closer they are, the more they interfere: Semantic similarity of word distractors increases competition in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 753–763. https://doi.org/10.1037/xlm0000592

Rossiter, C., & Best, W. (2013). "Penguins don't fly": An investigation into the effect of typicality on picture naming in people with aphasia. *Aphasiology*, *27*(7), 784–798. https://doi.org/10.1080/02687038.2012.751579

Sandberg, C., Sebastian, R., & Kiran, S. (2012). Typicality mediates performance during category verification in both ad-hoc and well-defined categories. *Journal of Communication Disorders*, *45*(2), 69–83. https://doi.org/10.1016/j.jcomdis.2011.12.004

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press. https://doi.org/10.1017/CBO9780511526817

Siew, C. S. Q. (2020). Feature distinctiveness effects in language acquisition and lexical processing: Insights from megastudies. *Cognitive Processing, 21*(4), 669–685. https://doi.org/10.1007/s10339-019-00947-6

Stanczak, L., Waters, G., & Caplan, D. (2006). Typicality-based learning and generalisation in aphasia: Two case studies of anomia treatment. *Aphasiology*, *20*(2–4), 374–383. https://doi.org/10.1080/02687030600587631

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp. 143–186). Lawrence Erlbaum Associates.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1140–1154. https://doi.org/10.1037/0278-7393.21.5.1140

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*, 1–16. https://doi.org/10.3389/fpsyg.2011.00356

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928. https://doi.org/10.1093/cercor/bhp153

Swaab, T. Y., Baynes, K., & Knight, R. T. (2002). Separable effects of priming and imageability on word processing: An ERP study. *Cognitive Brain Research*, *15*(1), 99–103. https://doi.org/10.1016/S0926-6410(02)00219-7

Taylor, K. I., Devereux, B. J., Acres, K., Randall, B., & Tyler, L. K. (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. *Cognition*, *122*(3), 363–374. https://doi.org/10.1016/j.cognition.2011.11.001

Taylor, K. I., Moss, H., Randall, B., & Tyler, L. K. (2004). The interplay between distinctiveness and intercorrelation in the automatic activation of word meaning. *Proceedings of the 45th Annual Meeting of the Psychonomic Society*, *9*, 109.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(6), 244–252. https://doi.org/10.1016/S1364-6613(00)01651-X

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain and Language*, *75*(2), 195–231. https://doi.org/10.1006/brln.2000.2353

Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience*, *8*, 1–13. https://doi.org/10.3389/fnins.2014.00390

Vieth, H. E., McMahon, K. L., & de Zubicaray, G. I. (2014a). The roles of shared vs. distinctive conceptual features in lexical access. *Frontiers in Psychology*, *5*, 1–12. https://doi.org/10.3389/fpsyg.2014.01014

Vieth, H. E., McMahon, K. L., & de Zubicaray, G. I. (2014b). Feature overlap slows lexical selection: Evidence from the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology*, *67*(12), 2325–2339. https://doi.org/10.1080/17470218.2014.923922

Vigliocco, G., & Vinson, D. P. (2007). Semantic representation. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 195–214). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198568971.001.0001

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Vinson, D. P., Andrews, M., & Vigliocco, G. (2013). Giving words meaning: Why better models of semantics are needed in language production research. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *Oxford library of psychology. The Oxford handbook of language production* (pp. 134–151). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199735471.013.006

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183–190. https://doi.org/10.3758/BRM.40.1.183

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 664–680. https://doi.org/10.1037/0278-7393.17.4.664

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 243–259. https://doi.org/10.1037/0278-7393.19.2.243

Woollams, A. M. (2012). Apples are not the only fruit: The effects of concept typicality on semantic representation in the anterior temporal lobe. *Frontiers in Human Neuroscience*, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00085

Woollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia*, *46*(10), 2503–2514. https://doi.org/10.1016/j.neuropsychologia.2008.04.005

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 1–10. https://doi.org/10.3389/fnhum.2012.00072

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin and Review*, *18*(4), 742–750. https://doi.org/10.3758/s13423-011-0092-y

CHAPTER 2

Effects of semantic variables on word production in aphasia

Leonie F. Lampe
Solène Hameau
Nora Fieder
Lyndsey Nickels

Accepted for publication in *Cortex*

An earlier version of this paper was presented at the *Science of Aphasia conference:* Lampe, L., Fieder, N., Hameau, S., & Nickels, L. (2018). Effects of semantic variables on picture naming in a large group of people with aphasia. *Stem-, Spraak- en Taalpathologie, 23*(Suppl.), 109–112.

Pages 60-160 of this thesis have been removed due to copyright restriction. Please refer to the following citation for details of the article contained in these pages.

Leonie F. Lampe, Solène Hameau, Nora Fieder, Lyndsey Nickels, Effects of semantic variables on word production in aphasia, *Cortex*, 141, 363-402

DOI: 10.1016/j.cortex.2021.02.020

CHAPTER 3

Semantic variables both help and hinder word production: Behavioural evidence from picture naming

Leonie F. Lampe
Solène Hameau
Lyndsey Nickels

Accepted for publication in
Journal of Experimental Psychology: Learning, Memory, and Cognition

Abstract

This research investigated how word production is influenced by six feature-based semantic variables (number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness). We simultaneously investigated effects of the six semantic variables on spoken picture naming in a large group of participants (n = 87), while controlling for other psycholinguistic variables. Across analyses, number of semantic features was the most consistent predictor with a facilitatory effect on naming latency and accuracy. In addition, inhibitory effects were found on naming accuracy for intercorrelational density and on naming latency for distinctiveness. The facilitatory effect of number of semantic features is suggested to stem from stronger semantic activation with an increasing number of semantic features, which results in facilitated selection of the word's lexical representation. In contrast, the inhibitory effect of intercorrelational density is most easily accounted for by increased competition at the lexical level. The mechanism underlying the inhibitory effect of distinctiveness is unclear. These findings indicate that future research on factors affecting word retrieval should also control for effects of number of semantic features, intercorrelational density, and distinctiveness. They also suggest that effects of the other semantic variables (e.g., semantic neighbours) reported in the literature were potentially overestimated due to insufficient control of other semantic and/or psycholinguistic variables.

Introduction

Decades of research targeted at understanding processes during word production have led to several theoretical accounts of word production (e.g., Abdel Rahman & Melinger, 2009, 2019; Dell, 1986; Howard et al., 2006; Levelt et al., 1999), which mostly agree about the general architecture of word production and there is a consensus that it is a semantically mediated process (e.g., Bock & Levelt, 1994; Vitkovitch & Humphreys, 1991; but see Kremin, 1986). The most common experimental paradigm for these investigations is picture naming. In this paradigm, following picture processing, the activation of word meaning at a semantic level leads to lexical level activation, where the target word's lexical representation is selected among (possibly competing) alternatives. This is followed by activation of the target's word form and then by further phonological and phonetic processes that lead to word production. However, there is disagreement about crucial details of the model. More specifically, it is, for example, still unclear how exactly information is processed at each level of the model and researchers disagree about whether there is competition between co-activated representations, the mechanism underlying any such competition, and the level(s) it might affect.

Word production processes are often investigated by manipulating the context in which a target word appears (Picture Word Interference, e.g., Gauvin et al., 2018; Melinger & Rahman, 2013; Blocked-Cyclic Naming, e.g., Ewald et al., 2012; Python et al., 2018; Continuous Naming, e.g., Howard et al., 2006; Rose & Abdel Rahman, 2016). Rather than manipulating the context in which words occur, another approach to the investigation of processes in word production is to examine the effects of item-inherent characteristics that naturally differ from word to word (e.g., word frequency or imageability, e.g., Alario et al., 2004). The principle is that information processing at the different levels of word production may be affected by manipulations of, or naturally occurring variability in, these word characteristics. Hence, the effects of these word characteristics can in turn inform and advance our understanding of representation and processing in word production. One set of variables that attracts research attention are those that capture information about the semantic representation of a target word and its relationship to other words in the mental lexicon (e.g., Fieder et al., 2019; Rabovsky et al., 2016, 2021): *semantic variables.* Semantic variables operationalise statistical or distributional facts about the presumed semantic representation of a target word and can capture a wide range of aspects of the meaning of single concepts or of its relationships to other words in the mental lexicon. They can, for example, describe the richness of the semantic representation of the target word or the number of words that are closely related in meaning. It is the effects of these semantic variables on spoken word production that are the focus of the research presented here.

Most of the previous research into effects of semantic variables was conducted to investigate the underlying semantic representations themselves and thus focused on word comprehension, using tasks like semantic categorisation and feature verification, or conducted computational modelling from this perspective (e.g., Fujihara et al., 1998; McRae et al., 1997; Mirman & Magnuson, 2008; Pexman et al., 2003; Randall et al., 2004). However, in word production, semantic processing precedes and therefore determines lexical processing. Consequently, semantic variables are hypothesised to not only have consequences for semantic processing but may also affect subsequent lexical processing. However, current theories of word production are underspecified with respect to effects of semantic variables and details of how they might affect processing are not clearly stated. This makes word production a particularly interesting and relevant modality in which to study effects of semantic variables.

Given the possibility of widespread influence of semantic variables on the word production process, it is important to establish which semantic variables reliably affect word production (as measured in picture naming) and to understand the mechanisms underlying these effects. Understanding the influence of these semantic variables will facilitate our understanding of how information is represented and processed at the different stages of word production and help us to better understand word production mechanisms. More specifically, studying semantic variables will shed light on the structure of semantic representations and their influence on retrieval and selection of words in the mental lexicon. This also allows us to evaluate the adequacy of current theories of word production, given that some theories may not include the architectural elements required to explain the effects observed. Moreover, knowledge of the semantic variables that influence word production is also of methodological importance, as it can, for example, be used to achieve improved matching and control of influential variables.

Depending on the theory of semantic representation, that is, the nature of the semantic dimensions encoded in the human mind, various relationships are suggested to be important, each of which represents a different approach to specifying aspects of semantic representation. Semantic relationships can be described on the basis of participant-generated associations (e.g., De Deyne et al., 2019; Nelson et al., 2004), co-occurrence patterns in text corpora (e.g., Latent Semantic Analysis, Landauer et al., 1998; Continuous bag-of-words model, Mikolov et al., 2013), or participant-generated semantic features (e.g., Devereux et al., 2014; McRae et al., 2005; Vinson & Vigliocco, 2008).

In this study, we focus on effects of semantic variables that can be operationalised based on semantic features, as provided in the feature norm database by McRae et al. (2005). Importantly, we believe that the use of feature-based semantic variables is not only compatible with semantic knowledge being decomposed into semantic features. In contrast, feature-based semantic variables could also be indicative of relations between holistic lexical concepts like, for example, the number or strength of connections between them (e.g., Abdel Rahman & Melinger, 2009; Collins & Loftus, 1975; Levelt et al., 1999). In our theoretical interpretations of the findings, we use a cascading, interactive model of word production that assumes lexical competition (Abdel Rahman & Melinger, 2009, 2019), however, we also assess the ability of other current theories of word production to account for the observed effects of semantic variables and consider both feature-based and holistic semantic representations.

Previous research into effects of semantic variables was often inconclusive and subject to methodological flaws (discussed below). Consequently, here, we aimed to extend and improve the evidence base, by investigating the effects of these semantic variables on the word production process.

Review of the previous literature on feature-based semantic variables

Effects of semantic variables that can be operationalised based on semantic features were studied to varying degrees in previous research. Table 1 summarises effects of feature-based semantic variables from studies that used simple picture naming paradigms with neurotypical adults, and, as is apparent, both facilitatory and inhibitory effects have been observed. While facilitation in word production is usually attributed to spreading activation at the semantic level (e.g., Abdel Rahman & Melinger, 2009) or feedback from the lexical to the semantic level (e.g., Dell et al., 1986), inhibition has been linked to competitive processes at the lexical (Abdel Rahman & Melinger, 2009, 2019; Howard et al., 2006; Levelt et al., 1999) or post-lexical stages (Mahon et al., 2007), or to a non-competitive learning mechanism influencing the strength of connections between semantic and lexical units (Oppenheim et al., 2010). We now address each of these variables in turn.

Number of semantic features

Number of semantic features, often referred to as *semantic richness*, is a simple count of the total number of features generated by participants in response to a target word, for example when producing features for a feature-norm database (e.g., McRae et al., 2005). A higher number of features has been consistently found to facilitate both response times in picture naming and, where investigated, naming accuracy (Rabovsky et al., 2016, 2021; Taylor et al., 2012). Looking at specific types of semantic features (e.g., distinctive perceptual or shared functional features) Rico Duarte and Robert (2014) also reported a facilitatory effect from a higher number of semantic features.

Intercorrelational Density

Intercorrelational density captures the extent to which feature pairs of a concept co-occur across concepts. Rabovsky and colleagues (2016, 2021) reported lower naming accuracy for items with higher intercorrelational density. In contrast, there is inconsistent evidence regarding an effect on response latencies: While Rabovsky et al. (2016) found significantly slower latencies with increasing intercorrelational density, this effect only reached significance in a second repetition of the item set in Rabovsky et al., (2021), and was non-significant in Taylor et al. (2012).

Table 1

Effects of semantic variables on picture naming in neurotypical adults: previous research

Study	Participants (*n*)	Items (*n*)	Design/analyses	Operationalisation	RT	Accuracy	Semantic errors	Omissions
Number of semantic features								
Rabovsky et al., 2016	16	541	continuous, (g)lmer	feature database (McRae et al., 2005)	↗	↗		
Rabovsky et al., 2021	31	345	continuous, lmer, Bayesian model	feature database (McRae et al., 2005)	↗	↗		
Taylor et al., 2012	20	302	continuous, (g)lmer	feature database (McRae et al., 2005)	↗			
Intercorrelational density								
Rabovsky et al., 2016	16	541	continuous, (g)lmer	feature database (McRae et al., 2005)	↙	↙		
Rabovsky et al., 2021	31	345	continuous, lmer, Bayesian model	feature database (McRae et al., 2005)	Ø [a]	↙		
Taylor et al., 2012	20	302	continuous, (g)lmer	feature database (McRae et al., 2005) [b]	Ø			
Number of near semantic neighbours								
Fieder et al., 2019 [c]	30	180	continuous, (g)lmer	feature database (Devereux et al., 2014)	↙	↙	↗ (also for coordinate errors)	↗
Mirman, 2011 [c]	35 older adults	57	matched sets	feature database (McRae et al., 2005)	Ø	↙	↗	
Hameau et al., 2019	40	84	continuous, (g)lmer	feature database (McRae et al., 2005) and rating; feature-based semantic neighbourhood factor: number of near	Ø [d]	Ø [d]		

Study	Participants (*n*)	Items (*n*)	Design/analyses	Operationalisation	RT	Accuracy	Semantic errors	Omissions
				semantic neighbours and rated competitors				
Lampe et al., 2017	15 older adults	44	matched sets	feature database (McRae et al., 2005)	Ø	Ø		
Bormann, 2011	18	54	matched sets, (g)lmer	rated within category competitors [e]	Ø			
				Semantic similarity				
Fieder et al., 2019 [c]	30	180	continuous, (g)lmer	feature database (Devereux et al., 2014)	Ø	↙	↗	↗
				Typicality				
Dell'Acqua et al., 2000	84	266	continuous, multiple regression	rating	↗			
Grossman et al., 1998	14 older adults	72	matched sets	rating	↗			
Holmes & Ellis, 2006 [f]	25	84	matched sets	rating	↗	Ø		
Jolicoeur et al., 1984	18	48	matched sets	rating	↗			
Fieder et al., 2019 [c]	30	180	continuous, (g)lmer	rating	Ø	↗	Ø (also for coordinate errors)	↙
Morrison et al., 1992 [g]	20	48	continuous, multiple regression	rating	Ø			
Woollams, 2012 [h]	16	80	matched sets	rating	Ø	Ø		
Rogers et al., 2015	12 older adults	48	matched sets, no statistics reported	rating		↙		

Study	Participants (n)	Items (n)	Design/analyses	Operationalisation	RT	Accuracy	Semantic errors	Omissions
				Distinctiveness				
Rabovsky et al., 2016	16	541	continuous, (g)lmer	feature database (McRae et al., 2005)	↗ [i]	↗ [i]		
Taylor et al., 2012	31	345	continuous, (g)lmer	feature database (McRae et al., 2005)	↗ [j]			
Miozzo et al., 2015 [k]	17	146	continuous, stepwise regression	feature database (McRae et al., 2005)		Ø		
Humphreys et al., 1988	20	76	matched sets	category distinctiveness ("structural similarity"); rated features in common with exemplars of a category (control participants)	↗			

Note. RT = response time, (g)lmer = (generalised) linear mixed effects model analyses, RTs and accuracy: ↙ = poorer performance – slower RTs and decreased accuracy with higher values of the measure, ↗ = improved performance – faster RTs with increased accuracy and higher accuracy with higher values of the variable, Semantic errors and omissions: ↙ = reduced numbers of errors of this type with higher values of the variable, ↗ = increased numbers of errors of this type with higher values of the variable, Ø = no effect, blank cells = not investigated.

Participants—where not otherwise specified, young adults (typically undergraduate students).

[a] inhibitory effect on RT significant in a second round of naming.

[b] correlation measure based on 'intercorrelational strength' in Taylor et al. (2012).

[c] speeded naming paradigm (500ms deadline).

[d] results based on Holm-Bonferroni corrected *p*-values to account for testing for multiple variables of interest.

[e] Bormann's (2011) measure was based on ratings capturing the number of category coordinates of a target, which was then dichotomised into words with many and few competitors.

[f] similar results in picture naming after familiarisation and a subsequent second round of naming.

[g] rated typicality in two categories: man-made versus naturally occurring objects.

[h] pre rTMS.

[i] distinctiveness added in additional analysis in discussion; non-significant when intercorrelational density was in the model at the same time.

[j] naming was also faster for concepts with more highly correlated distinctive features.

[k] Miozzo et al.'s (2015) *Specific Semantic Features* measure contained two measures grouped through Principal Component Analysis: number of distinctive features and number of encyclopaedic features.

Number of near semantic neighbours

Near semantic neighbours are defined as words that are semantically very similar to the target word and share many semantic features with it (e.g., a cosine similarity of at least 0.4 between feature vectors, Mirman, 2011). Increased numbers of near semantic neighbours resulted in less accurate picture naming accuracy and lead to more semantic errors and omissions (versus correct responses) under special conditions in a speeded picture naming task (a variant of simple picture naming designed to increase naming errors, Fieder et al., 2019; Mirman, 2011). Fieder et al. also found that a higher number of near semantic neighbours led to slower responses in speeded naming. However, Hameau et al. (2019), Lampe et al. (2017), and Bormann (2011, using a similar measure) found no significant effects in standard picture naming, on either errors or response latencies.

Semantic similarity

Fieder et al.'s (2019) semantic neighbourhood similarity measure captured the average similarity of a target word and its semantic neighbours, based on the cosine similarity of their feature vectors (note that in other research the term semantic similarity is used to refer to the relationship between a target and a stimulus context (e.g., Rose et al., 2019) or the semantic overlap between two concepts in a particular context (e.g., Rose & Abdel Rahman, 2017)). Fieder et al. were the first to examine effects of semantic neighbourhood similarity on participants' picture naming performance but did so using speeded picture naming. They found that naming accuracy in a speeded picture naming task was reduced for words with higher semantic neighbourhood similarity (i.e., semantically more similar neighbours) and participants were more likely to make a semantic error, coordinate error, or an omission over a correct response. In contrast, the effect of semantic neighbourhood similarity on naming latencies was not significant.

Typicality

Typicality captures the extent to which a target word is representative of its taxonomic category. Previous studies investigating typicality in word production used ratings to determine concept typicality (e.g., Dell'Acqua et al., 2000; Holmes & Ellis, 2006); however, it can also be obtained using feature norms (e.g., Rosch & Mervis, 1975). Compared to the other semantic variables, a relatively large body of literature has investigated effects of rated typicality on picture naming latency and accuracy. Where significant effects of typicality on response times were found (Dell'Acqua et al., 2000; Grossman et al., 1998; Holmes & Ellis, 2006; Jolicoeur et al., 1984), participants responded faster to words of higher typicality. However, note that significant effects on response latencies are not found consistently (see Fieder et al., 2019; Morrison et al., 1992; Woollams, 2012, for non-significant effects). For naming accuracy, Rogers et al. (2015) and Fieder et al. (speeded picture naming) reported conflicting results, while other studies reported non-significant results (Holmes & Ellis, 2006; Woollams, 2012). However, the higher statistical power of Fieder et al.'s speeded naming study, the alignment of their effect with the facilitatory effect of typicality on response times in other studies, as well as the lack of statistical analyses by Rogers et al. make a strong case for a generally facilitatory effect of typicality.

Distinctiveness

Distinctiveness indicates how special or unique the features of a concept are with regard to all other concepts of the database. Higher concept distinctiveness has been found to lead to faster (Rabovsky et al., 2016; Taylor et al., 2012) and more accurate responses (Rabovsky et al., 2016), although the effects in Rabovsky et al. were only significant once intercorrelational density was excluded from the analyses. Miozzo et al. (2015) investigated a similar measure, the number of distinctive features of a concept. The authors used Principal Component Analysis and found that this measure loaded on the same component, labelled *Specific Semantic Features*, as the number of encyclopaedic features. However, there was no evidence that this Specific Semantic Features component predicted naming latency in a stepwise regression analysis. Finally, also Humphreys et al. (1988) studied a related measure, structural similarity, which captures the extent to which a category is rated as having members which share their features. Participants were faster to name items from categories that were rated as having few shared features (the members of which were therefore presumably of higher average distinctiveness) in contrast to items from categories with many shared features (and therefore lower distinctiveness). While effects of distinctiveness were not assessed at the item level, the reported effect matches what Rabovsky et al. and Taylor et al. found for concept distinctiveness.

Limitations of previous research

This brief overview of the literature demonstrates that some semantic variables (e.g., typicality) have been investigated more than others (e.g., intercorrelational density). Despite this imbalance and the sparseness of the evidence base, previous research reporting *significant* effects of a single semantic variable has mostly provided a consensus on the *direction* of any observed effect (i.e., positive or negative effect), although some studies did not find conclusive (significant) effects. The occurrence of both significant and non-significant findings for some semantic variables could be due to experimental and methodological differences between the studies or shortcomings in their design. More specifically, studies differed widely in statistical power due to varying numbers of items and/or participants, characteristics of the participants (e.g., age), and experimental factors, such as the experimental language or the task design (e.g., speeded vs standard naming).

However, two crucial shortcomings apply to most previous studies: First, previous studies of effects of semantic variables on word production have focused on only one or two types of semantic variable at the same time. Importantly, this approach neglects the issue that word meaning is likely to be a multidimensional construct and that different aspects of word meaning probably each explain unique variance with their effects occurring simultaneously (Pexman et al., 2013). Hence, given that semantic variables would be expected to operate simultaneously and not selectively, they would also be expected to influence participants' picture naming behaviour simultaneously. Yet, despite this, previous research has not studied the joint effects of different types of semantic variables on the word production process.

Second, previous research has often insufficiently controlled for (non-semantic) psycholinguistic variables that are known to influence word production (e.g., Alario et al., 2004; Perret & Bonin, 2019). This could have led to statistically significant effects of semantic variables in fact being artifacts of the lack of control. This is of particular concern given the intercorrelations within and between semantic and other psycholinguistic variables (e.g., words with more semantic features tend to have higher typicality and be of lower age of acquisition). Consequently, a significant effect of a semantic variable in the absence of sufficient control of psycholinguistic control variables and other semantic variables cannot confidently be interpreted as an independent effect of the variable of interest and the chance of a false positive finding, a Type 1 error, is increased. It is hence crucial that potentially confounding variables are controlled in the experimental design (e.g., via matching) or the statistical analyses (e.g., by using regression analyses) to avoid false positives. Only then can the unique effect of a variable of interest be identified with confidence. This was the aim of our research.

Two studies that deserve closer attention in the context of our research, are those of Rabovsky and colleagues (Rabovsky et al., 2016, 2021) who conducted behavioural and EEG analyses of number of semantic features and intercorrelational density. Rabovsky and colleagues (2016, 2021) improved on most previous research by studying these two semantic variables simultaneously. However, they only included familiarity, number of orthographic neighbours, and lexical frequency in their analyses (Rabovsky et al., 2016; visual complexity was additionally included in Rabovsky et al., 2021) and did not account for other psycholinguistic variables that influence word production (e.g., age of acquisition, name agreement, imageability, image agreement, Perret & Bonin, 2019). However, the rigorous control of these variables is especially crucial when the effect of interest is potentially rather small.

Rabovsky et al. (2016, 2021) used McRae et al.'s (2005) feature database to derive stimuli and associated measures of number of semantic features and intercorrelational density. However, given that the studies were conducted with German-speaking participants, Rabovsky and colleagues translated the (English) database to German. While one can probably assume that cultural differences between English and German speakers would not have a dramatic impact on conceptual representations (and features) of the items in the database, some are undoubtedly problematic for a picture naming experiment in German. More specifically, some of the translated names for concepts have very low frequency in German and are unlikely to be elicited with high name agreement and other concepts are highly specific to North American culture (e.g., *pie*, *gopher, cedar*) (note however, that Rabovsky et al. (2016) familiarised half of their participants with the correct names before the experiment). Importantly, such differences may arise across languages, but may also influence performance of participants of different cultures with the same native language. Specifically, in this study we tested Australian English speakers, who may be as unfamiliar with some of the North American concepts as German participants. Thus, in the norming study presented in Study 1 we identified the items in the McRae et al.'s (2005) database that had high name agreement in Australian English, and selected these for Study 2, in order to reduce any impact that may have arisen due to cultural differences between speakers of American and Australian English.

Moreover, Rabovsky et al. (2016, 2021) used black and white photographs of the objects, which may have hindered recognition of some objects (e.g., lime vs lemon). Consequently, even though the studies by Rabovsky et al. (2016, 2021) provide some of the most comprehensive analyses to date, it remains unclear how reliable the effects are and, as in the other studies in the literature, whether some effects are miscalibrated (i.e., over- or underestimated) due to insufficient experimental control of psycholinguistic and semantic variables.

The current research

In the light of the shortcomings of the previous studies to have explored semantic variables, we investigated the influences of an increased number (six) of different semantic variables on picture naming performance. This work aimed to determine if these semantic variables affect word production latency and accuracy by examining their effects simultaneously, while taking into account effects of relevant psycholinguistic variables. As outlined above, this knowledge will inform and advance word production theories by clarifying and dissociating effects of these semantic variables. Hence, this exploratory study both aims to make methodological contributions and to further extend and constrain theories of word production.

However, we also wished to ensure that our findings could be directly compared to the previous literature, rather than add another study with differing methodology. Consequently, as Rabovsky et al. (2016) provide one of the most rigorous investigations, we first replicate their analysis approach in order to directly compare our data to theirs and to test the reliability of their findings. Second, we incrementally increase the complexity of the analysis to ultimately, simultaneously, examine the effects of six feature-based semantic variables (number of near semantic neighbours, semantic similarity, number of semantic features, typicality, intercorrelational density, distinctiveness) while also controlling for important psycholinguistic variables known to affect picture naming. Hence, a, preregistered (Open Science Framework: https://osf.io/yw6ma/, Lampe et al., 2019), threefold approach to the analysis was used which conceptually replicates and extends Rabovsky et al.'s study. The three separate analyses addressed the following two research questions:

1. *How is speech production affected by the number of semantic features*[1] *and intercorrelational density of the semantic representation of the target word?*

Analysis 1A: First, we conceptually replicated Rabovsky et al. (2016) and analysed effects of number of semantic features and intercorrelational density on naming latency and accuracy in English speaking participants, while controlling for concept familiarity, number of orthographic neighbours, and lexical frequency. In contrast to Rabovsky et al., we used only those McRae et al. (2005) items with high name agreement (in Australian English, as determined in Study 1). Moreover, we used an improved set of pictures (i.e., colour photographs instead of black and white pictures) and a slightly modified procedure (i.e., a familiarisation phase was used for half of the participants in the original study, while none of our participants were familiarised with any of the materials before the beginning of the task).

Analysis 1B: Subsequently, we sought to establish if any effects of number of semantic features and intercorrelational density in Analysis 1A were statistical artifacts due to insufficient control of psycholinguistic variables. For this purpose, we extended Rabovsky et al.'s (2016) analysis (and Analysis 1A) by controlling for the psycholinguistic variables that were identified to influence picture naming by Perret and Bonin (2019): name agreement, image agreement, imageability, age of acquisition, conceptual familiarity, and lexical frequency. Moreover, we added the ordinal category position of an item (e.g., Howard et al., 2006) and the item's trial number in the experiment as independent variables (e.g., Baayen & Milin, 2017).

[1] This was termed *featural richness* in Rabovsky et al. (2016) and our preregistration.

2. *How is speech production additionally affected by other semantic variables, specifically, the number of near semantic neighbours, semantic similarity, concept typicality, and distinctiveness, when the psycholinguistic variables in Analysis 1B are also controlled for? If effects of the proposed semantic variables are significant, are they facilitative or inhibitory?*

Analysis 2: In this analysis, we determined if, when studied simultaneously, any of the other feature-based semantic variables have effects on picture naming behaviour, and whether any effects of number of semantic neighbours and intercorrelational density are retained. We therefore analysed effects of the six semantic variables of interest simultaneously while controlling for the psycholinguistic variables detailed in Analysis 1B.

Study 1: Norming study

Word and picture characteristics are known to influence performance in object naming studies (e.g., Alario et al., 2004; Perret & Bonin, 2019). It is therefore important to account for such variables when studying effects of other variables of interest. Psycholinguistic control variables can be included in the analysis (e.g., as fixed effects), so that any variance in performance associated with them can be attributed to them and subtracted from the estimates of the variables of interest.

Some confounding variables are specific to the pictures or the population used in a study (e.g., name agreement) and can hence not be extracted from previously published norms. Therefore, in the absence of previous studies using these pictures, we collected normative data to be able to account for all psycholinguistic variables identified to influence picture naming behaviour in a Bayesian meta-analysis by Perret and Bonin (2019).

Methods

Participants

Name agreement, age of acquisition, and imageability ratings were collected from 45 Macquarie University undergraduate students (21 female; *M* = 20.2 years old, range = 18–33 years, *SD* = 3.1) who were recruited from Macquarie University's Psychology participant pool and participated for course credit. A different population of 48 undergraduate students (42 female; *M* = 20.9 years old, range = 17–35 years, *SD* = 4.9) provided image agreement ratings. All participants were native

Australian English speakers, right-handed and reported no history of neurological, cognitive, speech or language impairments. Participants were tested individually after giving written informed consent. The study was approved by Macquarie University's Human Ethics Committee.

Stimuli, Material, and Procedure

Colour photographs corresponding to all 541 items of the McRae et al. (2005) feature database were retrieved from the internet. If necessary, the pictures were edited, such that they showed the object on white background.

For the name agreement, age of acquisition, and imageability ratings, the pictures were divided into four lists containing 135 or 136 items each. Items from semantic categories were evenly distributed across the lists (e.g., buzzard, eagle, and vulture appeared in different lists). Items were presented in a randomised order, using Qualtrics software (Qualtrics, 2018).

Participants were asked to write the name of the pictures to obtain a measure of Australian name agreement (at least 11 participants per list). Subsequently, participants were presented with the items from one of the other lists as written words on the computer screen. In case of a word with ambiguous meaning (e.g., bat), a disambiguating description was provided together with the target word (e.g., bat (animal) vs bat (baseball)). Participants were asked to rate the age of acquisition and then the imageability of each item. The age of acquisition rating followed Johnston et al. (2010) and Gilhooly and Logie (1980): Participants were asked to estimate the age at which they thought they had first learned the name of the depicted object, choosing between seven age bands ranging from 0 to 13+ years. For the imageability rating, participants had to estimate the item's capacity to arouse a mental image of an object on a 7-point Likert scale, following Cortese and Fugett (2004), Toglia and Battig (1978), and Paivio et al. (1968).

Image agreement ratings were obtained in an online study also using Qualtrics software. Each participant saw on average 307 of the 541 pictures in a randomised order (range = 104–330 pictures, *SD* = 60.3). At least 22 participants rated each picture. The participants were asked to indicate the picture's image agreement following Snodgrass and Vanderwart (1980). Participants first saw a written word on the screen. Once ready, they advanced to a blank screen during which they were asked to create a mental image of the object they had read the name of. After 3 seconds, a single picture appeared on the screen and the participants were asked to judge how closely this picture matched their mental image of the object. Participants used a 5-point Likert scale, ranging from 1 "low agreement" to 5 "high agreement". Two additional options were given to indicate that the object was unknown or that the participant had thought of a different object when reading the word.

Analysis and Results

We will first detail how the item set for the picture naming study was selected based on the name agreement data. Subsequently, values of the other variables will be reported for that item subset with high name agreement.

As we were interested in the name associated with an item rather than the spelling accuracy of production, responses with typing/spelling errors were corrected (e.g., achordian for target word accordion) and considered as correct responses. Items where less than 75% of participants produced the same response to a picture were then excluded from further analyses, which led to removal of 196 items. We retained four items where participants agreed on a different name to that used in the McRae et al. (2005) American English target list. For all of these items, we considered the appropriate Australian English labels for the same concept (i.e., "motorbike" for the target word "motorcycle", "pram" for the target word "buggy", "prawn" for the target word "shrimp", and "teacup" for "cup").

From the remaining 345 items we discarded items when the agreed name:

1. was another target word (e.g., 82% of participants used the target word "church" to name the picture of a chapel).
2. was a superordinate (e.g., 91% of participants used "bird" to name the picture of a partridge).
3. referred to a different concept that was not another target (e.g., 91% used "wolf" to name the picture of a coyote).

This process led to a final set of 297 items. For each of these items, average age of acquisition, imageability, and image agreement ratings were calculated. For the image agreement ratings, responses where a participant had indicated that the word was unknown or that they had thought of a different object were excluded from the analysis. The characteristics of the 297 items are listed in Table 2.

Table 2

Descriptive statistics for name agreement (proportion of participants providing the target name) and rated variables (age of acquisition, imageability, and image agreement) for the final 297 item set

	Name agreement (proportion of participants)	Age of acquisition (7 age bands)	Imageability (7-point scale)	Image agreement (5-point scale)
Average	0.94	3.21	5.43	4.36
Minimum	0.75	1.25	2.09	2.91
Maximum	1.00	6.64	6.64	5.00
SD	0.07	1.03	0.73	0.42

Study 2: Picture naming experiment

A standard continuous picture naming experiment using simple picture naming without context manipulation was conducted to examine the effect of semantic variables on word production processes. The experimental procedure of this study and the three main analyses were preregistered on the Open Science Framework (Lampe et al., 2019; https://osf.io/yw6ma/) and the data and analyses scripts are available there. The materials are available on request from the corresponding author. The study was approved by the Macquarie University Human Ethics Committee.

Methods

Participants

Participants were recruited from Macquarie University's Psychology participant pool and received course credit or monetary compensation (AUD15 per hour) for their time. All participants provided informed consent. Participants were eligible to participate in this study if they were Australian English native speakers, were right-handed, 17–35 years old, and had normal or corrected-to-normal vision. Exclusion criteria were a history of neurological or cognitive impairments or a history of speech and language impairments. 89 participants took part in this study, of whom 2 were excluded because they did not fulfil the eligibility criteria or did not comply with the instructions. The data from 87 participants (68 female; *M* = 20.1 years old, range = 17–33 years, *SD* = 2.3) were therefore used in the analyses.

Stimuli

The stimuli consisted of pictures of the 297 items from the McRae et al. (2005) database that were retained following the norming process described in Study 1 above. Another feature database (Devereux et al., 2014) provided information on semantic categories. We used the categories assigned by Devereux et al. for the 267 of our items that appeared in this database and assigned the remaining items to one of Devereux et al.'s 24 semantic categories. As 45 of our items were allocated to the "miscellaneous" category by Devereux et al., we re-assigned these items to new categories where possible (e.g., 10 items (e.g., skis) to a category of "sport equipment", and 7 items (e.g., pipe) to a category of "plumbing"). This procedure led to 35 different semantic categories, containing between 1 and 32 items (*M* = 8.49 items per category).

Stimuli were divided into four blocks (Blocks 1–4). Block 1 was designed as a set of stimuli that would not be prone to effects of cumulative semantic inhibition (i.e., the finding that the production of a target word is slower the more items from the same semantic category were previously named in the experiment; e.g., Howard et al., 2006) by comprising items from different semantic categories. This block consisted of 35 items, nine of which came from the miscellaneous category and one item of each of the different semantic categories, under the condition that no items in this block were near semantic neighbours (feature vector similarity cosine between pairs of items < .4; Mirman, 2011). The remaining items were divided evenly between Blocks 2–4 (Blocks 2 and 3 *n* = 87, Block 4 *n* = 88 items). Three pseudorandomised orders of items within each of the four blocks were created. For Blocks 2–4 a minimum of two items from different semantic categories intervened between items of the same semantic category in order to reduce interference/facilitation from closely related items. Finally, six different experimental lists were created by manipulating the order and pseudorandomisations of the different blocks. Block 1, controlling for the cumulative semantic inhibition effect, always appeared in the first position in each list, while Blocks 2–4 were presented in varying orders afterwards (i.e., positions 2–4, see Appendix A). Each participant saw one of the pseudorandomised experimental lists.

For all 297 items, information was available for six psycholinguistic control variables as suggested by Perret and Bonin (2019) to influence picture naming: name agreement, image agreement, imageability, age of acquisition, familiarity (all values obtained from our norming study), and frequency (Zipf, based on television subtitles, SUBTLEX-UK; van Heuven et al., 2014). In addition, to replicate Rabovsky et al.'s (2016) analysis, the number of orthographic neighbours of the items was retrieved from the N-Watch database (Coltheart's N; Davis, 2005). Two further control variables were based on the experimental lists: A measure accounting for the number of previously seen items of the same semantic category to control for the cumulative semantic inhibition effect (Howard et al., 2006) and the rank-order of an item within the experiment to account for habituation to the experimental situation or fatigue (Baayen & Milin, 2010).

Information on six feature-based semantic predictor variables was derived from information given in McRae et al. (2005): number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness.

Number of semantic features was a simple count of the semantic features generated for a target word (e.g., Rabovsky et al., 2016). For intercorrelational density, the shared variance of a concept's correlated feature pairs was determined and then summed (e.g., Rabovsky et al., 2016). To qualify as a near semantic neighbour, the cosine feature vector similarity with the target had to be at least .4 (Hameau et al., 2019; Mirman, 2011; Mirman & Graziano, 2013). Following Mirman and Magnuson (2008), we operationalised semantic similarity as the average similarity between the feature vectors of the target and all other words in McRae et al.'s (2005) feature database (note that the only previous investigation of semantic similarity in word production (Fieder et al., 2019) excluded from the calculation concepts that had a feature vector similarity of 0 with the target. Other studies that have used measures of semantic similarity, are not relevant here as they have investigated different issues, e.g., the relationship between a target word and the stimulus context (e.g., Rose et al., 2019) or the semantic overlap between two concepts in a context (e.g., Rose & Abdel Rahman, 2017)). Typicality was calculated in a similar way to Rosch and Mervis' (1975) family resemblance score: Each feature of an item received a score based on the number of other items in the same semantic category that were credited with that particular feature. This feature weight was then divided by the number of items in the semantic category and ultimately multiplied by its production frequency (number of participants who produced that feature for the item) before the feature weights of all features of an item were summed. Finally, distinctiveness was the inverse of the number of concepts in which a certain feature occurs across the database, which was then averaged across the features of a concept (e.g., Rabovsky et al., 2016). See Appendix B for a more detailed description of the calculation of the semantic variables and Appendix C for a demonstration that the variability of the semantic variables in the item set selected for this study was comparable to the full feature database by McRae et al. (2005).

Procedure

Testing took place in a quiet room at Macquarie University. In addition to naming latency and accuracy, a continuous EEG signal was recorded using a 64 channels ActiveTwo BioSemi system (BioSemi, Amsterdam, The Netherlands). The EEG data are reported elsewhere (Lampe, Bürki, et al., 2021). The simple picture naming task analysed here was the first task in a larger study that involved the participants subsequently naming the same pictures twice more in different experimental paradigms (e.g., Lampe, Hameau, et al., 2021).

Picture presentation and trial-sequence were controlled by Presentation® software (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). Trial sequence was as follows: First a fixation cross appeared in the centre of the screen for a random duration of between 500 and 1000ms, such that participants would be unable to predict the exact onset of the picture. Next, a picture was displayed for 2000ms on a white background in the centre of the screen and the participants were instructed to name the picture as quickly and accurately as possible, using a single word only. After the picture offset the screen was blank for 1000ms before the start of the next trial.

The experiment was presented on a Dell Precision tower 3620 running Windows 10 and using an AOC FreeSync LED monitor. Verbal responses and response latencies were recorded using a voice trigger in Presentation® and a Behringer preamplifier (Tube Ultragain Mic100) together with a Rode NTG1 shotgun microphone. The keyboard was used to navigate through the experiment.

The experiment began with 5 practice trials in which participants named pictures that were not part of the 297 experimental stimuli and came from different semantic categories to the experimental stimuli. There was a break after the practice phase and after each experimental block for the participants to ask questions and to rest. The first trial after each break included another practice picture, which again came from a different semantic category from all experimental stimuli. The task took about 30 minutes.

Response coding

After the experiment, all responses were transcribed and coded for naming accuracy. Response latencies recorded by the voice trigger were checked and manually adjusted as necessary by visual and auditory inspection of the waveform using the program Praat (Version 6.0.49; Boersma & Weenink, 2019).

Accuracy of the first response was coded. Responses were coded as correct if the first response consisted of (only) the correctly pronounced target word. Only correct responses were considered for the naming latency analyses. For the accuracy analyses, in addition to correct responses, responses with a determiner preceding the correct name (e.g., “an octopus”), disfluencies on the target’s initial phoneme (e.g., “sss sofa”), hesitations (e.g.,”erm apple”), and elaborations (e.g., “yellow truck” for “truck”) were coded as correct. Responses consisting of a synonym or acceptable response (e.g., “sofa” for “couch”), abbreviations (e.g., “croc” for “crocodile”), or a response continuing from the previous item with subsequent correct naming of the target word (e.g., “emu .. bed” for “bed” with “emu” being the previous item) were coded as ‘not analysed’ (NA) and excluded from the analyses. Disfluencies with self-corrections (e.g., “oct squid” for “squid”), incomplete responses (e.g., “so” for “sofa”), incorrect responses (e.g., “chair” for “sofa”), as well as instances where the participants did not respond or made a comment that expressed a failure to respond (e.g., “I don’t know”; i.e., omissions) were coded as errors (see Appendix D, for examples and a summary of the coding system).

Data analysis

The data was analysed in RStudio (Version 1.2.5033; RStudio Team, 2019) using (generalised) linear mixed effect models ((G)LMMs) as implemented in the lme4-package (Version 1.1.21; Bates et al., 2017). *p*-values were derived using lmerTest (Version 3.1.1; Kuznetsova et al., 2017). For the specification of the random effects structure, we followed Bates et al. (2015) and the model fit of nested models was assessed using likelihood ratio tests (stats package, Version 3.6.1; R Core Team, 2018). Plots were created using the package sjPlot (Version 2.6.3; Lüdecke, 2019).

Ten data points from two participants had to be removed because of a programming error. Moreover, four items (board, bridge, racquet, and pie) were excluded from the analyses as many participants produced an elaboration that comprised a subordinate of the target word (e.g., tennis racquet). Outliers were identified for naming latency and accuracy, both for participants and items based on visual inspection of boxplots. Two participants who performed considerably less accurately than the other participants (69% and 62% naming accuracy versus mean accuracy of 86%, *SD* = 5%) were excluded from further analyses. Two items (crowbar and raft) were identified as outliers for naming accuracy (26% and 34% accuracy versus mean accuracy of 86%; *SD* = 15%) and excluded from further analyses.

For the naming latency analyses, only correct responses were considered and all other responses (3,782 data points, 15.30% of the data) were disregarded. 20,943 data points from 291 items and 85 participants entered the analyses, and mean naming latency was 900.81ms. For naming accuracy, responses coded as NA (synonyms, acceptable alternatives, abbreviations, and continuing responses; see section *Response coding*) were not analysed (171 data points, 0.69% of the data). 24,554 data points from 291 items and 85 participants entered the accuracy analyses, of which 85.69% were correct responses.

All predictor variables were standardised using a z-transformation[2]. Based on the output of the boxcox function (EnvStats Version 2.3.1; Millard, 2013), naming latency was negative reciprocally transformed to approximate a normal distribution (note that the negative reciprocal transformation preserves order among values of the same sign).

Three separate analyses on response latencies and naming accuracy were conducted:

---

[2] The standardisation of ordinal category position and item number in the experiment was not preregistered. However, standardisation of all variables was necessary to facilitate model convergence.

Analysis 1A: Replication of Rabovsky et al. (2016)*.* In this analysis, the number of semantic features and intercorrelational density were the only semantic variables included in the model. Following Rabovsky et al. (2016), we also included rated familiarity, number of orthographic neighbours, and lexical frequency. We also replicated Rabovsky et al.'s random effects structure by initially including random intercepts for participants and items as well as random slopes for participants for both semantic variables. However, random slopes were only retained in the models if they were supported by the data, following the model definition approach described by Bates et al. (2015).

Analysis 1B: Replication of Rabovsky et al. (2016) including more psycholinguistic control variables*.* In this analysis, the number of semantic features and intercorrelational density remained the only semantic predictor variables included in the model. However, in contrast to Rabovsky et al. (2016), we included a wider range of psycholinguistic control variables in the analysis. Following Perret and Bonin (2019) and Baayen and Milin (2017), the models included the control variables name agreement, image agreement, imageability, age of acquisition, familiarity, frequency, as well as a measure of ordinal category position and rank-order of an item within the list. Number of orthographic neighbours from Analysis 1A was not included. Again, random intercepts for participants and items were included in the models as well as random slopes for the two semantic variables by participants, which were retained if they were supported by the data.

Analysis 2: Extension of Rabovsky et al.'s (2016) analyses to include six semantic variables*.* In the final analyses, the models included four further semantic variables (number of near semantic neighbours, semantic similarity, typicality, and distinctiveness), in addition to number of semantic features and intercorrelational density and the psycholinguistic control variables described in Analysis 1B. Random intercepts for participants and items were included in the models as well as random slopes for semantic variables by participants if they were supported by the data.

Results

Correlations

Pearson's correlations between variables were determined (stats package Version 3.6.1; R Core Team, 2018) to investigate potential multicollinearity. Hutcheson and Sofroniou (1999) suggested that correlation coefficients of larger than .80 indicate multicollinearity. While most correlations between semantic predictor variables were significant (Table 3), the observed coefficients were not indicative of any problematic levels of collinearity between variables. In addition, we calculated variance inflation factors (VIF) for the fixed effects of each (G)LMM in the analyses to further identify any possibly problematic levels of multicollinearity between the variables entered in the individual analyses. The focus of those calculations were the semantic predictor variables, and not the control variables. VIFs for the semantic predictors in all models (Tables 4–6) were below the values that have been recommended as acceptable levels (depending on the author, VIFs > 2.5 (Allison, 2012), or around 5 (Hair et al., 2014; Rogerson, 2011) indicate potentially problematic multicollinearity). Hence, multicollinearity was not a problem between the predictors of the models used in the analyses.

Table 3

Pearson's correlations between the semantic predictor variables and the psycholinguistic control variables

Variable	Number near semantic neighbours	Semantic similarity	Number semantic features	Typicality	Intercorrelational density	Distinctiveness
Semantic similarity	0.58***					
Number semantic features	-0.08	-0.01				
Typicality	0.37***	0.14*	0.25***			
Intercorrelational density	0.42***	0.07	0.40***	0.44***		
Distinctiveness	-0.51***	-0.58***	0.03	-0.25***	-0.41***	
Name agreement	-0.07	-0.04	-0.01	-0.05	-0.07	0.04
Age of acquisition	0.05	0.04	-0.20***	-0.14*	-0.19***	0.09
Imageability	-0.05	-0.08	0.24***	0.12*	0.16**	-0.02
Image agreement	0.31***	0.24***	0.03	0.04	0.12*	-0.19***
Frequency	-0.22***	-0.11	0.24***	0.03	0.09	0.09
Familiarity	-0.27***	-0.37***	0.15*	0.12*	0.05	0.20***
Number orthographic neighbours	-0.12*	0.03	0.03	-0.04	-0.04	0.05

*** p < .001, ** p < .01, * p < .05.

Analysis 1A: Replication of Rabovsky et al. (2016)

Naming latency. This analysis replicated the analysis run by Rabovsky et al. (2016) and, therefore, we fitted a model including the same control and semantic variables as in that paper. Following Bates et al.'s (2015) approach to define a random effects structure that is supported by the data, we identified a model with number of semantic features as random slope for participants, including correlations between the random slope and the intercept, as the model explaining the data best (Table 4, Model 1A.1). Responses were faster for words that were more familiar and had higher word frequency, however, the two semantic variables number of semantic features and intercorrelational density did not reach significance (see also Figure E1, Panel A, for a graphical display of the fixed effects and their confidence intervals).

Naming accuracy. To replicate the analysis of Rabovsky et al. (2016), we used a model with the same fixed effects structure as for the naming latency analysis of Analysis 1A.1. A model with crossed random intercepts for items and participants and no random slope provided the best fit of the data (Table 4, Model 1A.2). Increasing familiarity and frequency were found to facilitate naming accuracy. Moreover, response accuracy was facilitated by a higher number of semantic features (see also Figure E1, Panel B).

Table 4

Analysis 1A replicating Rabovsky et al. (2016): Summarised output of linear mixed model analysis of picture naming latency (Models 1A.1) and naming accuracy (Model 1A.2)

	Naming latency (Model 1A.1)						Naming accuracy (Model 1A.2)					
	lmer(RT ~ Familiarity + OrthNeigh + Frequency + NoFeats + IntercorrDens + (1 | Item) + (1 + NoFeats | Participant), data = data, REML = FALSE)						glmer(Accuracy ~ Familiarity + OrthNeigh + Frequency + NoFeats + IntercorrDens + (1 | Item) + (1 | Participant), data = data, family = binomial)					
Random effect	Variance	*SD*	Correlation				Variance	*SD*				
Item (Intercept)	0.02	0.16					2.46	1.57				
Participant (Intercept)	0.01	0.12					0.28	0.53				
Participant | NoFeats	0.00	0.01	0.54									
Residuals	0.05	0.23										
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF	Estimate	*SE*	*CI*	*z*-value	*p*-value	VIF
Intercept	-1.17	0.02	-1.20 – -1.14	-73.56	< .001		2.76	0.12	2.53 – 2.98	23.79	< .001	
Familiarity	-0.06	0.01	-0.08 – -0.04	**-5.69**	**< .001**	1.25	0.38	0.11	0.16 – 0.59	**3.44**	**< .001**	1.24
OrthNeigh	0.02	0.01	-0.00 – 0.04	1.85	.066	1.17	-0.13	0.11	-0.34 – 0.08	-1.19	.233	1.17
Frequency	-0.03	0.01	-0.05 – -0.00	**-2.38**	**.018**	1.47	0.48	0.12	0.25 – 0.72	**4.03**	**< .001**	1.46
NoFeats	-0.02	0.01	-0.04 – 0.00	-1.54	.125	1.25	0.23	0.11	0.01 – 0.44	**2.06**	**.039**	1.25
IntercorrDens	0.01	0.01	-0.01 – 0.03	0.76	.448	1.19	-0.19	0.11	-0.39 – 0.02	-1.77	.076	1.20
	Observations: 20,943						Observations: 24,554					
	Marginal R^2 / Conditional R^2: 0.058 / 0.457						Marginal R^2 / Conditional R^2: 0.089 / 0.503					

Note. VIF = Variance Inflation Factor, OrthNeigh = Number of orthographic neighbours, NoFeats =number of semantic features, IntercorrDens = intercorrelational density, Participant | X = random slope of X by participants.

Values of significant effects (p < .05) are printed in bold.

Analysis 1B: Replication of Rabovsky et al. (2016) including more psycholinguistic control variables

Naming latency*.* We extended the original Rabovsky et al. (2016) analyses by including more control variables, while the semantic variables remained the same. Again, a model with a by-participants random slope for number of semantic features, including the correlation between the random slope and intercept, provided the best fit of the data (Table 5, Model 1B.1).

Most control variables had a significant effect on response latencies and their effects pointed in the expected directions: Responses were faster for words with higher name agreement, higher image agreement, and higher familiarity, and responses were slower for words acquired later in life and the more items from the same category that had been previously seen in the experiment (Cumulative Semantic Interference effect). Finally, responses became slower over the course of the experiment. Neither of the two semantic variables showed significant effects on naming latency (Table 5, Figure E2, Panel A).

Naming accuracy*.* In replicating Rabovsky et al.'s (2016) experiment, we extended Analysis 1A.2 with further control variables (as in Analysis 1B.1). The full complex effect structure was not supported by the data, and a model without by-participant random slopes provided the best fit (Table 5, Model 1B.2).

All control variables reached significance, except for familiarity, and their effects were in the expected directions: Higher name agreement, image agreement, imageability, and frequency facilitated naming accuracy, while higher age of acquisition and a higher number of items from the same semantic category previously seen in the experiment (Cumulative Semantic Interference effect) led to less accurate responses. Moreover, participants were more accurate as the experiment progressed. The two semantic variables were also significant: While higher numbers of features improved naming accuracy, higher intercorrelational density reduced naming accuracy (Table 5, Figure E2, Panel B).

Table 5

Analysis 1B replicating Rabovsky et al. (2016) taking more psycholinguistic control variables into account: Summarised output of linear mixed model analysis of picture naming latency (Models 1B.1) and naming accuracy (Model 1B.2)

	Naming latency (Model 1B.1)						Naming accuracy (Model 1B.2)					
	lmer(RT ~ NameAgr + ImageAgr + Imageability + AoA + Familiarity + Frequency + OrdCatPos + Order + NoFeats + IntercorrDens + (1 | Item) + (1 + NoFeats | Participant), data, REML = FALSE)						glmer(Accuracy ~ NameAgr + ImageAgr + Imageability + AoA + Familiarity + Frequency + OrdCatPos + Order + NoFeats + IntercorrDens + (1 | Item) + (1 | Participant), data, family = binomial)					
Random effect	Variance	*SD*	Correlation				Variance	*SD*				
Item (Intercept)	0.02	0.13					1.43	1.20				
Participant (Intercept)	0.01	0.12					0.28	0.53				
Participant | NoFeats	0.00	0.01	0.53									
Residuals	0.05	0.23										
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF	Estimate	*SE*	*CI*	*z*-value	*p*-value	VIF
Intercept	-1.17	0.01	-1.20 – -1.14	-78.68	< .001		2.73	0.10	2.53 – 2.92	27.50	< .001	
NameAgr	-0.06	0.01	-0.07 – -0.04	**-6.99**	**< .001**	1.11	0.72	0.08	0.57 – 0.88	**9.11**	**< .001**	1.08
ImageAgr	-0.05	0.01	-0.07 – -0.04	**-6.44**	**< .001**	1.17	0.34	0.08	0.18 – 0.51	**4.04**	**< .001**	1.20
Imageability	-0.02	0.01	-0.04 – 0.00	-1.91	.058	1.82	0.26	0.11	0.05 – 0.47	**2.46**	**.014**	1.78
AoA	0.03	0.01	0.01 – 0.05	**2.53**	**.012**	2.30	-0.31	0.12	-0.54 – -0.07	**-2.59**	**.010**	2.27
Familiarity	-0.04	0.01	-0.06 – -0.02	**-4.44**	**< .001**	1.44	0.14	0.09	-0.04 – 0.32	1.54	.122	1.43
Frequency	-0.02	0.01	-0.03 – 0.00	-1.59	.114	1.54	0.27	0.10	0.08 – 0.46	**2.80**	**.005**	1.53
OrdCatPos	0.01	0.00	0.00 – 0.02	**3.17**	**.002**	3.11	-0.13	0.05	-0.24 – -0.03	**-2.59**	**.010**	3.06
Order	0.01	0.00	0.00 – 0.01	**2.20**	**.028**	3.08	0.11	0.05	0.02 – 0.20	**2.52**	**.012**	3.01
NoFeats	-0.02	0.01	-0.03 – 0.00	-1.70	.089	1.28	0.21	0.09	0.04 – 0.38	**2.38**	**.017**	1.29
IntercorrDens	0.01	0.01	-0.00 – 0.03	1.45	.147	1.24	-0.22	0.09	-0.39 – -0.06	**-2.61**	**.009**	1.27
	Observations: 20,943						Observations: 24,554					
	Marginal R^2 / Conditional R^2: 0.139 / 0.454						Marginal R^2 / Conditional R^2: 0.231 / 0.494					

Note. VIF = Variance Inflation Factor, NameAgr = name agreement, ImageAgr = image agreement, AoA = age of acquisition, OrdCatPos = ordinal category position, NoFeats = number of semantic features, IntercorrDens = intercorrelational density, Participant | X = random slope of X by participants.

Values of significant effects ($p < .05$) are printed in bold.

Analysis 2: Extension of Rabovsky et al.'s (2016) analyses to include six semantic variables

Naming latency. In this analysis, we added the four additional semantic variables number of near semantic neighbours, semantic similarity, typicality, and distinctiveness to Analysis 1B.1. A model with random by-participant slopes for semantic similarity, number of near semantic neighbours, and number of semantic features, including correlations between the slopes and with the intercept provided the best fit of the data (Table 6, Model 2.1).

Effects of the control variables (see Table 6) were comparable to the effects found in Analysis 1B.1 with faster responses for words with higher name agreement, higher image agreement, and higher familiarity. Responses were slower for words with a higher age of acquisition, the more items from the same category were previously seen in the experiment (Cumulative Semantic Interference effect), and responses slowed over the course of the experiment. Importantly, the number of semantic features, as well as the distinctiveness of an item, significantly predicted naming latency: Responses were faster for items with more features and slower the higher the distinctiveness of an item.

These findings are displayed in Figure 1 (Panel A), where red lines (left of centre) indicate a facilitatory effect of a variable with faster responses as that variable increases in value, while blue lines (right of centre) indicate an inhibitory effect with slower responses as that variable increases in value. Confidence intervals that cross the black zero line are indicative of non-significant effects.

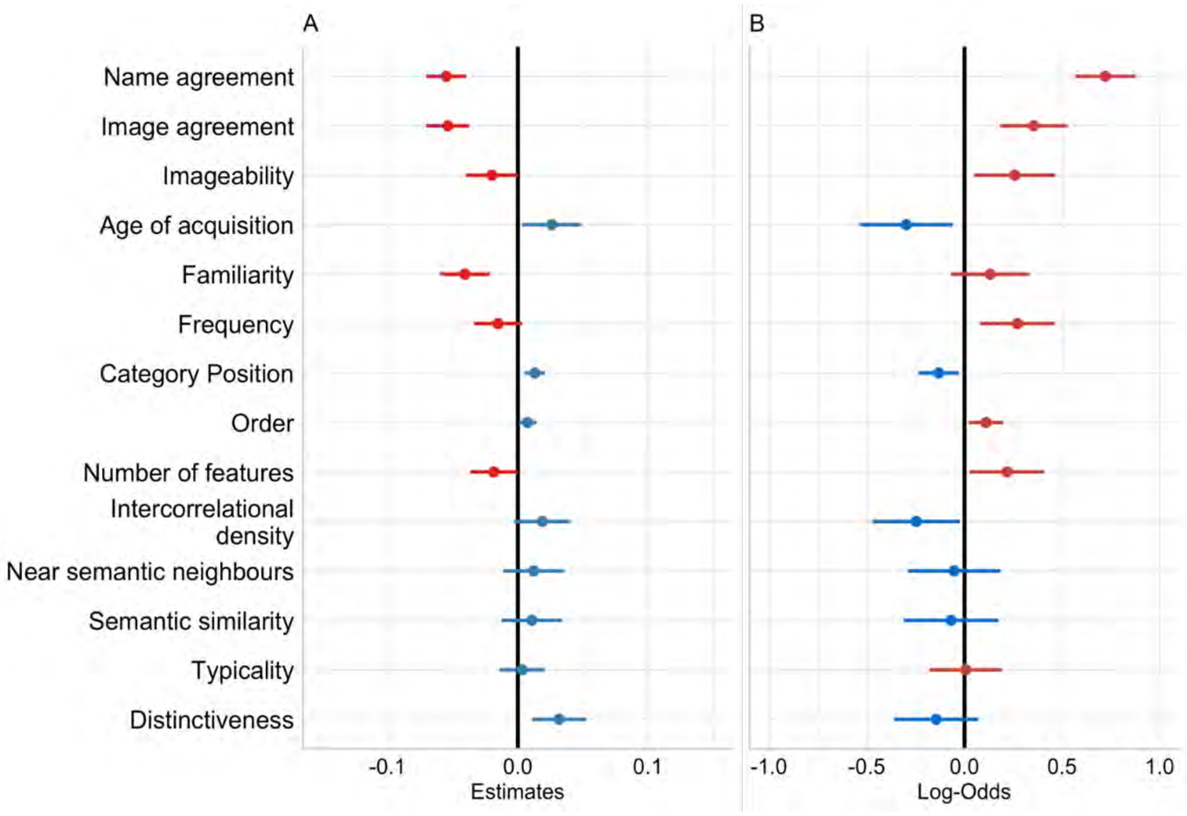
Naming accuracy. When the four additional semantic variables number of near semantic neighbours, semantic similarity, typicality, and distinctiveness were added to Analysis 1B.2, a model with a by-participant random slope for number of near semantic neighbours and its correlation with the intercept provided the best fit of the data (Table 6, Model 2.2).

Effects for the control variables (Table 6) remained comparable to Analysis 1B.2, with higher name agreement, image agreement, imageability, and frequency facilitating naming accuracy, and higher age of acquisition and a higher number of items from the same semantic category previously seen in the experiment (Cumulative Semantic Interference effect) leading to more naming errors. Moreover, participants' accuracy increased over the course of the experiment. Only two semantic variables reached significance: Higher numbers of semantic features led to more correct responses, while higher intercorrelational density led to reduced naming accuracy.

The findings are plotted in Figure 1 (Panel B), where red lines (right of centre) indicate a facilitatory effect with more accurate responses with higher values of a variable, while blue lines (left of centre) indicate an inhibitory effect with reduced naming accuracy with higher values of a variable. Again, confidence intervals that cross the black zero line are indicative of non-significant effects.

Figure 1

Analysis 2 extending Rabovsky et al. (2016) with other semantic variables: Fixed effects estimates with 95% confidence interval of picture naming latency analysis (Panel A; Model 2.1) and accuracy analysis (Panel B; Model 2.2)



Note. Panel A shows the output of the naming latency analysis and Panel B the output of the naming accuracy analysis; red lines (to the left of centre for latency, and right for accuracy) indicate increased values of the variable lead to better performance, and blue lines (to the right of centre for latency and left for accuracy) indicate worse performance.

Table 6

Analysis 2 extending Rabovsky et al. (2016) with other semantic variables: Summarised output of linear mixed model analysis of picture naming latency (Model 2.1) and naming accuracy (Model 2.2)

	Naming latency (Model 2.1)						Naming accuracy (Model 2.2)					
	lmer(RT ~ NameAgr + ImageAgr + Imageability + AoA + Familiarity + Frequency + OrdCatPos + Order + NoFeats + IntercorrDens + NrSemNeigh + SemSim + Typicality + Distinct + (1 | Item) + (1 + NoFeats + NrSemNeigh + SemSim | Participant), data = data, REML = FALSE)						glmer(Accuracy ~ NameAgr + ImageAgr + Imageability + AoA + Familiarity + Frequency + OrdCatPos + Order + NoFeats + IntercorrDens + NrSemNeigh + SemSim + Typicality + Distinct + (1 | Item) + (1 + NrSemNeigh | Participant), data = data, family = binomial)					
Random effect	Variance	*SD*	Correlation				Variance	*SD*	Correlation			
Item (Intercept)	0.02	0.12					1.42	1.19				
Participant (Intercept)	0.01	0.12					0.28	0.53				
Participant | NoFeats	0.00	0.01	0.50									
Participant | NrSemNeigh	0.00	0.02	-0.24	0.02			0.02	0.15	0.34			
Participant | SemSim	0.00	0.01	0.07	-0.61	-0.49							
Residuals	0.05	0.23										
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF	Estimate	*SE*	*CI*	*z*-value	*p*-value	VIF
(Intercept)	-1.17	0.02	-1.20 – -1.14	-78.43	< .001		2.73	0.10	2.54 – 2.92	27.67	< .001	
NameAgr	-0.06	0.01	-0.07 – -0.04	**-7.07**	**< .001**	1.11	0.72	0.08	0.57 – 0.87	**9.13**	**< .001**	1.08
ImageAgr	-0.05	0.01	-0.07 – -0.04	**-6.37**	**< .001**	1.27	0.35	0.09	0.18 – 0.53	**4.03**	**< .001**	1.30
Imageability	-0.02	0.01	-0.04 – -0.00	-1.96	.051	1.86	0.26	0.11	0.05 – 0.47	**2.43**	**.015**	1.84
AoA	0.03	0.01	0.00 – 0.05	**2.24**	**.026**	2.42	-0.30	0.12	-0.54 – -0.06	**-2.45**	**.014**	2.44
Familiarity	-0.04	0.01	-0.06 – -0.02	**-4.19**	**< .001**	1.74	0.13	0.10	-0.07 – 0.33	1.29	.198	1.75
Frequency	-0.02	0.01	-0.03 – 0.00	-1.61	.108	1.58	0.27	0.10	0.08 – 0.46	**2.78**	**.005**	1.57
OrdCatPos	0.01	0.00	0.01 – 0.02	**3.25**	**.001**	3.18	-0.13	0.05	-0.24 – -0.03	**-2.52**	**.011**	3.14
Order	0.01	0.00	0.00 – 0.01	**2.16**	**.030**	3.13	0.11	0.06	0.02 – 0.20	**2.43**	**.015**	3.05

NoFeats	-0.02	0.01	-0.04 – -0.00	**-1.99**	**.048**	1.55	0.22	0.10	0.03 – 0.41	**2.24**	**.025**	1.57
IntercorrDens	0.02	0.01	-0.00 – 0.04	1.71	.088	2.20	-0.25	0.11	-0.47 – -0.02	**-2.17**	**.030**	2.26
NrSemNeigh	0.01	0.01	-0.01 – 0.04	1.02	.307	2.44	-0.05	0.12	-0.29 – 0.18	-0.45	.657	2.47
SemSim	0.01	0.01	-0.01 – 0.03	0.90	.369	2.50	-0.07	0.12	-0.31 – 0.17	-0.56	.577	2.51
Typicality	0.00	0.01	-0.01 – 0.02	0.37	.713	1.44	0.01	0.09	-0.18 – 0.19	0.08	.940	1.49
Distinct	0.03	0.01	0.01 – 0.05	**2.98**	**.003**	2.04	-0.15	0.11	-0.36 – 0.07	-1.32	.186	2.03
	Observations: 20,943						Observations: 24,554					
	Marginal R^2 / Conditional R^2: 0.143 / 0.456						Marginal R^2 / Conditional R^2: 0.231 / 0.495					

Note. VIF = Variance Inflation Factor, NameAgr = name agreement, ImageAgr = image agreement, AoA = age of acquisition, OrdCatPos = ordinal category position, NoFeats = number of semantic features, IntercorrDensity = intercorrelational density, NrSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity, Distinct = distinctiveness, Participant | X = random slope of X by participants.

Values of significant effects ($p < .05$) are printed in bold.

Discussion

Understanding how word-inherent semantic properties of an item affect word production is critical for improving our understanding of the mechanisms and refining theoretical models of spoken word production. However, previous research investigating effects of semantic variables on production has had numerous methodological and statistical limitations, which may have distorted the findings. Here, we overcame these shortcomings and investigated the role of six feature-based semantic variables on the word production process: number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness, while controlling for other psycholinguistic variables that have been found to influence word production.

The design of our study also enabled direct comparison with previous literature through a conceptual replication of Rabovsky et al. (2016) that evaluated the reliability of their findings. We therefore investigated effects of number of semantic features and intercorrelational density on word production under reduced (Analysis 1A and as in Rabovsky et al.) and subsequently under more adequate control of psycholinguistic variables (Analysis 1B), before investigating the effects of the six semantic variables of interest simultaneously (Analysis 2).

Table 7 summarises the results of our analyses. Taken together, number of semantic features was the most important and consistent predictor of naming accuracy across analyses and there was also a relatively weak effect on naming latency in the most complex analysis (Analysis 2). Moreover, there were effects of intercorrelational density on naming accuracy (Analyses 1B and 2) and an effect of distinctiveness on latency (Analysis 2). Importantly, there was a lack of evidence for effects of number of near semantic neighbours, semantic similarity, or typicality on naming. Below, we first discuss possible reasons for the relatively restricted effects of semantic variables that were found, then we relate our results to those of Rabovsky et al. (2016) and discuss the changes in effects in the different analyses before turning to the theoretical interpretation of individual significant effects.

Table 7

Summary of the findings for semantic variables across all analyses

Semantic variable	Rabovsky et al. (2016)		Analysis 1A		Analysis 1B		Analysis 2	
	latency	accuracy	latency	accuracy	latency	accuracy	latency	accuracy
NoFeats	↗	↗	Ø	↗	Ø	↗	↗	↗
IntercorrDens	↙	↙	Ø	Ø	Ø	↙	Ø	↙
NearSemNeigh							Ø	Ø
SemSim							Ø	Ø
Typicality							Ø	Ø
Distinct							↙	Ø

Note. Ø = non-significant effect, ↙ = poorer performance (slower responses and decreased accuracy with higher values of the semantic variable), ↗ = improved performance (faster responses and higher accuracy with higher values of the semantic variable), blank cells were not investigated, NoFeats = number of semantic features, IntcorrDens = intercorrelational density, NearSemNeigh = number of near semantic neighbours, SemSim = semantic similarity, Distinct = distinctiveness.

Limited effects of some semantic variables

Over and above the effects of number of semantic features and intercorrelational density, as also investigated by Rabovsky et al. (2016), distinctiveness was the only other semantic variable to exert a significant effect on the participants' naming performance (Analysis 2, Tables 6 and 7). That is, there was no evidence that picture naming performance was affected by number of near semantic neighbours, semantic similarity, and typicality.

The absence of effects of these variables is in contrast to most previous work, where significant effects for these semantic variables were reported. This reinforces the validity of the concern, raised in the Introduction, that significant effects in some of the previous studies could have been Type 1 errors, false positives, arising due to a lack of control of other semantic and psycholinguistic variables (see below for more detail). However, particularly for effects of number of semantic neighbours and semantic similarity, differences in significance could also be due to task differences (i.e., speeded vs simple picture naming). Moreover, as evident in Table 1, compared to our study, most previous reports had fewer data points, due to smaller samples of participants and/or items, leading to lower statistical power and reliability of the results. In addition, differences in the statistical approach (i.e., taking into account the variability induced by specific participants and items using mixed effects models or disregarding such differences in regression analyses; using semantic variables as continuous or categorical predictors) may have caused differences in the significance of effects of semantic variables. Finally, of course, there is also the very real concern that this literature, as is common, suffers from publication bias towards significant findings.

Importantly, our null effects of the number of near semantic neighbours are in line with previous work using a simple picture naming paradigm (Hameau et al., 2019; Lampe et al., 2017); the studies that found significant effects used a speeded naming paradigm. Similarly, not all previous studies have found significant effects of typicality (Morrison et al., 1992; Woollams, 2012).

Development of effects across analyses

We now review how the effects of semantic variables changed across the analyses. We do not comment on marginally significant findings and only focus on significant effects (p < .05).

Conceptual replication of Rabovsky et al. (2016)*.* As is summarised in Tables 1 and 7, Rabovsky et al. (2016) found significant effects of both number of semantic features and intercorrelational density on picture naming latency and accuracy. In Analysis 1A, where we replicated Rabovsky et al.'s analyses by taking the same semantic and control variables into account, we replicated their facilitatory effect of number of semantic features on naming accuracy, while the effect of intercorrelational density was non-significant. Moreover, we did not replicate Rabovsky et al.'s results for naming latency: Neither number of features nor intercorrelational density showed significant effects on naming latency. However, in the more complex analyses, particularly Analysis 2 that accounted for further control and semantic variables, our findings were broadly consistent with those of Rabovsky et al.: Number of semantic features affected both naming latency and accuracy (facilitatory) and intercorrelational density affected accuracy (inhibitory). Hence, the only effect reported by Rabovsky et al. that failed to reach significance in any of our analyses was the effect of intercorrelational density on naming latency. However, in Rabovsky et al.'s own work this effect also did not prove to be robust: While the group effect was significant in Rabovsky et al. (2016), few individual participants were significantly affected, and in Rabovsky et al. (2021) the effect was non-significant when the participants first named the pictures.

What could be the cause of the differences between Rabovsky et al.'s (2016) results and those of our Analysis 1A? Given that we had almost three times as many data points as the original study (24,725 data points vs 8,656 data points), our failure to replicate Rabovsky et al. (and other studies reported in Table 1) cannot be due to lack of statistical power relative to the original study; if anything, our findings should be more reliable. The fact that effects that are in line with Rabovsky et al. emerged in our subsequent analyses may suggest that the effects of semantic variables are relatively weak and subject to a certain degree of variability. By including more control variables in the models of Analysis 1B and 2 we statistically controlled some of this variability, which resulted in significant effects of more semantic variables (further discussion of this in the next section). Although the difference in the language spoken by the participants (English vs German) may have influenced the results, as the semantic variables are argued to reflect our underlying conceptual representation and lexical processing, there is no reason to expect that these factors should operate differently across languages. However, there are other methodological factors that may have been important: For example, we ensured that name agreement of the experimental items was high for our population, the features were generated in the same language as the experiment, and we used colour photographs, which can improve concept recognition for some items (e.g., lime vs lemon; Bonin et al., 2019; Rossion & Pourtois, 2004). We would argue that these differences are, once again, likely to make the results of our study more reliable but may also have led to differences in the outcomes of the two studies.

Changes in effects across analyses. As summarised in Table 7, the significance of some semantic variables changed between analyses: The most significant effects of semantic variables were found in the most complex analysis (Analysis 2). Why do more semantic variables effects reach significance with improved statistical control of factors affecting naming? It is well known that (inter-) correlations between variables can cause the effect of one variable to be significant only when another variable is not included in the analysis. Moreover, mediator or suppressor variables can also influence the findings, potentially leading to disappearance, overestimation, or underestimation of the effect of an independent variable on the dependent variable (Type 1 or 2 errors) (Hair et al., 2014). The inclusion of such variables in the model can increase the regression coefficient of another variable in the model by removing irrelevant variance from that predictor (e.g., Pandey & Elliott, 2010)[3]. Therefore, it is vital even if a certain influential variable is of no direct interest to the experiment, that it is included in the analysis as the omission of such confounding variables can substantially distort the effects of other variables and lead to spurious results. By keeping effects of independent variables (e.g., psycholinguistic control variables) constant by including them in the statistical analysis, we can study maximally *pure* contributions of the variables of interest.

In our analyses, most effects were significant under conditions of maximal control (Analysis 2). This suggests that in Analyses 1A and 1B some effects were *underestimated* due to uncontrolled suppressor variables or correlations with variables that were not included in the analyses. As explained above, any effects of semantic variables emerging under more rigorous control of other variables can be considered more reliable, because there is less free variance that can, spuriously, be taken up by the variables of interest. Therefore, it seems likely that effects of the semantic variables are rather weak and therefore only able to cross the significance threshold when the variance of other variables is controlled.

Theoretical explanations for significant effects

The effects of the semantic variables that we observed constrain the architecture of word production models, as, to be considered an adequate model of word production, a theory has to be able to account for all these effects. Below we discuss the theoretical explanations of the significant

[3] For example, both familiarity and imageability affect picture naming (higher values making picture naming more accurate and faster) but effects of frequency tend to be stronger. Imagine the situation where in a set of items more familiar words tend to be of lower imageability (there is a negative correlation). If the effect of each variable is examined independently (in a correlation for example), the two variables are in competition. There may be an effect of familiarity, the variable with the strongest effects. However, as the words of higher familiarity are of lower imageability, the effect of imageability will be masked: Familiarity is a suppressor variable. Only when both variables are analysed together and the independent variance of each can be evaluated will a significant effect of imageability become apparent (assuming there is sufficient independent variance).

effects found here—effects of number of semantic features, intercorrelational density, and distinctiveness—in the context of current theories of word production.

Number of semantic features. Number of semantic features represented a simple count of the features of a concept according to the McRae et al. (2005) feature database and was the most important predictor of naming accuracy and significant in all three of our analyses. Moreover, in the most complex analysis (Analysis 2), there was a weak effect of number of semantic features on naming latency. Naming was more accurate and faster for words with more semantic features, which is in agreement with previous work by Rabovsky et al. (2016, 2021) and Taylor et al. (2012).

Using a neural network model, Rabovsky and McRae (2014; Simulation 2) simulated processing of words with high or low numbers of semantic features. For words with a higher number of semantic features, there was higher semantic activation, corresponding to the summed activation across all semantic features of an item, which Rabovsky and McRae argued could explain facilitatory behavioural effects of number of semantic features. In word production, this increased semantic activation is argued to result in stronger activation of the target word's lexical representation, facilitating its selection and leading to faster and more accurate responses (see also Rabovsky et al., 2016). Current theories of word production would predict facilitated naming of targets with stronger lexical activation (e.g., Abdel Rahman & Melinger, 2009; Dell, 1986; Howard et al., 2006; Levelt et al., 1999; Oppenheim et al., 2010). Simultaneously, though, a higher number of semantic features may result in the co-activation of a larger number of semantically related lexical representations as these features will be represented in other concepts (e.g., Rabovsky et al., 2016). However, our finding of a facilitatory effect of number of semantic features suggests that any increased lexical competition from co-activated representations (in theories that incorporate competitive mechanisms) is outweighed by the conceptual facilitation from the many semantic features. In contrast, concepts with a lower number of semantic features will result in relatively less activation of that concept's lexical representation, with selection therefore being more error prone and slower.

The facilitatory effect from higher numbers of semantic features may also be compatible with other theoretical frameworks: Assuming holistic concepts, the measure of number of semantic features may capture the number of links between a target concept and other concepts. Activation would spread along those links, which ultimately facilitates target processing. Moreover, in the context of an attractor network, concepts with richer semantic representations (e.g., higher numbers of features) build stronger attractor basins, which enable the system to settle faster and more accurately into a stable pattern of activation (Plaut & Shallice, 1993; see also Pexman et al., 2007).

Intercorrelational Density. Intercorrelational density had an inhibitory effect on naming accuracy, as found previously (Rabovsky et al., 2016, 2021): Participants made significantly more errors on words with higher intercorrelational density. Such words have feature pairs that tend to co-occur across concepts and therefore share a lot of variance. Hence, intercorrelational density measures how strongly the features of a concept cluster together. For example, the features *has fur* and *has four legs* also tend to occur with the features *has a tail*, *has ears*, etc. (McRae & Cree, 2002). Features in such intercorrelated clusters have been proposed to boost each other's activation (e.g., via bidirectional links between the features; McRae, 2004; McRae et al., 1997) and simulations (McRae et al., 1997) and previous experimental work (McRae et al., 1999; see also Taylor et al., 2004) have found that correlations between semantic features *facilitate* conceptual processing. However, in word production, intercorrelational density may also determine the size of the co-activated lexical cohort and the strength of its activation as it reflects the activation that is shared between the target concept and related concepts (Rabovsky et al., 2016).

In McRae et al. (2005) intercorrelational density is calculated as the *summed* shared variance of a concept's correlated feature pairs. Therefore, higher intercorrelational density can arise in two ways: 1) A concept has *many* feature pairs that are at least correlated just above threshold (6.5% of shared variance) or, 2) a concept has at least a few feature pairs that are *strongly* correlated and share considerable variance. In the first scenario (1), in addition to the target's lexical representation, lexical representations of many other concepts that share the (slightly) correlated semantic features with the target word would be (slightly) co-activated, with the extent of the co-activation depending on the strength of the intercorrelation of their features. It is important to note that in our analyses (but not in all the others in the literature), this is *not* confounded with the number of semantic features or number of near semantic neighbours as we controlled for these measures. In the second scenario (2), where an item has a small number of strongly correlated pairs, a different pattern emerges with lexical representations of few other concepts receiving activation. However, their co-activation would be stronger due to the stronger correlations among the feature pairs, making them stronger candidates for selection. These examples clearly illustrate that whether a large number of less intercorrelated concepts or a small number of highly intercorrelated concepts are co-activated critically depends on the number of concepts in which the pairs of correlated features occur and on the strength of their correlation. This not only suggests that the consequences of similar intercorrelational density values for lexical activation can vary between items, but that across sets of items the precise effects of intercorrelational density may be unpredictable depending on the balance of scenario 1 and scenario 2 items. In our analyses, given the absence of evidence for an effect of number of near semantic neighbours, it is likely that the stronger co-activation of closely semantically related representations (scenario 2) is the driving force underlying the negative effects of intercorrelational feature density.

Similarly, intercorrelational density may represent the strength of the labelled links between a target and other holistic lexical concepts. In that case, spreading activation along those links may cause the co-activation of various concepts and their lexical representations. Irrespective of the type of semantic representation, assuming bidirectional links between conceptual and lexical levels of word production (e.g., Abdel Rahman & Melinger, 2009; Dell, 1986), co-activated lexical representations will mutually increase their activation. This mutual increase in activation is particularly powerful for co-activated words that are closely related, due to the strength of the correlation of their features.

Yet, how do these co-activated representations negatively affect the production of the target word? The account outlined above of a negative effect of intercorrelational density on accuracy assumes a theoretical framework incorporating competitive lexical selection: During lexical selection, the co-activated lexical representations compete with the target word for selection. For words with higher intercorrelational density this competition is stronger, particularly when it arises from more closely related co-activated representations (e.g., Abdel Rahman & Melinger, 2019; Rose & Abdel Rahman, 2017). The stronger competition negatively impacts the selection of the target's lexical representation, leading to increased chances of incorrectly selecting a co-activated representation, which can ultimately cause more naming errors. The precise mechanism underlying this competition is still unclear and different ways have been suggested to implement it (e.g., Luce choice mechanism, e.g., Luce, 1959; Abdel Rahman & Melinger, 2009, 2019; Levelt et al., 1999; Roelofs, 1997; lateral inhibitory links between active representations at the lexical level, e.g., Howard et al., 2006; Caramazza, 1997; McClelland & Rumelhart, 1981). However, regardless of the details of the competitive mechanism, the presence of more or more strongly co-activated representations will result in poorer performance.

On the other hand, Oppenheim et al.'s (2010) learning mechanism may also be able to explain the poorer performance on items with higher intercorrelational density. According to Oppenheim et al., interference is due to implicit incremental learning, which causes weakening of the connections between the semantic representation of the target and the lexical representations of co-activated but unnamed alternatives after retrieval of the target word. Each act of lexical retrieval results in persistent learning that adjusts the weights of the connections between semantic and lexical representations. Importantly, in our analyses we have equated the influence of other factors, like frequency, between words of lower and higher intercorrelational density. Thus, in this scenario, across a lifetime, words of lower and higher intercorrelational density will be selected at a comparable rate. Consequently, strengthening of conceptual to lexical connections after successful selection would apply equally to words of lower and higher intercorrelational density. However, it seems possible that, everything else being equal, words with higher intercorrelational density would be co-activated more often via their strongly intercorrelated features, in contrast to words with lower intercorrelational density. When these items are co-activated, but not selected, the connections between their semantic and lexical representations would be increasingly and persistently weakened. When a word with such weakened connections becomes the target in the experiment, its lexical representation is harder to access, resulting in relatively poorer performance. In sum, although computational modelling may be needed to test this prediction, Oppenheim et al.’s theoretical framework also seems to be able to account for our finding of an inhibitory effect of intercorrelational density.

Finally, inhibitory effects of intercorrelational density on naming accuracy could also be explained in Dell's (1986) theory where the most active lexical representation is selected, irrespective of co-activated lexical representations. In the interactive architecture, the bi-directional information flow between the semantic and the lexical levels could be particularly strong amongst strongly intercorrelated semantic features, leading to stronger activation of non-target lexical representations, which can erroneously exceed the target word's activation and therefore be selected. This would increase the number of naming errors for words with higher intercorrelational density, as reported here.

Distinctiveness. Distinctiveness was defined as the mean uniqueness of the features of a concept in the feature norm database, with concepts with higher distinctiveness, by definition, having more unique and less shared features. The effect of distinctiveness was inhibitory[4], which is in contrast to previous studies that have investigated an effect of concept distinctiveness in picture naming: Rabovsky et al. (2016) and Taylor et al. (2012) found faster and more accurate responses for words with higher distinctiveness (see also Humphreys et al., 1988, for a facilitatory effect of a similar measure). As an explanation for facilitatory effects of distinctiveness, distinctive features have been suggested to play a privileged role in conceptual processing (e.g., Cree et al., 2006; Taylor et al., 2012), leading to faster and stronger activation of such features and to stronger activation of the respective concept in feature verification tasks[5].

If distinctive features are preferentially processed at the conceptual level, how can our finding of a negative effect of higher distinctiveness on naming be explained? When processing a word with higher distinctiveness, due to its rather unique semantic features, only a few semantically related lexical representations will be co-activated via the features they share with the target word. Alternatively, in a

[4] The effect was also inhibitory in an exploratory, not preregistered, analysis in which intercorrelational density was not included in the analysis, following Rabovsky et al. (2016) (please refer to the Introduction for details on the analysis conducted by Rabovsky et al.).

[5] Cree et al. (2006) explained the facilitatory effect of distinctiveness in the context of an attractor network: The strength and speed of activation of privileged distinctive features allows the network to enter the correct attractor basin faster, which enables it to settle in a stable state more rapidly.

holistic semantic architecture, there are (holistic) concepts which represent features (e.g., seeds), which are connected to a range of other concepts (e.g., apple, grapes, grass). Distinctiveness of a target (e.g., apple) could capture the average number of these connections for each of the (feature) concepts to which the target is connected (e.g., red, seeds). For words with fewer connections (i.e., words with higher distinctiveness), target processing would result in reduced activation spreading to other holistic lexical concepts. Thus, irrespective of the semantic architecture, enhanced co-activation and thus competition for lexical selection would be expected for words with *lower* distinctiveness, which predominantly have features that are shared with other concepts (see Vieth et al., 2014 Figure 1, for a visualisation) or may be connected to many other concepts, which are highly connected themselves. Thus, our inhibitory effect of higher distinctiveness is unlikely to be due to stronger competition between co-activated (lexical) representations, as fewer competing candidates are expected to be co-active when processing words with higher distinctiveness.

Importantly, this account of competition from strongly co-activated representations is also heavily based on the notion of the number of near semantic neighbours (i.e., words that share many semantic features with the target word) competing with the target for selection. Yet, in our analyses, and in contrast to the previous studies by Rabovsky et al. (2016) and Taylor et al. (2012), effects of the number of near semantic neighbours were accounted for and held constant by including this measure in the analysis. Thus, our statistical approach enabled us to obtain a more precise (purer) estimate of the effect of mean concept distinctiveness on naming, over and above number of near semantic neighbours, to evaluate the effect of a concept having distinctive features.

Critically, it has not yet been fully established exactly how numerous or closely related to the target co-activated lexical representations must be to compete for selection; however, there is research suggesting that more closely related representations cause stronger competition (e.g., Rose et al., 2019). Thus, while competition from many co-activated competitors is unlikely to be the mechanism underlying the inhibitory effect of distinctiveness, it may be possible that the effect could be caused by few highly related co-activated representations. There may, for example, be only one or two concepts that are closely related to the target, sharing many of its relatively unique features, and thus acting as strong competitors for selection, causing the observed inhibitory effect. In future studies, this could perhaps be tested by operationalising the semantic similarity between the target and its closest semantic neighbour(s) as a further semantic variable.

Similarly, in Oppenheim et al.'s (2010) framework we would expect an *advantage* for words that have higher distinctiveness, through weakening of connections between *shared* features and non-target lexical representations. Every time a lexical item is retrieved, when there are shared features, links from those features to non-target representations would be subject to increased weakening. A word with lower distinctiveness (which has more shared features) would be subject to this weakening to a higher degree than a word with higher distinctiveness, resulting in poorer performance on items of *lower* distinctiveness and not of *higher* distinctiveness as we found in our data.

In contrast to the accounts of a privileged status of distinctive features described above, the Conceptual Structure Account (e.g., Tyler et al., 2000; Tyler & Moss, 2001) of semantic memory argues that there is a processing *disadvantage* for distinctive features of living things and an advantage for shared features, while there is no such dramatic difference for artifacts. The disadvantage for distinctive features of living things is thought to arise because they are only weakly correlated with other features and hence do not benefit from mutual activation among correlated features. This causes slowed access to such features and makes them more vulnerable to loss in impaired semantic processing. In contrast, distinctive features of artifacts are correlated more strongly with other features and are therefore thought to benefit from mutual activation among the correlated features. Hence, this account predicts an *inhibitory* effect of distinctiveness for living things, which was found in experimental work using feature verification and definition tasks with individuals with semantic impairment (Moss et al., 1998, 2002) and unimpaired participants (Randall et al., 2004). Randall et al. found the predicted advantage in feature verification for shared over distinctive features for living things (while there was no robust main effect of distinctiveness across domains on feature verification speed or accuracy). Moreover, distinctive features of living things were verified more slowly than distinctive features of non-living things and higher error rates were reported for distinctive rather than for shared features when this data was computationally simulated. However, Cree et al. (2006) pointed out several shortcomings in Randall et al.'s work (e.g., insufficient matching), and reported findings that were inconsistent with Randall et al.: Effects of distinctiveness that were comparable for living and nonliving things.

While the account by Randall et al. (2004) predicts an inhibitory effect for distinctiveness for some items (i.e., living things), it also relies heavily on intercorrelations between features to explain the effects. To determine whether the inhibitory effect in our data was particularly strong for living things, as predicted by the Conceptual Structure Account, we ran exploratory post-hoc analyses[6] of the naming latency and accuracy data that included an interaction between distinctiveness and animacy (reported in Appendix F). While the interaction was non-significant for naming accuracy, it was significant for the naming latency analysis. However, the findings were in the opposite direction to that predicted by the Conceptual Structure Account: Responses to living things were unaffected by distinctiveness, but there was a significant inhibitory effect for non-living things. Hence, while the Conceptual Structure Account is consistent with the general direction of our inhibitory finding for distinctiveness, its more specific predictions do not hold for our data. Importantly, the Conceptual Structure Account and most previous investigations of distinctiveness were designed in the context of category specific impairments and the processing of single features, which makes it hard to generalise the frameworks and findings to our work. Moreover, in contrast to the current study, many previous studies did not test for an effect of overall concept distinctiveness and did not investigate picture naming.

In our study, we held both intercorrelational density and number of near semantic neighbours constant and still found a significant inhibitory effect of distinctiveness. As the origin of our effect and its theoretical interpretation are unclear, we suggest that future computational modelling could be helpful for uncovering the mechanisms underlying this effect and lead to better understanding of it.

[6] As these were unplanned post-hoc analyses to explore a particular finding in more detail, these analyses were not preregistered.

Contributions and limitations of this research

In this study, we determined feature-based semantic variables that are important for word production. Our findings of effects of semantic variables allowed us to critique theories of spoken word production and to evaluate how they must operate to be able to explain the results. Interestingly, we identified semantic variables that had both facilitatory and inhibitory influences on picture naming. Thus, theories of word production need mechanisms that can account for both effects. While, as detailed in the previous sections, semantic facilitation is implemented in current theories as spreading activation or feedback from lexical to semantic representations, interference can be explained via lexical competition or could be due to long-term adjustments of conceptual to lexical connections.

Importantly, current theories of word production are underspecified regarding the way semantic variables may affect performance. Of the currently available theories of word production, Abdel Rahman and Melinger (2009, 2019) explicitly address word-inherent variables and their Swinging Lexical Network Hypothesis comprises the architectural elements (i.e., spreading activation and lexical competition) to parsimoniously explain both facilitatory and inhibitory effects of such variables. Other theories (e.g., Dell, 1986; Oppenheim et al., 2010) were able to only account for some of our results. Crucially, for current theories to explain the findings, we were required to make additional assumptions regarding processing or the encoding of semantic relations that are not explicitly specified in the theories. Thus, the current theories of word production need further modification and specification to explain effects of semantic variables.

Important methodological contributions of this paper have been highlighted throughout this section: the importance of rigorous control of other variables influencing word production (i.e., both control and other semantic variables), highly powered statistical analyses, and considering variables of interest simultaneously.

This study was also subject to some limitations. While we showed that multicollinearity between the measures included in our analyses did not occur at compromising levels, there were some relatively high correlations between semantic variables (Table 3). However, using other statistical techniques (e.g., Principal Component Analysis) was neither suitable for addressing our research goals, nor possible given a single measure of each variable. In addition, according with our study's aim to identify individual semantic variables that affect picture naming performance, in the sections above, we assessed how the effect of each of the significant semantic variables could be explained in the context of current theories of word production. While we covered a range of semantic variables that capture different aspects of the flow of activation at the semantic level and co-activation at the lexical level, it may also be the case that there are other (potentially composite) measures that outperform the variables used here in capturing the dynamics of activation spread and lexical co-activation. Further research and modelling could address this.

Use of semantic feature norms

We chose which semantic variables to study based on whether it was possible to quantify these variables objectively rather than using ratings. We chose variables that could be operationalised based on the information provided in semantic feature databases (and specifically that of McRae et al., 2005). However, in theories of word production there is no theoretical agreement that conceptual representations are actually decomposed into semantic features: While some accounts assume decomposed conceptual representations (e.g., Dell, 1986; Oppenheim et al., 2010), others presume non-decomposed, holistic, representations of meaning (e.g., Levelt et al., 1999; Roelofs, 1992).

While we use the context of semantic features to characterise aspects of the semantic representation of concepts and relationships between concepts, and to interpret our findings, this does not imply that we favour theories of decomposed conceptual representations over non-decomposed theories. The semantic variables we studied here could also reflect semantic relationships as described by non-decomposed theories of semantics such as direct connections between concepts (Collins & Loftus, 1975) or indirect connections via property nodes (Collins & Quillian, 1969). Throughout the theoretical explanations of our significant findings, we pointed towards possible ways the significant variables may operate in a holistic semantic architecture. Importantly though, computational modelling may be helpful to further assess our hypotheses concerning the mechanisms underlying the effects of semantic variables in the context of different semantic architectures and to aid the interpretation of the effects of these variables.

Hence, while the effects of this study can more readily be explained in the context of models assuming feature-based semantic representations, our findings do not allow us to adjudicate between decomposed and non-decomposed semantic representations.

Participant generated semantic features are a verbalisation and temporary abstraction of complex semantic representations, which have developed through multisensory experience with the concept in real life (e.g., McRae, 2004). While semantic feature databases provide one of the best currently available approaches to specifying conceptual content, researchers have also pointed out some limitations (e.g., Amsel & Cree, 2013; Randall et al., 2004). While operationalising our semantic variables based on a semantic feature database (McRae et al., 2005), we accept the shortcomings of this resource, such as, for example, the underrepresentation of highly frequent features (e.g., *breathes*) or the absence of features that are difficult to verbalise (Cree & McRae, 2003).

However, in contrast to ratings, calculating semantic variables based on semantic features represents a more objective way of operationalising conceptual knowledge, as participants are not asked to directly rate a certain aspect of the semantic representation (e.g., concept typicality within its semantic category). Nevertheless, while measures based on semantic feature databases are transparent and reproducible, their calculation is based on the rather limited number of items in the corpus.

As described in the Introduction, semantic feature databases present only one way of describing conceptual knowledge. Semantic relationships and similarity could also be objectively captured by other semantic dimensions, such as associations (e.g., De Deyne et al., 2019; Nelson et al., 2004) or contextual co-occurences (e.g., Latent Semantic Analysis (LSA), Landauer et al., 1998; Continuous bag-of-words model (CBOW), Mikolov et al., 2013; Global Vectors for word representation (GloVe), Pennington et al., 2014). In this study, we conducted a thorough investigation of the feature-based dimension of semantic relationships as previous research suggested a large variety of feature-based variables that affected word production. In addition, when comparing different measures of semantic neighbourhood density, Hameau et al. (2019) found a feature-based measure to be a better predictor of picture naming in people with aphasia than association-based or contextual semantic neighbourhood density. Future research could extend our approach by including semantic variables derived from other semantic dimensions to determine the extent to which contextual and association-based similarity measures can also affect word production. Such an approach would also allow assessment of how far the observed dynamics of semantic and lexical activation in word production depend on the structural assumptions of semantic feature databases, and, finally, to identify the dimension of semantics that captures most of the variability in our data.

Conclusion

We have reported an investigation of effects of six feature-based semantic variables (number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness) on the picture naming performance of a large group of young English speakers. We found that number of semantic features was the most important predictor of naming accuracy and, to a lesser extent, of naming latency, with further effects of intercorrelational density and distinctiveness. These semantic variables should therefore be considered in further research.

This research allowed us to determine which variables influence behaviour and develop hypotheses regarding the mechanisms and dynamics underpinning these effects in theories of spoken word production. The findings of our study indicate that both aspects of the richness of the semantic representation and of relationships between concepts play an important role in semantic and lexical processing during word production. Importantly though, some previously reported effects of several semantic variables (namely number of near semantic neighbours and semantic similarity (in speeded naming), typicality, and the direction of the effect of distinctiveness) did not replicate here, which highlights once again the importance of a well-controlled study design, to minimise the risk of miscalibration in estimating such effects.

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*(5), 713–734. https://doi.org/10.1080/01690960802597250

Abdel Rahman, R., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience*, *34*(9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Allison, P. (2012). *When can you safely ignore multicollinearity?* Statistical Horizons. https://statisticalhorizons.com/multicollinearity

Amsel, B. D., & Cree, G. S. (2013). Semantic richness, concreteness, and object domain: An electrophysiological study. *Canadian Journal of Experimental Psychology*, *67*(2), 117–129. https://doi.org/10.1037/a0029807

Baayen, H. R., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28. https://doi.org/10.21500/20112084.807

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv*, 1–27. https://doi.org/arXiv:1506.04967

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., & Green, P. (2017). *lme4: Linear Mixed-Effects Models using "Eigen" and S4* (R package version 1.1-21).

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press. https://doi.org/10.3115/981923.981940

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 6.0.49). https://doi.org/http://www.praat.org/

Bonin, P., Méot, A., Laroche, B., Bugaiska, A., & Perret, C. (2019). The impact of image characteristics on written naming in adults. *Reading and Writing*, *32*(1), 13–31. https://doi.org/10.1007/s11145-017-9727-3

Bormann, T. (2011). The role of lexical-semantic neighborhood in object naming: Implications for models of lexical access. *Frontiers in Psychology*, *2*, 1–11. https://doi.org/10.3389/fpsyg.2011.00127

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240–247. https://doi.org/https://doi.org/10.1016/S0022-5371(69)80069-1

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384–387. https://doi.org/10.3758/BF03195585

Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a priviledged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychological Learning, Memory and Cognition*, *32*(4), 643–658. https://doi.org/10.1037/0278-7393.32.4.643

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*(2), 163–201. https://doi.org/10.1037/0096-3445.132.2.163

Davis, C. J. (2005). N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words"

English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, *32*(4), 588–615. https://doi.org/10.3758/BF03200832

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Devereux, B. J., Tyler, L. K., Geertzen, J., & Randall, B. (2014). The Centre for Speech, Language and the Brain (CSLB) concept property norms. *Behavior Research Methods*, *46*(4), 1119–1127. https://doi.org/10.3758/s13428-013-0420-4

Ewald, A., Aristei, S., Nolte, G., & Rahman, R. A. (2012). Brain oscillations and functional connectivity during overt language production. *Frontiers in Psychology*, *3*, 1–12. https://doi.org/10.3389/fpsyg.2012.00166

Fieder, N., Wartenburger, I., & Abdel Rahman, R. (2019). A close call: Interference from semantic neighbourhood density and similarity in language production. *Memory & Cognition*, *47*(1), 145–168. https://doi.org/10.3758/s13421-018-0856-y

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International Journal of Psychophysiology*, *29*(1), 65–75. https://doi.org/10.1016/S0167-8760(97)00099-8

Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language*, *93*(1), 79–94. https://doi.org/10.1016/j.bandl.2004.08.003

Gauvin, H. S., Jonen, M. K., Choi, J., McMahon, K., & de Zubicaray, G. I. (2018). No lexical competition without priming: Evidence from the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology*, *71*(12), 2562–2570. https://doi.org/10.1177/1747021817747266

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427. https://doi.org/10.3758/BF03201693

Grossman, M., Robinson, K., Biassou, N., White-Devine, T., & D'Esposito, M. (1998). Semantic memory in Alzheimer's disease: Representativeness, ontologic category, and material. *Neuropsychology*, *12*(1), 34–42. https://doi.org/10.1037/0894-4105.12.1.34

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson.

Hameau, S., Nickels, L., & Biedermann, B. (2019). Effects of semantic neighbourhood density on spoken word production. *Quarterly Journal of Experimental Psychology*, *72*(12), 2752–2775. https://doi.org/10.1177/1747021819859850

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition*, *13*(7–8), 884–910. https://doi.org/10.1080/13506280544000093

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*(1), 67–104. https://doi.org/10.1080/02643298808252927

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage Publications.

Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and naming times for a set of 539 pictures: The role of age of acquisition. *Behavior Research Methods*, *42*(2), 461–469. https://doi.org/10.3758/BRM.42.2.461

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*(2), 243–275. https://doi.org/10.1016/0010-0285(84)90009-4

Kremin, H. (1986). Spared naming without comprehension. *Journal of Neurolinguistics*, *2*(1–2), 131–150. https://doi.org/10.1016/S0911-6044(86)80008-2

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lampe, L. F., Bürki, A., Sowman, P. F., Hameau, S., & Nickels, L. (2021). Effects of semantic variables on processes during word planning for production: Evidence from electrophysiological data. *Manuscript in Preparation*.

Lampe, L. F., Fieder, N., Krajenbrink, T., & Nickels, L. (2017). Semantische Nachbarschaft in der Wortproduktion bei Aphasie [Semantic neighbourhood in word production in aphasia]. In A. Adelt, Ö. Yetim, C. Otto, & T. Fritzsche (Eds.), *Spektrum Patholinguistik - Panorama Patholinguistik: Sprachwissenschaft trifft Sprachtherapie* (Vol. 10, pp. 103–114). Universitätsverlag Potsdam. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/39701

Lampe, L. F., Hameau, S., Bürki, A., Fieder, N., Sowman, P. F., & Nickels, L. (2019). *Investigating effects of semantic variables on word production*. Preregistration on the Open Science Framework. osf.io/yw6ma

Lampe, L. F., Hameau, S., & Nickels, L. (2021). Are they really stronger? Comparing effects of semantic variables in speeded deadline and standard picture naming. *Manuscript Submitted for Publication*.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2–3), 259–284. https://doi.org/10.1080/01638539809545028

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.

Lüdecke, D. (2019). *sjPlot: Data visualization for statistics in social science* (R package version 2.6.3). https://cran.r-project.org/package=sjPlot

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*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375

McRae, K. (2004). Semantic memory: Some insights from feature-based connectionist attractor networks. In B. H. Ross (Ed.), *The psychology of learning and motivation - Advances in research and theory* (pp. 41–86). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(03)45002-0

McRae, K., & Cree, G. S. (2002). Factors underlying category-specific impairments. In E. M. E. Forde & G. W. Humphreys (Eds.), *Category specificity in brain and mind* (pp. 211–249). Psychology Press.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

McRae, K., Cree, G. S., Westmacott, R., & Sa, V. R. De. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology*, *53*(4), 360–373. https://doi.org/10.1037/h0087323

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99–130. https://doi.org/10.1080/10430719008404646

Melinger, A., & Abdel Rahman, R. (2013). Lexical selection is competitive: Evidence from indirectly activated semantic associates during picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 348–364. https://doi.org/10.1037/a0028941

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations (ICLR)*, 1–12. https://arxiv.org/abs/1301.3781

Millard, S. P. (2013). *EnvStats: An R package for environmental statistics*. Springer.

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, *25*(10), 3343–3355. https://doi.org/10.1093/cercor/bhu137

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive,* *Affective and Behavioral Neuroscience*, *11*(1), 32–43. https://doi.org/10.3758/s13415-010-0009-7

Mirman, D., & Graziano, K. M. (2013). The neural basis of inhibitory effects of semantic and phonological neighbors in spoken word production. *Journal of Cognitive Neuroscience*, *25*(9), 1504–1516. https://doi.org/10.1162/jocn_a_00408

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 65–79. https://doi.org/10.1037/0278-7393.34.1.65

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705–714. https://doi.org/10.3758/BF03202720

Moss, H. E., Tyler, L. K., & Devlin, J. T. (2002). The emergence of category specific deficits in a distributed semantic system. In E. Forde & G. W. Humphreys (Eds.), *Category-specificity in brain and mind* (pp. 115–145). Psychology Press.

Moss, H. E., Tyler, L. K., Durrant-Peatfield, M., & Bunn, E. M. (1998). 'Two eyes of a see-through': Impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase*, *4*(4–5), 291–310. https://doi.org/10.1080/13554799808410629

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

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*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1, Pt.2), 1–25. https://doi.org/10.1037/h0025327

Pandey, S., & Elliott, W. (2010). Suppressor variables in social work research: Ways to identify in multiple regression models. *Journal of the Society for Social Work and Research*, *1*(1), 28–40. https://doi.org/10.5243/jsswr.2010.2

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. https://doi.org/10.3115/v1/D14-1162

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). The neural consequences of semantic richness: When more comes to mind, less activation is observed. *Psychological Science*, *18*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2007.01913.x

Pexman, P. M., Holyk, G. G., & Monfils, M.-H. (2003). Number-of-features effects and semantic processing. *Memory & Cognition*, *31*(6), 842–855. https://doi.org/10.3758/BF03196439

Pexman, P. M., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience*, *7*, 1–3. https://doi.org/10.3389/fnhum.2013.00723

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*(5), 377–500. https://doi.org/10.1080/02643299308253469

Python, G., Fargier, R., & Laganaro, M. (2018). ERP evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex*, *99*, 1–12. https://doi.org/10.1016/j.cortex.2017.09.008

Qualtrics. (2018). *Qualtrics* (Version 04/18). https://doi.org/https://www.qualtrics.com

R Core Team. (2018). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org/

Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, *132*(1), 68–89. https://doi.org/10.1016/j.cognition.2014.03.010

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, *146*, 240–244. https://doi.org/10.1016/j.cognition.2015.09.016

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2021). Semantic richness and density effects on language production: Electrophysiological and behavioral evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47(3), 508–517. https://doi.org/10.1037/xlm0000940*

Randall, B., Moss, H. E., Rodd, J. M., Greer, M., & Tyler, L. K. (2004). Distinctiveness and correlation in conceptual structure: Behavioral and computational studies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 393–406. https://doi.org/10.1037/0278-7393.30.2.393

Rico Duarte, L., & Robert, C. (2014). Semantic richness influences naming pictures in old age and in Alzheimer’s disease. *Journal of Neurolinguistics*, *31*, 55–68. https://doi.org/10.1016/j.jneuroling.2014.06.004

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (1997). A case for nondecomposition in conceptually driven word retrieval. *Journal of Psycholinguistic Research*, *26*(1), 33–67. https://doi.org/10.1023/A:1025060104569

Rogers, T. T., Patterson, K., Jefferies, E., & Lambon Ralph, M. A. (2015). Disorders of representation and control in semantic cognition: Effects of familiarity, typicality, and specificity. *Neuropsychologia*, *76*, 220–239. https://doi.org/10.1016/j.neuropsychologia.2015.04.015

Rogerson, P. (2011). *Statistical methods for geography*. SAGE Publications, Ltd. https://doi.org/10.4135/9781849209953

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rose, S. B., & Abdel Rahman, R. (2016). Cumulative semantic interference for associative relations in language production. *Cognition*, *152*, 20–31. https://doi.org/10.1016/j.cognition.2016.03.013

Rose, S. B., & Abdel Rahman, R. (2017). Semantic similarity promotes interference in the continuous naming paradigm: Behavioural and electrophysiological evidence. *Language, Cognition and Neuroscience*, *32*(1), 55–68. https://doi.org/10.1080/23273798.2016.1212081

Rose, S. B., Aristei, S., Melinger, A., & Abdel Rahman, R. (2019). The closer they are, the more they interfere: Semantic similarity of word distractors increases competition in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 753–763. https://doi.org/10.1037/xlm0000592

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236. https://doi.org/10.1068/p5117

RStudio Team. (2019). *RStudio: Integrated development for R* (1.2.5033). RStudio, Inc. http://www.rstudio.com

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Taylor, K. I., Devereux, B. J., Acres, K., Randall, B., & Tyler, L. K. (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. *Cognition*, *122*(3), 363–374. https://doi.org/10.1016/j.cognition.2011.11.001

Taylor, K. I., Moss, H., Randall, B., & Tyler, L. K. (2004). The interplay between distinctiveness and intercorrelation in the automatic activation of word meaning. *Proceedings of the 45th Annual Meeting of the Psychonomic Society*, *9*, 109.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Lawrence Erlbaum Associates.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(6), 244–252. https://doi.org/10.1016/S1364-6613(00)01651-X

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain and Language*, *75*(2), 195–231. https://doi.org/10.1006/brln.2000.2353 van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vieth, H. E., McMahon, K. L., & de Zubicaray, G. I. (2014). Feature overlap slows lexical selection: Evidence from the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology*, *67*(12), 2325–2339. https://doi.org/10.1080/17470218.2014.923922

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183–190. https://doi.org/10.3758/BRM.40.1.183

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 664–680. https://doi.org/10.1037/0278-7393.17.4.664

Woollams, A. M. (2012). Apples are not the only fruit: The effects of concept typicality on semantic representation in the anterior temporal lobe. *Frontiers in Human Neuroscience*, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00085

Appendices

Appendix A: Composition of the six experimental lists

Table A1

Composition of the different experimental lists

	Position 1	Position 2	Position 3	Position 4
List 1	Block 1 – Pseudorand. 1	Block 2 – Pseudorand. 1	Block 3 – Pseudorand. 1	Block 4 – Pseudorand. 1
List 2	Block 1 – Pseudorand. 1	Block 2 – Pseudorand. 1	Block 4 – Pseudorand. 1	Block 3 – Pseudorand. 1
List 3	Block 1 – Pseudorand. 2	Block 3 – Pseudorand. 2	Block 4 – Pseudorand. 2	Block 2 – Pseudorand. 2
List 4	Block 1 – Pseudorand. 2	Block 3 – Pseudorand. 2	Block 2 – Pseudorand. 2	Block 4 – Pseudorand. 2
List 5	Block 1 – Pseudorand. 3	Block 4 – Pseudorand. 3	Block 2 – Pseudorand. 3	Block 3 – Pseudorand. 3
List 6	Block 1 – Pseudorand. 3	Block 4 – Pseudorand. 3	Block 3 – Pseudorand. 3	Block 2 – Pseudorand. 3

Note. Pseudorand. = Pseudorandomisation

Appendix B: Detailed description of calculation of semantic variables based on information given in McRae et al. (2005)

Number of semantic features

Number of semantic features was the total number of semantic features produced by at least 5 out of 30 participants when generating features for the concept. Taxonomic features were excluded from this calculation and the calculation of all other semantic variables. This measure was directly retrieved from McRae et al. (2005).

Intercorrelational density

Intercorrelational density was calculated based on Pearson's product moment correlations between the pairs of features in McRae et al.'s (2005) concept-feature matrix (cells were filled with the production frequency of the particular feature for a concept). Features that appeared in less than three concepts were excluded to avoid spurious correlations.

Feature pairs were counted as significantly correlated if the two features shared at least 6.5% of their variance (r^2). From this information, intercorrelational density of a concept was calculated as the sum of the shared variance of its significantly correlated feature pairs. This measure was directly retrieved from the database by McRae et al. (2005).

Number of near semantic neighbours

A near semantic neighbour was defined as a word that had feature overlap of at least .4 (cosine similarity between feature production frequency vectors in the database by McRae et al. (2005)) with the target word (following Mirman, 2011). The instances of near semantic neighbours were added for each concept to determine the number of near semantic neighbours.

Semantic similarity

Semantic similarity was the average similarity of the target word's feature production frequency vector and the feature vectors of all other concepts from the McRae et al. (2005) database (following Mirman & Magnuson, 2008).

Typicality

Typicality was calculated in a similar way to Rosch and Mervis' (1975) family resemblance score. First, the features of an item were weighted based on category affiliations. Each feature of an item was attributed with the number of other items in the same semantic category that were credited with that particular feature.

This value was then divided by the number of items in the semantic category, which results in a proportion. Such a division was not part of the original measure by Rosch and Mervis (1975), who used raw values to determine the family resemblance score; however, we decided for this step because the number of items in each semantic category was not uniform in our data, as opposed to Rosch and Mervis.

Subsequently, we amended the original family resemblance score calculations further by multiplying this value for each feature by its production frequency (number of participants who produced that feature for the target word when generating the feature norms) to account for feature importance (e.g., Garrard et al., 2005). As a final step, the feature weights of all features of an item were summed.

We considered four different ways to operationalise typicality: 1) taking production frequency of individual features into account or not, and 2) using either raw values or proportions based on the number of items in a particular semantic category. The four versions of the typicality measure were highly correlated with one another, however, the measure as described above (using proportions based on the number of items in a particular semantic category and taking production frequencies of individual features into account) correlated least strongly with the other semantic variables.

Distinctiveness

Distinctiveness of each feature of an item was calculated as the inverse of the number of concepts in which that feature occurred across the database. To determine conceptual distinctiveness, this feature distinctiveness was then averaged across the features of a concept. This measure was also directly retrieved from McRae et al. (2005).

Appendix C: Comparison of semantic variables in the full McRae et al. (2005) database and the 297 items included in the experimental investigation in Study 2

The feature database by McRae et al. (2005) contains 541 concepts, however, here we only used a subset of these items (*n* = 297) with high name agreement in Australian English, as identified in the norming study presented in Study 1. Importantly, and as shown in Table C1, the items selected for the experimental investigation in Study 2 were a good representation of the items of the full McRae et al. set as semantic variables were largely comparable for the selected items and the whole database. This suggests that the selection of high name agreement items in Australian English and the resulting reduction of the stimuli compared to the full McRae et al. database did most likely *not* affect the comparability of the findings of Study 2 to previous work that used the full McRae et al. database (i.e., Rabovsky et al., 2016).

Table C1

Descriptive statistics for the semantic variables included in Study 2 and the full McRae et al. (2005) database

	Study 2: *n* = 297 items			McRae et al. (2005) full database: *n* = 541 items		
Semantic variable	Mean	*SD*	Range	Mean	*SD*	Range
Number of semantic features	12.71	2.99	6–20	12.07	3.21	5–21
Intercorrelational density	153.40	172.19	0.00–1296.22	175.06	205.66	0.00–1419.41
Semantic similarity	0.04	0.02	0.00–0.09	0.04	0.02	0.00–0.09
Number of near semantic neighbours	6.14	7.55	0–38	7.14	8.64	0–40
Typicality	32.64	16.14	4.22–91.25	32.47	16.85	4.22–95.44
Distinctiveness	0.37	0.16	0.04–0.80	0.35	0.17	0.03–0.80

Appendix D: Detailed response coding

Table D1

Response coding examples

Target word	Response	Response type	RT	Accuracy
Sofa	Sofa	Correct	RT	1
Sofa	A sofa (determiner)	Correct with determiner	-1	1
Sofa	S sofa	Disfluency on the initial phoneme of target word	-1	1
Sofa	Erm sofa	Hesitation	-1	1
Sofa	Red sofa	Elaboration	-1	1
Sofa	Couch	Synonym or acceptable alternative	-1	NA
Submarine	Sub	Abbreviation	-1	NA
Bed	Emu (previous item: emu) .. Bed	Previous	-1	NA
Squid	O squid or Oct squid	Disfluency with self-correction	-1	0
Sofa	So	Incomplete response	-1	0
Sofa	Chair	Error	-1	0
Sofa	I don't know	Omission	-1	0

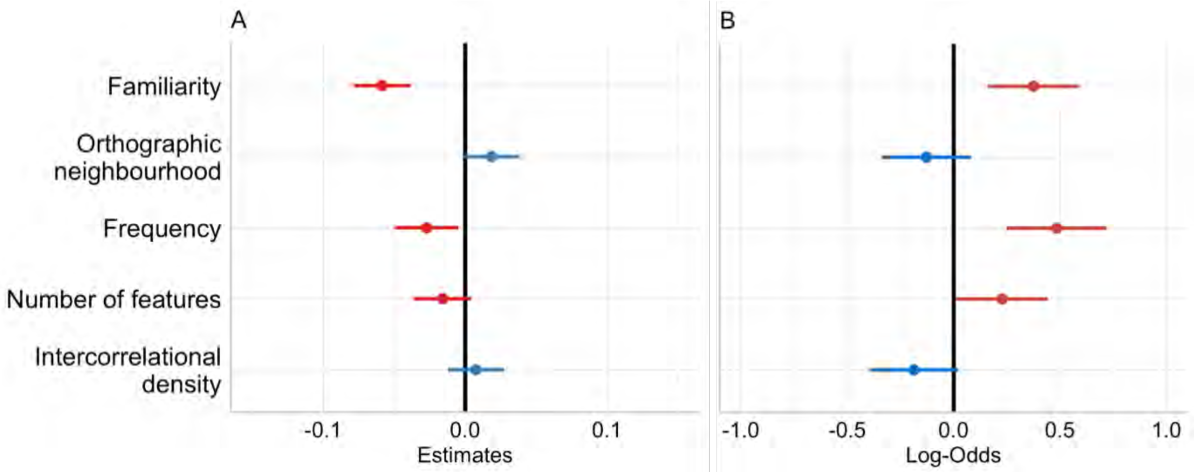
Note. RT = Response time, NA = not analysed; responses with RT -1 were excluded from the naming latency analyses.

Appendix E: Visualisations of findings of Analyses 1A and 1B

The findings of Analysis 1A are graphically displayed in Figure E1, with Model 1A.1 in Panel A and Model 1A.2 in Panel B. Figure E2 displays the findings of Analysis 1B, with Model 1B.1 in Panel A and Model 1B.2 in Panel B. Red lines indicate a facilitatory effect of a variable with speeding of naming latency or higher naming accuracy as that variable increases in value, while blue lines indicate an inhibitory effect with slowing of naming latency or reduced naming accuracy as that variable increases in value. Confidence intervals that cross the black zero line are indicative of non-significant effects.

Figure E1

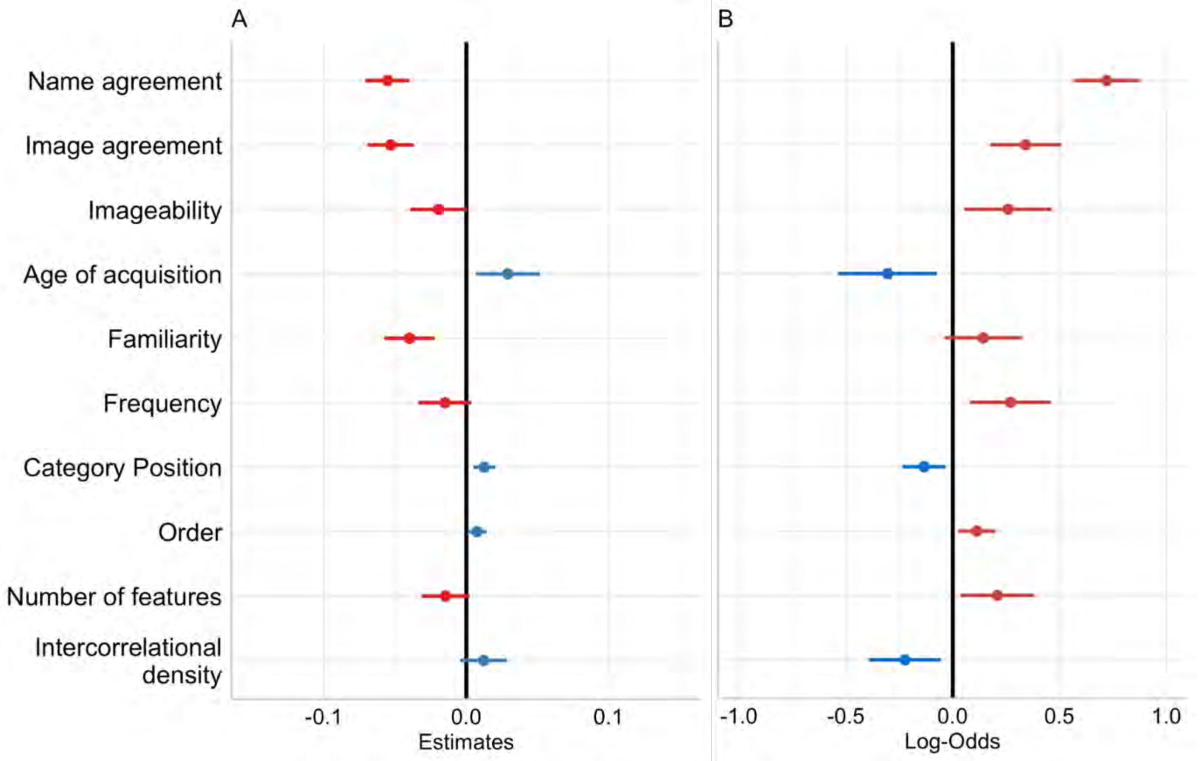
Analysis 1A replicating Rabovsky et al. (2016): Fixed effects estimates with 95% confidence interval of picture naming latency analysis (Panel A; Model 1A.1) and accuracy analysis (Panel B; Model 1A.2)



Note. Panel A shows the output of the naming latency analysis and Panel B the output of the naming accuracy analysis; red lines (to the left of centre for latency, and right for accuracy) indicate increased values of the variable lead to better performance, and blue lines (to the right of centre for latency, and left for accuracy) indicate worse performance.

Figure E2

Analysis 1B replicating Rabovsky et al. (2016) taking more psycholinguistic control variables into account: Fixed effects estimates with 95% confidence interval of picture naming latency analysis (Panel A; Model 1B.1) and accuracy analysis (Panel B; Model 1B.2)



Note. Panel A shows the output of the naming latency analysis and Panel B the output of the naming accuracy analysis; red lines (to the left of centre for latency, and right for accuracy) indicate increased values of the variable lead to better performance, and blue lines (to the right of centre for latency, and left for accuracy) indicate worse performance.

Appendix F: Post-hoc analyses of distinctiveness in living things: Testing the prediction of the Conceptual Structure Account

The Conceptual Structure Account predicts a disadvantage for distinctive features of living things. We therefore tested the explanatory strength of this account for our inhibitory effect of distinctiveness on naming latency and for naming accuracy.

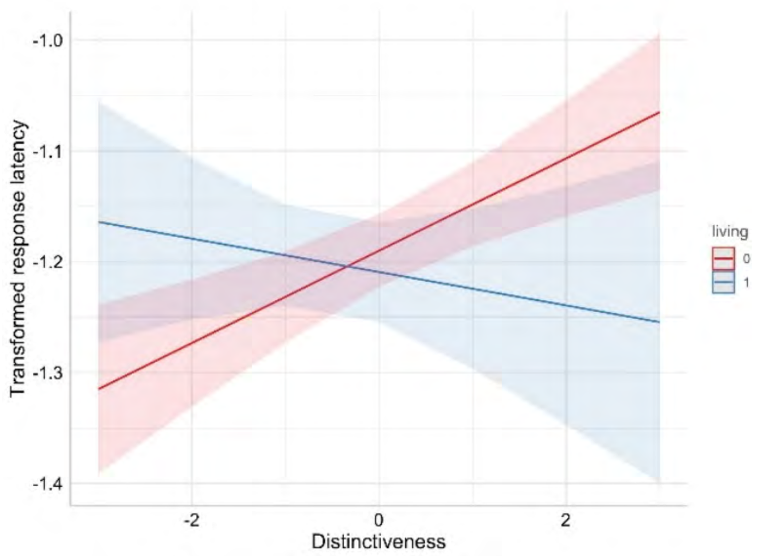
We coded items of the semantic categories *birds, fruit, invertebrate, land animal, plant, sea creature,* and *vegetable* as living (n = 100) and all other items were coded as non-living.

The structure of the linear mixed effect model used for this exploratory analysis included the same fixed effects as the model of Analysis 2.1 with an additional interaction between distinctiveness and animacy. A random by-participant slope for number of near semantic neighbours was included.

The interaction between distinctiveness and animacy was significant (Estimate = -0.06, SE = 0.02, t = -2.69, p = .008) and is displayed in Figure E1. While words of the *living* category were unaffected by distinctiveness, the inhibitory main effect of distinctiveness in Analysis 2 seemed to be driven by the items of the *non-living* category. This finding allows us to exclude the Conceptual Structure Account as an appropriate framework for the finding of this experiment.

Figure F1

Interaction between distinctiveness and animacy



For naming accuracy, the generalised linear mixed effect model included the same fixed effects as Analysis 2.2 with the additional interaction between distinctiveness and animacy and a random by-participant slope for number of near semantic neighbours. The interaction between distinctiveness and animacy was non-significant (Estimate = 0.10, *SE* = 0.22, z = 0.45, p = .642).

Acknowledgements

We thank Nora Fieder and Audrey Bürki for helpful discussions when planning the experiment and Audrey Bürki for her comments on a previous version of this paper. We also thank Serje Robidoux for statistical advice and Andrew Etchell for his help when programming the experiment. Moreover, we would like to thank the Editor, Prof Keith A. Hutchison, and two anonymous reviewers for helpful comments on an earlier version of this manuscript.

CHAPTER 4

Effects of semantic variables on processes during word planning for production: Evidence from electrophysiological data

Leonie F. Lampe
Audrey Bürki
Paul F. Sowman
Solène Hameau
Lyndsey Nickels

Abstract

Semantic variables capture aspects of the activation environment of semantic and lexical processing in word production. While there is growing evidence regarding their effects on behavioural measures of word production, little previous research has examined their influence on electrophysiological data. Here we investigated electrophysiological correlates of six feature-based semantic variables, which were previously found to influence behavioural measures of picture naming: number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness. Data from 78 participants naming 291 colour photographs were analysed, while controlling for psycholinguistic variables affecting word production. We analysed the mean amplitude of event-related potential data in a posterior region of interest and conducted a microstate analysis on a trial-by-trial basis. Results revealed a stronger posterior positivity for words with more semantic features and number of semantic features, intercorrelational density, semantic similarity, and number of near semantic neighbours affected the number of timeframes associated with two microstates. The findings are interpreted as reflecting increased activity in the semantic and lexical network involved in word production, which was due to enhanced activation of the target word itself or activation distributed across a cohort of co-activated lexical representations caused by the significant semantic variables.

Introduction

For neurotypical native speakers of a language, the production of single words is a quick and effortless process. However, our minds and brains have to successfully complete a number of steps before we can utter a word. Different aspects of this process have been the focus of considerable research over the past decades. Using primarily behavioural data, such as naming latencies, accuracy, or error types, this research led to the formulation of several theoretical models of word production from pictures (e.g., Abdel Rahman & Melinger, 2009; Dell, 1986; Howard et al., 2006; Levelt et al., 1999; Oppenheim et al., 2010; Roelofs, 1992). The different models generally agree on the broader cognitive processes involved in word production: Visual recognition of the to-be-named object enables the activation of the target's concept in semantic memory, which in turn leads to the retrieval and selection of a lexical representation (lemma) and to the phonological and phonetic encoding of the word. However, there remain debates focusing on details of the encoding process and in particular on the processing of semantic information during target word planning (e.g., Abdel Rahman & Melinger, 2009; Mahon et al., 2007; Oppenheim et al., 2010). In this research we were particularly concerned with the cognitive mechanisms underlying encoding of semantic information in word production.

Semantic variables in word production

One way of exploring processing during word production is to study effects of word-specific characteristics (psycholinguistic variables, e.g., word frequency, word length) on the word production process. Naturally or experimentally occurring differences in such word characteristics and their influence on tasks that involve spoken word production can inform our understanding of both information representation and processing during word production. A sub-group of these word-specific characteristics are *semantic variables*, which describe characteristics of the semantic representation of the target concept and its relationship to other concepts (e.g., number of semantic features, intercorrelational density, number of near semantic neighbours). Recently, interest in effects of semantic variables on word production has increased (e.g., Fieder et al., 2019; Lampe et al., in press; Rabovsky et al., 2016) and semantic variables have been used to investigate details of lexical selection (e.g., whether lexical selection is a competitive process) and word production more generally. Research conducted on item-inherent semantic variables allows researchers to use a simple and relatively natural task: standard picture naming. Studying semantic variables can hence provide a valuable alternative to frequently used complex experimental paradigms, which manipulate target context (e.g., Picture Word Interference, Blocked Cyclic Naming) when studying semantic effects on word production.

Effects of semantic variables originate at the semantic level as they are hypothesised to relate to aspects of the semantic representation of the target word. In word production, activation at the semantic level (also referred to as the conceptual level) has direct consequences for lexical processing (i.e., lemma activation and selection), where semantic variables can influence the activation strength of the target word's lexical representation and/or the number, or strength of activation, of co-activated semantically related lexical representations (e.g., Abdel Rahman & Melinger, 2019; Rabovsky et al., 2016). Hence, knowledge about effects of semantic variables can inform our understanding of the activation environment in which word production takes place and thus advance models of word production. However, current models of word production are underspecified with regard to predicting and explaining effects of semantic variables. Consequently, for semantic variables to be utilised in targeted research that allows adjudication between different theoretical accounts, it is first necessary to understand which semantic variables reliably affect word production.

Table 1 provides definitions of the semantic variables that we focus on here and summarises evidence from previous behavioural and evoked responses (i.e., event-related potential, ERP, from electroencephalography, EEG or event-related field, ERF, from magnetoencephalography, MEG) research (see also Lampe et al., in press, for review). There is still relatively little empirical evidence regarding effects of word-specific semantic variables on word production and limited understanding of the underlying mechanisms. Moreover, most previous research was behavioural in nature and few studies have complemented behavioural investigations with investigations of ERPs or ERFs. The aim of this study was therefore to further investigate effects of semantic variables, by examining their effects on electrophysiological responses during word production.

Different semantic variables have been found to be either positively or negatively correlated with response times and naming accuracy (Table 1). For example, we (Lampe et al., in press) found that naming was faster and more accurate for words with many semantic features (replicating Rabovsky et al., 2016, 2021; Taylor et al., 2012). In contrast, naming was more error-prone for words with higher intercorrelational density, a measure argued to capture the size and activation strength of the co-activated lexical cohort (replicating Rabovsky et al., 2016, 2021). Positive effects of semantic variables (i.e., faster and more accurate responses with increasing values of a variable) are usually attributed to activation spreading between related concepts at the semantic level, which, in turn, increases target activation at the lexical level (e.g., Abdel Rahman & Melinger, 2009). However, if this spread of activation causes the co-activation of many semantically related lexical representations, negative effects arise (i.e., slower and more error prone responses with increasing values of a variable), which have been hypothesised to be due to enhanced competition between lexical representations (e.g., Abdel Rahman & Melinger, 2009, 2019). Thus, what determines the direction of the behavioural effect of a semantic variable is, most likely, the degree to which it captures semantic activation spread or lexical co-activation, and the balance between these two processes. Importantly however, not all theoretical accounts include competition (Dell, 1986; Oppenheim et al., 2010) and some authors argue that negative influences of co-activated representations on word production occur at a later, non-lexical, stage of processing (Mahon et al., 2007).

Table 1

Semantic variables: Definitions and behavioural and ERP/ERF effects on picture naming in neurotypical adults

Study	Participants / Items (*n*)	RT	Accuracy	ERP/ERF evidence
Number of semantic features				
average number of properties generated for a concept by participants in feature generation tasks; operationalised using feature databases				
Lampe et al., in press	85 / 291	↗	↗	
Rabovsky et al., 2016	16 / 541	↗	↗	
Rabovsky et al., 2021	31 / 345	↗	↗	enhanced posterior positivity for words with higher numbers of semantic features in mean amplitude analysis (200–550ms) [a] and in time course analysis ~330–600ms [b]
Taylor et al., 2012	20 / 302	↗		
Clarke et al., 2013 [c]	11 / 350			ø
Intercorrelational density				
highly intercorrelated features (e.g., *has fur, has four legs, has paws, has whiskers*) co-occur across concepts and characterise clusters of closely related concepts (e.g., 'cat', 'dog', 'tiger'); operationalised using feature databases				
Lampe et al., in press	85 / 291	ø	↙	
Rabovsky et al., 2016	16 / 541	↙	↙	
Rabovsky et al., 2021	31 / 345	ø [d]	↙	enhanced posterior positivity for words with higher intercorrelational density in mean amplitude analysis (200–550ms) [a] and in time course analysis ~335–450ms [b]
Taylor et al., 2012 [e]	20 / 302	ø		
Clarke et al., 2013 [c e]	11 / 350			increased MEG response for items with more weakly correlated semantic features from ~224ms (right ventral and anterior temporal regions and bilateral prefrontal cortex)
Number of near semantic neighbours				
number of words that share a substantial part of their semantic information with the target word (e.g., feature vector cosine similarity of > .4); operationalised using feature databases				
Lampe et al., in press	85 / 291	ø	ø	
Fieder et al., 2019 [f]	30 / 180	↙	↙	
Mirman, 2011 [f]	35 (older adults) / 57	ø	↙	
Hameau et al., 2019	40 / 84	ø [g]	ø [g]	
Lampe et al., 2017	15 (older adults) / 44	ø	ø	

Study	Participants / Items (*n*)	RT	Accuracy	ERP/ERF evidence
Bormann, 2011 [h]	18 / 54	∅		
Semantic similarity average featural similarity of the target word and (all) other words in the mental lexicon; operationalised using feature databases				
Lampe et al., in press	85 / 291	∅	∅	
Fieder et al., 2019 [f]	30 / 180	∅	↙	
Typicality representativeness of a concept for its semantic category; often operationalised as rating, but also using feature databases				
Lampe et al., in press	85 / 291	∅	∅	
Dell'Acqua et al., 2000	84 / 266	↗		
Grossman et al., 1998	14 (older adults) / 72	↗		
Holmes & Ellis, 2006 [i]	25 / 84	↗	∅	
Jolicoeur et al., 1984	18 / 48	↗		
Fieder et al., 2019 [f]	30 / 180	∅	↗	
Morrison et al., 1992 [j]	20 / 48	∅		
Woollams, 2012 [k]	16 / 80	∅	∅	
Rogers et al., 2015	12 (older adults) / 48		↙	
Distinctiveness average degree to which the semantic features of a concept are shared with other concepts (e.g., *has four legs*) or is more unique to a particular concept (e.g., *meows*); operationalised using feature databases				
Lampe et al., in press	85 / 291	∅	↙	
Rabovsky et al., 2016 [l]	16 / 541	↗	↗	
Taylor et al., 2012	31 / 345	↗ [m]		
Humphreys et al., 1988	20 / 76	↗		
Miozzo et al., 2015 [n]	17 / 146		∅	effect from ~150ms (posterior inferior temporal area)
Clarke et al., 2013 [c]	11 / 350			increased MEG response for items with more shared relative to distinctive semantic information ~84–120ms (left ventral temporal cortex, anterior temporal lobe) and ~170–210ms and 240–300ms (left ventral stream); increased MEG response for items with more distinctive relative to shared semantic information ~240–300ms (left ventral temporal cortex)

Study	Participants / Items (*n*)	RT	Accuracy	ERP/ERF evidence

Note. RT = response time, ↙ = poorer performance (slower RTs and decreased accuracy with higher values of semantic variable), ↗ = improved performance (faster RTs with increased accuracy and higher accuracy with higher values of semantic variable), Ø = no effect, blank cells = not investigated, Participants = where not otherwise specified, young adults (typically undergraduate students).

[a] mean amplitude analysis conducted on combined data of two naming rounds.

[b] from visual inspection of plots with data of the first naming round.

[c] no behavioural analysis conducted; MEG analysis on the first 300ms; measures obtained from PCA: 'number of features' – number of semantic features; 'correlational strength' – combined mean correlational strength of shared features within and across concepts; 'relative distinctiveness' – combined relative amount of shared to distinctive information and correlational strength of the distinctive features.

[d] inhibitory effect on RT significant in a second round of naming.

[e] correlation measure based on 'intercorrelational strength'.

[f] speeded naming paradigm (500ms deadline).

[g] measure was derived from Principal Component Analysis and combined number of near feature-based neighbours and number of rated category neighbours; results based on Holm-Bonferroni corrected *p*-values to account for testing for multiple variables.

[h] measure was based on ratings capturing the number of category coordinates of a target, which was then dichotomised into words with many and few competitors.

[i] similar results in picture naming after familiarisation and a subsequent second round of naming.

[j] rated typicality in two categories: man-made versus naturally occurring objects.

[k] pre rTMS.

[l] distinctiveness added in additional analysis in discussion; non-significant when intercorrelational density was in the model at the same time.

[m] naming was also faster for concepts with more highly correlated distinctive features.

[n] measures obtained from PCA: 'Specific Semantic Features' – combined number of distinctive features and number of encyclopaedic features.

Evidence from evoked responses

In previous studies, the theoretical interpretations of mechanisms underlying effects of semantic variables have been inferred from behavioural measures, such as the duration of word planning (response latency), the accuracy of the response, or type of naming error produced. Yet, response time and naming accuracy measures do not elucidate processes that unfold *during* word production. Therefore, EEG and MEG methods have been introduced to word production research aiming to provide an insight into the brain correlates underlying our behaviour and into the temporal development (i.e., time course) of processes in word production (e.g., Aristei et al., 2011; Clarke et al., 2013; Costa et al., 2009; Rose & Abdel Rahman, 2017; Valente et al., 2014).

Most previous investigations of item-inherent word characteristics using ERPs or ERFs have examined a limited number of (i.e., one or two) lexical variables in factorial designs to establish an electrophysiological index of lexical access in word production (e.g., word frequency, Laganaro, 2014; Levelt et al., 1998; Piai et al., 2012; Strijkers et al., 2010; name agreement, Cheng et al., 2010; age of acquisition, Laganaro, 2014; Laganaro et al., 2012; Laganaro & Perret, 2011). In contrast, Valente et al. (2014) conducted a more comprehensive study to localise effects of various variables (i.e., visual complexity, image agreement, familiarity, frequency, name agreement, age of acquisition, word length, phonological neighbourhood, and phonotactic probability) hypothesised to be associated with several different stages of word production (see Figure 1 in Valente et al., 2014).

Similar to behavioural investigations of semantic effects in spoken word production, previous investigations using ERPs or ERFs have predominantly examined contextually-driven semantic facilitation and interference using manipulations like Picture Word Interference (e.g., Aristei et al., 2011; Blackford et al., 2012; Hirschfeld et al., 2008; Python et al., 2018b; Rose et al., 2019), Blocked Cyclic Naming (Aristei et al., 2011; Maess et al., 2002; Python et al., 2018a), or cumulative semantic interference in a Continuous Naming Paradigm (Costa et al., 2009; den Hollander et al., 2019; Rose & Abdel Rahman, 2017). However, in context manipulation paradigms, the observed brain responses and behaviours are caused by the interplay of processing the target word and the semantic context. In addition to often higher task demands compared to standard picture naming, processing in context manipulation tasks may also be influenced by control mechanisms and attentional processes. Consequently, findings from such paradigms may not be generalisable to standard picture naming. Using standard picture naming to study effects of item-inherent semantic variables provides a means of gaining insight into, and further broadening our understanding of, semantic and lexical processing in word production without the impact of these additional cognitive demands.

Very few previous studies have investigated effects of semantic variables on word production using ERPs or ERFs. In an MEG study, Clarke et al. (2013) studied how visual stimuli invoke object meaning by testing for effects of perceptual and feature-based semantic variables and their time course in the first 300ms of word planning. They used a Principal Component Analysis (PCA) to orthogonalize and combine variables and identified two perceptual components: image complexity and size, loaded on the component *image complexity* and concept familiarity and picture exemplarity, loaded on the component *familiarity*. In addition, there were four feature-based semantic components: Feature-based semantic measures capturing the amount of shared or distinctive information associated with a target concept and the correlation of its distinctive features loaded highly on a *relative distinctiveness* component and measures capturing the regularity with which shared features co-occurred, loaded on the component *correlational strength*. Moreover, the *number of semantic features* and *proportion of visual features* loaded highly on two further feature-based components.

Clarke et al. (2013) found that shared semantic feature information as captured by the relative distinctiveness component affected processing as early as 84ms post picture onset. Specifically, there was an increased MEG signal for items with more shared relative to distinctive information, suggesting that shared semantic information was rapidly extracted from the visual input. Around 240–300ms this effect was reversed with stronger MEG signals for words with more distinctive information, suggesting that distinctive aspects of meaning were being accessed for more fine-grained semantic processing. The correlational strength component was significant from around 224ms onwards, with increased MEG responses for concepts with more weakly correlated features. In contrast, number of semantic features and proportion of visual features did not affect processing in the first 300ms of word planning. However, given their focus on perceptual and semantic processing, Clarke et al. did not discuss consequences of the semantic variables on later stages for word production, most importantly lexical processing.

Similarly, in another MEG study, Miozzo et al. (2015) tested for effects of semantic and phonological variables in word production. They also combined several measures in a PCA. The semantic component contained a measure of the number of distinctive features (a measure related to the mean distinctiveness measure used in this study and by Clarke et al., 2013) and number of encyclopaedic features. Similar to Clarke et al., they reported an early onset of the effect of this semantic component (from around 150ms post picture onset) but in a posterior inferior temporal area.

Rabovsky et al. (2021) were the first to investigate effects of semantic variables on word production planning using EEG, studying the feature-based measures number of semantic features and intercorrelational density. In the context of a competitive theory of word production (Swinging Lexical Network Account; Abdel Rahman & Melinger, 2009, 2019), they assumed that any facilitatory effect of a semantic variable from semantic processing would be overpowered by inhibitory effects if a sufficiently strong cohort of competitors was co-activated at the lexical level. Rabovsky et al. predicted that number of semantic features should index the facilitatory processes associated with semantic processing and intercorrelational density the inhibitory processes associated with lexical processing.

The Swinging Lexical Network Hypothesis also proposes that semantic and lexical processing are interactive and overlapping in time. Consequently, Rabovsky et al. (2021) predicted that the ERP effects of number of semantic variables and intercorrelational density would occur around the same time. Following previous research (Costa et al., 2009; Rose et al., 2019; Rose & Abdel Rahman, 2017), lexical competition (predicted for words with higher intercorrelational density) was expected to be reflected by an increased posterior positivity occurring roughly around 200ms post picture onset. However, Rabovsky et al. had no clear expectations regarding the ERP signature of the predicted facilitatory effect of number of semantic features. Indeed, an increased posterior positivity was found for words with higher intercorrelational density *and* for words with a larger number of semantic features in the pre-defined time-window of interest between 200 and 550ms post picture onset. In a more fine-grained analysis between picture onset and 1000ms post-picture onset, they found significant effects between approximately 230 and 470ms for intercorrelational density and around 320 and 600ms for number of semantic features (not corrected for multiple comparisons) when combining across two presentation cycles of the items[1].

Rabovsky et al. (2021) argued that the posterior positivity represented higher levels of activation in the lexical semantic system when producing words with higher intercorrelational density or higher numbers of semantic features, rather than competitive processes as had been claimed previously (Costa et al., 2009; Rose et al., 2019; Rose & Abdel Rahman, 2017). For words with a higher number of semantic features, this higher activation was claimed to be mostly related to the target word's representation, due to its strong activation from many semantic features. In contrast, for words with higher intercorrelational density, it was thought to be more widespread across the target and its competitors. In sum, this study demonstrated that item-inherent variations in semantic variables can be used to investigate processes during word production using EEG. While the behavioural effects for number of semantic features (facilitatory) and intercorrelational density (inhibitory) were of opposing polarity, Rabovsky et al. (2021) found that their electrophysiological effects were very similar.

Reproducibility is a central scientific principle (e.g., Open Science Collaboration, 2015) and it is crucial to test the validity of previous findings in replications. Replications also enable extension of the to-be-replicated analyses with additional variables and testing of the robustness of the original finding following a different analysis approach. This was the aim of our study.

Rabovsky et al. (2021) included only two semantic variables and few psycholinguistic control variables in their analyses (i.e., familiarity, number of orthographic neighbours, lexical frequency, and visual complexity (subjective and objective) and their interactions with task repetition). As variables influencing behaviour do not operate selectively, but rather simultaneously, analyses can be strengthened by inclusion of a variety of variables known to have an effect on word production. This enables identification of the unique effect of a variable of interest and avoidance of false positive

[1] The effect of number of semantic features was significant in the first naming cycle between approx. 330 and 600ms and only sporadically in the second naming cycle. In contrast, the effect of intercorrelational density was significant in only some time-windows between approx. 335 and 450ms in the first naming cycle and continuously significant between about 240 and 470ms in the second naming cycle. Rabovsky et al. (2021) did not report the significance of the effects of the semantic variables in the mean amplitude analysis separately for the two presentation cycles.

findings (see Lampe et al., in press; Lampe, Hameau, Fieder, et al., 2021, for more detailed discussion of this topic). Consequently, in our study we replicated and extended Rabovsky et al. (2021) by including additional semantic and control variables.

Moreover, Rabovsky et al. (2021) only analysed a posterior ROI, which previous work (Costa et al., 2009; Rose et al., 2019; Rose & Abdel Rahman, 2017) had associated with competition during lexical selection. However, this focus risks missing effects related to semantic processing in other regions of the brain (e.g., Binder et al., 2009; Clarke et al., 2013; Indefrey, 2011). In addition, other research has also argued for additional brain regions being involved in lexical selection, particularly in the left mid MTG (middle temporal gyrus; see Indefrey, 2011, for a review) and the left IFG (inferior frontal gyrus; e.g., Schnur et al., 2005; Thompson-Schill et al., 1997). An approach that does not require pre-defined ROIs, but rather uses ERP data from the whole brain, can avoid these issues. One such approach that goes beyond descriptions of differences in waveforms depending on experimental conditions or variables is *spatio-temporal segmentation*, or *microstate analysis* (Lehmann et al., 1987; see e.g., Laganaro, 2014; Laganaro et al., 2012; Python et al., 2018a, 2018b; Valente et al., 2014). This analysis tests whether conditions or variables are associated with different brain configurations and does not require regions and time-windows of interest to be defined. Instead, it utilises the rich spatial information of the ERP signal (Jia, 2019).

In microstate analysis, the ERP signal is characterised based on the topography and temporal dynamics of electric fields at the scalp (Brunet et al., 2011; Poulsen et al., 2018) to determine periods of quasi-stable EEG topographies, so called *microstates*. Individual microstates are thought to be caused by synchronised activation of large neuronal networks, which correspond to different neural processes in the brain (Brunet et al., 2011). Hence, the idea is that different microstates reflect different neuronal networks involved in word production. Processing in these neuronal networks may be influenced by item-inherent variables which can be investigated by testing for effects of these variables. In addition, if effects of item-inherent word characteristics originate during different stages of word production (e.g., semantic or post-lexical processing stages) they would be expected to affect different microstates. For example, Valente et al. (2014) took this approach and used a microstate analysis to investigate effects of item-inherent word characteristics during word production. Conducting an analysis at the single trial level, they were able to test if different item-inherent word characteristics (e.g., image agreement, age of acquisition, and name agreement) affect different underlying neuronal networks (i.e., the different microstates). Importantly, this approach allowed the consideration of multiple continuous variables of interest simultaneously and to describe their influence on the whole word production process. Notably, microstate analyses allow researchers to go beyond the investigation of timing of processes of word production by investigating changes in the underlying networks across conditions or item sets.

The current research

With this study we wished to better understand the extent to which, and how, item-inherent semantic variables influence word production. Specifically, the aim was to determine whether, during word production, there were any electrophysiological signatures of particular item-inherent semantic variables by investigating the electrophysiological modulations they induced.

To date, the only study to use EEG in word production to study semantic variables, Rabovsky et al. (2021), focused on the number of semantic features and intercorrelational density. We considered it important to avoid the potential difficulty interpreting any lack of replication of their results in the context of more complex analyses. Consequently, in our first analysis, we conceptually replicated Rabovsky et al.'s approach with a waveform analysis in a posterior region of interest. In this analysis we focused on i) the effects of the semantic variables on the mean ERP amplitude between 200 and 550ms and ii) on the time course of effects of the semantic variables across 10ms time segments between 0 and 550ms. However, we also extended their approach by including additional feature-based semantic variables, which have been suggested by the behavioural literature to influence word production: number of near semantic neighbours, semantic similarity, typicality, and distinctiveness, in addition to number of semantic features and intercorrelational density. Hence, we *simultaneously* studied effects of the six feature-based semantic variables on brain processes during word planning to account for possibly confounding effects[2]. Moreover, we thoroughly controlled for a wide range of psycholinguistic variables that impact word production (Perret & Bonin, 2019): name agreement, image agreement, imageability, age of acquisition, conceptual familiarity, lexical frequency as well as ordinal category position (Howard et al., 2006), and the item's trial number in the experiment (Baayen & Milin, 2010).

In addition to the waveform analysis, we conducted a microstate analysis to determine whether underlying networks (or stable configurations at the surface) were modulated by the semantic variables. Running a trial-by-trial multiple regression analysis (see Valente et al., 2014) we determined if the six semantic variables affected the durations of periods of stable electrophysiological patterns, while including the same psycholinguistic control variables as in the mean amplitude analysis.

Methods

The behavioural methods for this study were identical to Lampe et al. (in press), which reports analyses of the behavioural picture naming responses that were collected together with the EEG data analysed here. Below we give an overview of the method and refer the reader to Lampe et al. for further detail. The approach to the waveform analysis (replication and extension of Rabovsky et al., 2021) was preregistered on the Open Science Framework (Lampe et al., 2019; https://osf.io/yw6ma/).

Participants

Eighty-nine participants took part in the picture naming study, 83 of whom had EEG data available (6 pilot participants were only included in the behavioural analysis in Lampe et al., in press). Two further participants were excluded as they did not perform the task as instructed or did not fulfil the eligibility criteria. In addition, three participants were excluded from the EEG analysis due to experimental control errors, resulting in a final sample size of 78 participants (64 female, age: M = 20.0

[2] Similar to the analysis of the behavioural data in Lampe, Hameau, and Nickels (in press), we conducted two further analyses to replicate Rabovsky et al. (2021) analyses as closely as possible by including only number of semantic features and intercorrelational density as semantic variables. In a first analysis, the choice of semantic and psycholinguistic control variables completely replicated Rabovsky et al.. Subsequently, we extended the analysis with further psycholinguistic control variables, and finally, in the analysis reported here, with the remaining semantic variables. This increasingly more complex data analysis allowed a direct comparison of our findings with Rabovsky et al.. The two additional analyses and their results are reported in Appendix A.

years, range = 17–33 years, *SD* = 2.2). All participants were Australian English native speakers, right-handed, had normal or corrected-to-normal vision, and no history of neurological, cognitive speech and language impairments. Participants were recruited through Macquarie University's Psychology participant pool and received course credit or a monetary compensation (AUD15 per hour). They were tested individually in a shielded room after giving informed consent. The study was approved by Macquarie University's Human Research Ethics Committee.

Stimuli

The stimuli consisted of colour photographs on white background of 297 items with high name agreement for Australian English from the feature database by McRae et al. (2005) (see Lampe et al., in press, for a description of the selection process). The stimuli were presented in six pseudorandomised lists and every participant saw one list. Each list consisted of four blocks. The items in Block 1 (*n* = 35) were selected from different semantic categories such that the response latencies for these items would be unaffected by cumulative semantic inhibition (e.g., Howard et al., 2006). In Blocks 2–4 (*n* = 87 or *n* = 88 items), items from the same semantic category were separated by at least two intervening items from different categories to reduce the influence from semantically related items. Three pseudorandomisations were created of each block. After Block 1, Blocks 2–4 were presented in a varying order in the different Lists (e.g., Block 3 after Block 2 in List 1 but Block 2 after Block 3 in List 4).

For all items, information on the six feature-based semantic variables was retrieved from, or calculated based on, information given in McRae et al. (2005): number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness.

Number of semantic features was a count of the features reported for each word in the McRae et al. (2005) database (e.g., Rabovsky et al., 2016). *Intercorrelational density* was the summed shared variance of all of a concept's correlated feature pairs (e.g., Rabovsky et al., 2016). *Number of near semantic neighbours* was a count of words whose feature vectors had a similarity of at least .4 with the feature vector of the target (Hameau et al., 2019; Mirman, 2011; Mirman & Graziano, 2013). *Semantic* *similarity* was the mean similarity between the feature vectors of the target word and all other words in the McRae et al. (2005) database (Mirman & Magnuson, 2008). *Typicality* was calculated similar to Rosch and Mervis' (1975) family resemblance score: First, each feature was weighted by the number of items of its semantic category that were also identified as having that feature by participants generating feature norms. Then, the feature weight was divided by the number of items in the semantic category, before each feature was weighted by its production frequency. Ultimately, the weights of all features of an item were summed to form the typicality measure. Lastly, *distinctiveness* was the inverse of the number of concepts that were credited with a particular feature in the whole database, which was averaged across all features of a concept (e.g., Rabovsky et al., 2016). Please refer to Lampe et al. (in press) and Lampe, Hameau, Fieder, et al. (2021) for more information on the calculation of the semantic variables.

The psycholinguistic control variables for this experiment included the item-characteristics name agreement, image agreement, imageability, age of acquisition, familiarity (all from Lampe et al., in press), rated visual complexity, and spoken word frequency (Zipf, SUBTLEX-UK; van Heuven et al., 2014). In addition, we derived two further control variables from the experiment itself: A category position measure to control for the cumulative semantic inhibition effect (Howard et al., 2006), which was the rank-order of an item in its same semantic category and the rank-order of an item in the experimental list to control for fatigue or habituation (Baayen & Milin, 2010).

Before running the analyses, six items (i.e., raft, crowbar, bridge, pie, racquet, and board) were excluded either because the number of errors was high compared to that of the other items or because participants frequently responded with a compound noun that was a subordinate to the target word (e.g., pie → meat pie). This resulted in a final number of 291 items (as in Lampe et al., in press).

Procedure

Each trial began with a fixation cross in the centre of the screen. To prevent the participants from predicting the exact onset of the picture on the screen, the fixation cross was shown for a random duration between 500 and 1000ms. Next, a single picture was displayed for 2000ms on white background, which participants were instructed to name as quickly and accurately as possible using a single word only. After the picture disappeared, the screen was blank for 1000ms before the next trial started.

There was no familiarisation phase prior to the experiment but there were five practice trials at the beginning of the experiment and each experimental block also started with a practice item. All practice items came from different semantic categories to the experimental stimuli. Between blocks, participants were given a break to ask questions or rest. The experiment took approximately 30 minutes to complete and, subsequently, the participants completed other tasks as part of a larger study (e.g., Lampe, Hameau, & Nickels, 2021).

The EEG data acquisition was controlled on an iMac with macOS version 10.14.1, while the experiment was presented in Presentation® (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) on a Dell Precision tower 3620 running Windows 10, using an AOC FreeSync LED monitor. Behavioural data were recorded with a Behringer preamplifier (Tube Ultragain Mic100) and a Rode NTG1 shotgun microphone. The spacebar of the keyboard was used to navigate through the experiment.

EEG recording and pre-processing

The continuous EEG signal was recorded using an ActiveTwo Biosemi system with 64 channels (Biosemi, Amsterdam, the Netherlands), positioned according to the extended 10–20 system (Jasper, 1958) and using Biosemi's active Ag/AgCl electrodes. Signals were sampled at 1024 Hz. EEG data processing was performed in MATLAB (R2016b, MathWorks Inc.) using the free EEGLAB toolbox (Version 14.1.2, Delorme & Makeig, 2004) for data cleaning and structuring.

After acquisition, the EEG data were first down-sampled to 500Hz and re-referenced against the average reference. The data was then filtered using a Butterworth bandpass filter with a high-pass cut-off of 1Hz and low-pass cut-off of 30Hz. All trials that did not contain exclusively the target word (e.g., naming errors, disfluencies, or elaborations; see section "Behavioural data response coding") and responses that were faster than 550ms were removed to get a maximally uncontaminated signal. On average, 57.71 trials were rejected per participant (range = 22–117, *SD* = 19.64). The EEG data were segmented into epochs of 750ms, starting 200ms before the onset of the target stimulus on the screen and ending 550ms after stimulus presentation. We were interested in processes occurring during semantic and lexical processing, long before the onset of articulation (Indefrey, 2011). Given that very fast responses (< 550ms) were removed, this time-window was wider than the fastest trial. Moreover, as mean picture naming latencies were 900ms for correct responses (Lampe et al., in press), and estimates of the time course of word production (Indefrey, 2011; Indefrey & Levelt, 2004) locate semantic and lexical processing between picture onset and 275ms post picture onset, this time-window was anticipated to include the processes of interest for this analysis.

Noisy channels were identified using the PREP pipeline (Bigdely-Shamlo et al., 2015) and temporarily excluded for the subsequent processing steps. Eye movement, heart, and muscle artifacts were identified and eliminated using the automatic independent component analysis (ICA) procedure of EEGLAB and the plugin ICLabel (Version 1.2.4; Pion-Tonachini et al., 2019) (mean number of components excluded = 2.34, range = 0–5, *SD* = 1.23). Then, any bad channels identified with the PREP pipeline were interpolated (mean number of channels interpolated per participant = 4.29, range = 0–17, *SD* = 3.25).

The epochs were baseline corrected using the average amplitude of the 200ms before the onset of picture presentation on the screen. Individual epochs were screened to detect channels with absolute amplitudes exceeding 100μV. On average, 2.47 epochs were excluded per participants (range = 0–19, *SD* = 3.45). The final trial number was 236.82 on average (range = 180–273, *SD* = 19.96).

Analyses

Behavioural data response coding

Only correct responses were included in the EEG data analyses. Correct responses were responses where the participant produced the exact target word as a first response (e.g., “truck”). In addition to overt naming errors and omissions, responses where a determiner preceded the correct name (e.g., “a truck”), participants were dysfluent (e.g., “t truck”) or hesitant (e.g., “erm truck”), or produced an elaboration (e.g., “yellow truck”), synonym (e.g., “lorry”), or abbreviation (e.g., “bike” instead of “motorbike”) were excluded from the analysis.

EEG data analysis

Analysis 1: Waveform analysis. Based on Rabovsky et al. (2021) (and therefore following Costa et al., 2009; Rose et al., 2019; Rose & Abdel Rahman, 2017), we analysed a cluster of posterior electrode sites comprising CP3, CP4, Pz[3], P3, P4, P5, P6, PO3, PO4, and POz. For this ROI, we ran single-trial linear mixed models using the lme4 package (Version 1.1-21, Bates, Mächler, et al., 2015; *p*-values were retrieved with lmerTest, Version 3.1.1, Kuznetsova et al., 2017) for RStudio (Version 1.3.959, RStudio Team, 2020) on mean ERP amplitudes in a 200-550ms time-window to test for influences of the semantic variables (*mean amplitude analysis*). This time-window was chosen following Rabovsky et al., who based it on previous research concerning semantic context effects in language production research (Aristei et al., 2011).

We extended Rabovsky et al. (2021) to new semantic variables to investigate the electrophysiological correlates of these semantic variables. The statistical analysis included all 6 semantic variables of interest (number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness). Moreover, the model also included psycholinguistic control variables that have been demonstrated to influence picture naming performance (name agreement, imageability, age of acquisition, familiarity, frequency, and measures of ordinal position within a category and within the list; Baayen & Milin, 2017; Howard et al., 2006; Perret & Bonin, 2019). All control and semantic variables were standardised using a *z*-transformation[4]. Finally, random intercepts for participants and items were included in the model as well as random slopes for semantic variables for participants. Following the model definition approach by Bates, Kliegl, et al. (2015) we only kept random slopes that increased the models' fit, which was assessed using likelihood ratio tests (stats package, Version 3.6.1; R Core Team, 2019). No random slopes were kept in any of the waveform analyses.

[3] Please note that Pz was erroneously not included in our preregistration, as it was not listed in the Methods section of the preprint (Rabovsky et al., 2018) of Rabovsky et al. (2021).

[4] The standardisation of the measures ordinal category position and item number in the experiment was not preregistered. However, standardisation of all variables was necessary to facilitate model convergence.

In addition, we replicated Rabovsky et al.'s (2021) approach to exploring the temporal dynamics of the observed effects in more detail by analysing consecutive 10ms segments between 0 and 550ms[5] post-stimulus within the same ROI (*time course analysis*). 18,424 data points from 78 participants and 291 items entered the analyses per time-window.

However, studies using a time course analysis often do not correct for the fact that numerous models were run (e.g., Rabovsky et al., 2021; Rose & Abdel Rahman, 2017), which is problematic as the probability of finding erroneously significant time-windows increases with the number of tests conducted. One reason for the omission of correction for multiple comparison may be that it is not quite clear which of the numerous approaches is best suited for time course data. Here we therefore explored different approaches to this problem and corrected for multiple testing in several different ways.

The most common correction for multiple testing is the Bonferroni correction, which corrects the p-value according to the family-wise error rate. To apply this correction, one divides the significance level (alpha) by the number of tests conducted (n = 55 in this study). In our case, this results in a significance level of .00091. However, importantly, the Bonferroni correction treats observations as independent, yet adjacent sampling points of ERP data tend to be correlated. In contrast, a correction of the false discovery rate estimation that controls the error rate among a set of tests is less conservative. One implementation is the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995), which allows for test statistics to be positively dependent (Benjamini & Yekutieli, 2001). For other dependencies, Benjamini and Yekutieli suggested the more conservative Benjamini-Yekutieli method. These approaches to control the false discovery rate were applied with a q-level of 0.05 (i.e., 95% of detections are true detections; Anders et al., 2019) and a less conservative q-level of 0.1 (Ewald et al., 2012; Genovese et al., 2002). Another approach to correct for multiple testing is to accept effects as significant only if they exceed a certain duration (e.g., 10ms, Laganaro et al., 2012; 20ms, Laganaro & Perret, 2011; 30ms, Valente & Laganaro, 2015) with effects being present at a few

[5] Rabovsky et al. (2021) ran this analysis on consecutive 10ms segments between 0 and 1000ms, however, we did not want to run an analysis on a time window larger than the shortest epoch in the dataset.

(e.g., at least five; Laganaro, 2014; Valente & Laganaro, 2015) adjacent electrodes at a conservative alpha criterion (e.g., .01). Consequently, here we also tested if effects of semantic variables survived a 20ms cut-off at a significance criterion of .01.

In two additional analyses, we investigated whether the effects reported by Rabovsky et al. (2021) were replicable when using 1) the same model structure as Rabovsky et al. or 2) improved control of the psycholinguistic variables. The model directly duplicating the analysis approach by Rabovsky et al. included number of semantic features and intercorrelational density as semantic variables and familiarity, frequency, orthographic neighbourhood density, and visual complexity as control variables. In the second analysis, we included a larger set of control variables known to influence spoken word production: name agreement, imageability, age of acquisition, familiarity, frequency, ordinal position within a category and within the list. As orthographic neighbourhood density and visual complexity are not commonly influential predictors of spoken word production (Perret & Bonin, 2019) they were excluded from this and the subsequent analysis. The results of these two analyses are reported in Appendix A.

Analysis 2: Microstate analysis. Following Valente et al. (2014) we conducted a microstate analysis, compressing the variability in the EEG signal to form template maps of quasi-stable global electrophysiological patterns at the scalp using spatio-temporal clustering. We used the Microstates plugin (MST, Version 1.0; Poulsen et al., 2018) for EEGLAB for this analysis.

Spatio-temporal segmentation was conducted on the subject-averaged ERP data between 50 and 550ms using a K-means clustering algorithm with 5000 runs of randomisations. The ideal number of ERP maps that best explained the averaged data was selected using a combination of Global Estimated Variance and Cross-Validation criteria (Pascual-Marqui et al., 1995; see also Murray et al., 2008). The sequence of maps also had to align with the maxima of the Global Field Power (i.e., the standard deviation of activity over all electrodes; Cohen, 2014). In order to remove maps with low explanatory power and periods of topographic instability, topographic maps were rejected if they were shorter than 20ms and it was checked that no maps were more than 97% correlated. The microstates determined from the group-averaged data using this procedure were then back-fitted to the single trials of each participant. In this fitting procedure, each time point of the single trials was attributed to a microstate on the basis of the spatial correlation between the template maps and every single time point of the individual trial's ERP.

Statistical analyses comprised linear mixed effects models for each topographical map in order to determine the effects of semantic and psycholinguistic control variables on the summed duration (i.e., the number of timeframes associated with each microstate) of each stable topographical map. Given the relatively low temporal precision of microstate analyses, we aimed to use the effect of lexical frequency as a marker of lexical processes (following e.g., Piai et al., 2012; Strijkers et al., 2010). Hence, the neuronal network supporting lexical selection was assumed to be reflected by the microstate with an effect of word frequency. Following previous work, this effect was expected to commence somewhere between 180 (Strijkers et al., 2010) and 290ms (Piai et al., 2012) post picture onset. Unfortunately, in our analyses (as in Valente et al., 2014), there was no evidence of lexical frequency affecting the number of timeframes associated with any microstate nor a significant effect in the waveform analysis. This lack of a significant effect of word frequency is most likely due to this measure being correlated with some of the other variables included in our analyses (r up to -.49 in correlation with age of acquisition, r = .31 with imageability, r = -.28 with image agreement, all other r < .25). Hence, frequency could not be used as a marker of lexical processing.

Results

Summary of behavioural results (Lampe et al., in press)

In the presence of only limited behavioural research into effects of feature-based semantic variables, we published a thorough investigation of the behavioural data collected in this study elsewhere (Lampe et al., in press; see also Table 1). When including the same six semantic variables that we investigated here with ERP, in models for the behavioural analysis (Analysis 2 in Lampe et al., in press), we found a facilitatory effect of number of semantic features on both naming latency and accuracy. In addition, higher intercorrelational density led to less accurate responses and higher distinctiveness had an inhibitory effect on naming latency. In contrast, there were no significant effects of semantic similarity, number of near semantic neighbours, and typicality on naming latency or accuracy.

Waveform analyses

Mean amplitude analysis

The results of the mean ERP amplitude analysis (200–550ms, posterior ROI) are summarised in Table 2 and depicted in Figure 1. There was an enhanced posterior positivity for words with higher image agreement, and lower age of acquisition. Moreover, the positivity was stronger the more items of the target word's semantic category were seen in the experiment before the target and for words that appeared earlier in the experiment. In addition, the posterior positivity was enhanced for words with a higher number of semantic features. No other semantic variables were significant in this analysis (Table 2; Figure 1).

Time course analysis

To understand the development of the effects of the semantic variables over time, separate linear mixed effect models were performed for each 10ms time segment between 0 and 550ms (see also Appendix B for the point-by-point correlations between response latency and ERP amplitudes). The development of ERP amplitudes for each semantic variable at the posterior ROI is depicted in Figure 2.

For all semantic variables, there were few significant time-windows in the first 200ms after picture onset. In addition, there were significant time windows around 200ms for number of semantic features, semantic similarity, typicality, and distinctiveness, as well as significant time windows around 350-400ms for number of semantic features, semantic similarity, and number of near semantic neighbours.

Table 2

Summarised output of linear mixed model analysis of mean amplitude

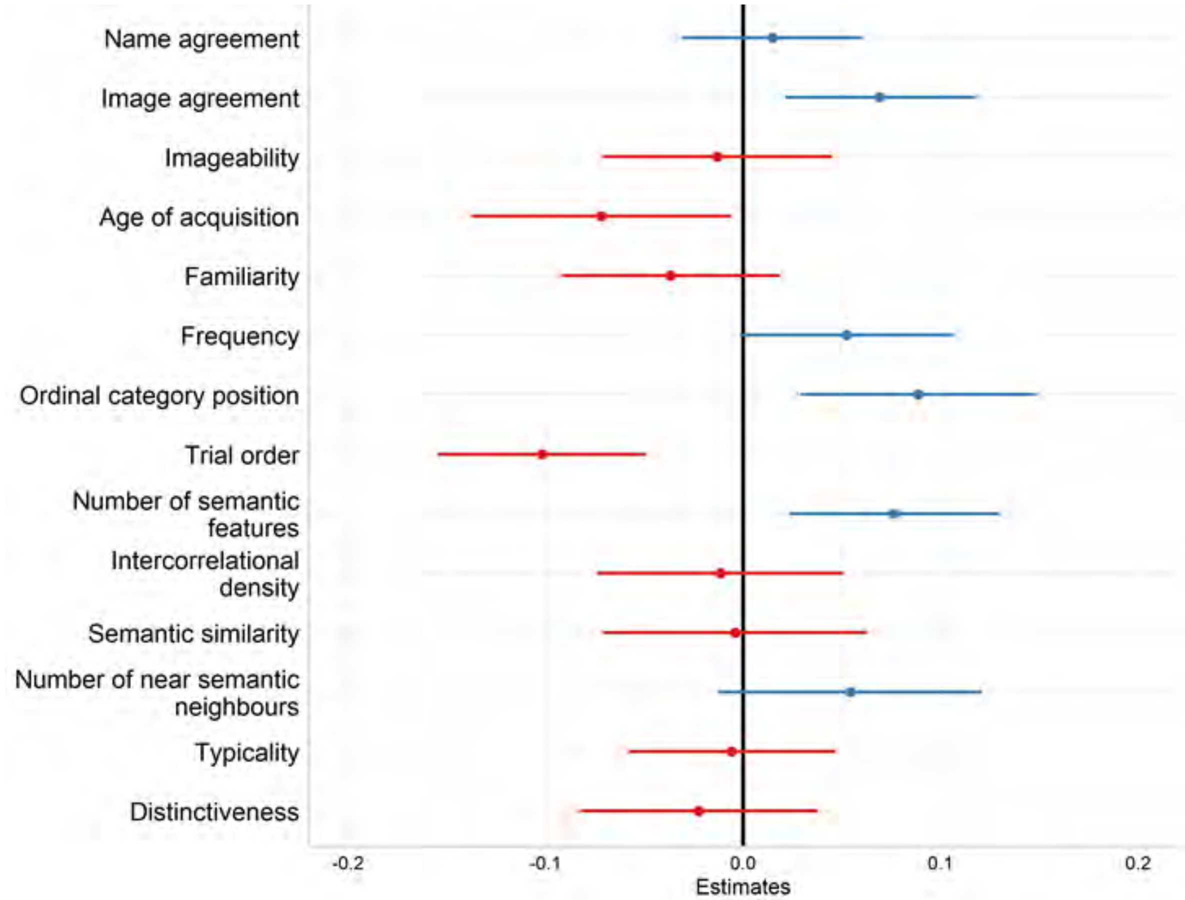
Random effect	Variance	*SD*				
Item (Intercept)	0.06	0.24				
Subject (Intercept)	0.53	0.73				
Residuals	4.55	2.13				
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF
(Intercept)	1.65	0.09	1.48 – 1.81	19.31	<0.001	
Name agreement	0.01	0.02	-0.03 – 0.06	0.64	.525	1.10
Image agreement	0.07	0.02	0.02 – 0.12	**2.82**	**.005**	1.27
Imageability	-0.01	0.03	-0.07 – 0.05	-0.44	.661	1.86
Age of acquisition	-0.07	0.03	-0.14 – -0.01	**-2.14**	**.032**	2.38
Familiarity	-0.04	0.03	-0.09 – 0.02	-1.31	.190	1.74
Frequency	0.05	0.03	-0.00 – 0.11	1.90	.057	1.57
Ordinal category position	0.09	0.03	0.03 – 0.15	**2.92**	**.004**	2.85
Trial order	-0.10	0.03	-0.15 – -0.05	**-3.79**	**<.001**	2.45
Number of semantic features	0.08	0.03	0.02 – 0.13	**2.83**	**.005**	1.61
Intercorrelational density	-0.01	0.03	-0.07 – 0.05	-0.37	.715	2.21
Semantic similarity	0.00	0.03	-0.07 – 0.06	-0.12	.904	2.52
Number of near semantic neighbours	0.05	0.03	-0.01 – 0.12	1.59	.112	2.46
Typicality	-0.01	0.03	-0.06 – 0.05	-0.22	.825	1.56
Distinctiveness	-0.02	0.03	-0.08 – 0.04	-0.73	.463	2.09
Observations	18,424					
Marginal R^2 / Conditional R^2	0.006 / 0.120					

Note. VIF = Variance Inflation Factor.

Values of significant effects (p < .05) are printed in bold.

Figure 1

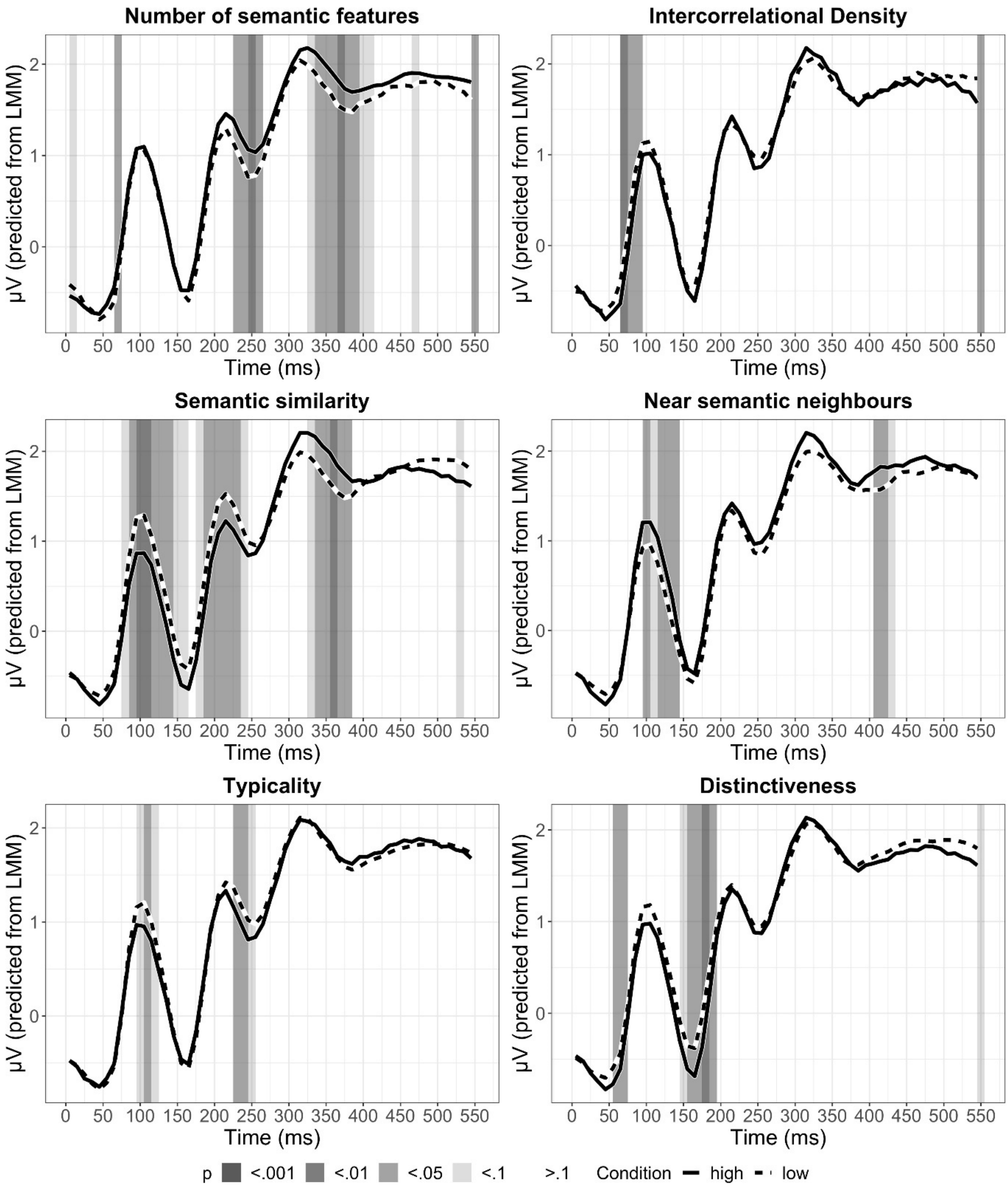
Fixed effects estimates of mean amplitude analysis with 95% confidence intervals



Note. Blue lines (to the right of centre) indicate that increased values of the variable lead to enhanced posterior positivity and red lines (to the left of centre) indicate decreased posterior positivity.

Figure 2

Linear mixed model estimates of the mean voltages ± the effect sizes of each semantic variable at the posterior region of interest in consecutive 10ms segments between 0 and 550ms



Note. Grey shading indicates levels of significance. High and low conditions of each semantic variable are based on the mean voltage (model Intercept) ± the effect size of the respective semantic variable.

However, when applying the different methods suggested to control for multiple comparisons, none of the significant time-windows reported in the time course analysis survived: None of the conducted tests reached the significance level of .00091 as suggested by the Bonferroni correction. Moreover, applying the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) or the Benjamini-Yekutieli method (Benjamini & Yekutieli, 2001) at *q*-levels of 0.05 and 0.10, again no time-window in the time course analysis remained significant for any of the semantic variables. Only when using a 20ms cut-off at a significance criterion of .01, the least conservative correction for multiple comparisons, did the enhanced posterior positivity for words with lower semantic similarity at 90–110ms survive. No effect of any other semantic variable met even this criterion. While the absence of significant effects of the semantic variables, following the different approaches to multiple comparison correction, cannot confirm the absence of any semantic effects in the posterior ROI, their consistent failure to survive multiple comparison correction suggests that the significant time-windows represent spurious effects and are likely false positives caused by multiple testing.

Microstate analysis

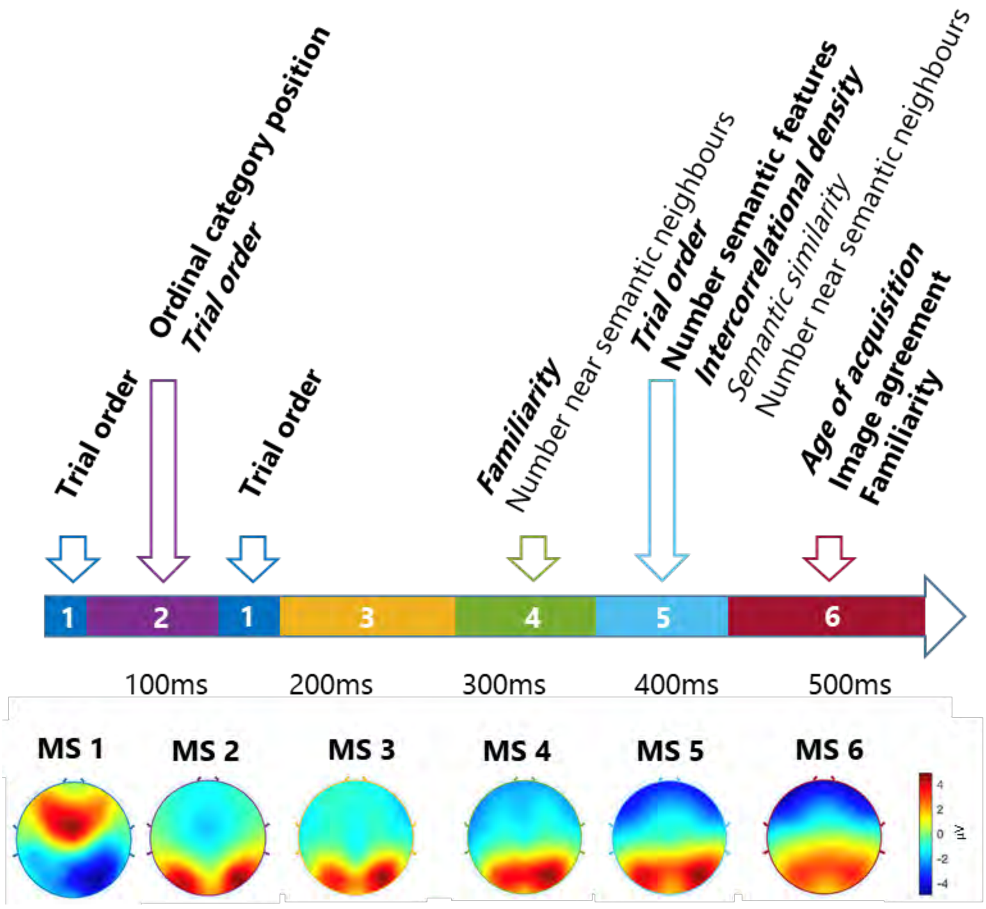
The mean amplitude and the time course analyses followed Rabovsky et al. (2021) and only included the ERP data collected at a posterior ROI. To test if any effects of semantic variables would be apparent when including the full spatial richness of the ERP data and to identify whether the underlying networks were modulated by the semantic variables, we ran a microstate analysis.

The spatio-temporal segmentation was conducted on the grand average data between 50 and 550ms after picture onset and yielded 6 topographic maps (Figure 3). During the back-fitting procedure each sampling point of every individual experimental trial was labelled with the template map with which it had the highest spatial correlation. The duration of each topographical map was then used as a dependent variable for linear mixed effects models. Note that 'duration' has to be understood as the number of timeframes associated with this map. In addition, the map onsets and offsets correspond to the first and last time point associated with a given map in the grand average ERPs.

The six semantic variables as well as the control variables were entered as fixed effects in each model. As for the previous analyses, random effects for participants and items with random slopes for the semantic variables by participants were included in the models and the random effects structure was gradually decreased following Bates, Kliegl, et al. (2015). Results of the models are summarised in Table 3 and Figure 3 (the complete statistical models are provided in Appendix C). To correct for multiple comparisons (we ran one model per topographic map), we applied the Bonferroni correction, resulting in a significance threshold of $p < .008$.

Figure 3

Microstate segmentation and summary of effects



Note. Top: Results of microstate analysis. Variables in italics resulted in shorter duration of the respective microstate the higher the value of the variable and variables in regular font resulted in longer duration of the respective microstate the higher the value of the variable. Variables that also affected behavioural measures (RTs or naming accuracy) are in bold.

Bottom: Topographic maps revealed by the spatio-temporal segmentation

Table 3

Summary of (Bonferroni corrected) significant effects of the linear mixed effects models for the duration of periods of stable electrophysiological activity (topographic maps)

Variable	Map 1 ~50–70ms and ~144–168ms	Map 2 ~70–144ms	Map 3 ~168–280ms	Map 4 ~280–364ms	Map 5 ~364–440ms	Map 6 ~440–550ms
Name agreement						
Age of acquisition						β = -0.11, t = -3.09, p = .002
Imageability						
Image agreement						β = 0.12, t = 4.51, p < .001
Frequency						
Familiarity				β = -0.08, t = -2.86, p = .005		β = 0.11, t = 3.71, p < .001
Ordinal category position		β = 0.07, t = 2.83, p = .005				
Trial order	β = 0.10, t = 2.89, p = .004	β = -0.10, t = -4.70, p < .001			β = -0.10, t = -5.05, p < .001	
Number of semantic features					β = 0.06, t = 3.26, p = .001	
Intercorrelational density					β = -0.07, t = -3.20, p = .002	
Semantic similarity					β = -0.07, t = -3.06, p = .002	
Number of near semantic neighbours				β = 0.13, t = 3.69, p < .001	β = 0.07, t = 2.75, p = .006	
Typicality						
Distinctiveness						

Note. Full statistical models are provided in Appendix C.

The duration of Map 1, which lasted from approximately 50 to 70ms as well as from 145 to 170ms post picture onset, increased with the position of the trial within the experiment (i.e., longer duration of Map 1 the later in the experiment an item occurred). In contrast, Map 2, lasting from about 70 to 145ms, decreased with the trial order of the items in the experiment and increased with ordinal category position. There was no evidence that the duration of Map 3 (approx. 170–280ms) was affected by any of the control or semantic variables included in the analysis. The duration of Map 4, which started around 280ms after picture onset and lasted for about 85ms, decreased with higher familiarity and increased with a greater number of near semantic neighbours. Map 5, which lasted from approximately 365 to 440ms, was most affected by the variables: Its duration decreased the later an item appeared in the experiment, but also with higher semantic similarity and higher intercorrelational density of an item. In contrast, its duration increased with higher numbers of near semantic neighbours and higher numbers of semantic features. Finally, the duration of Map 6 (440–550ms) decreased with higher age of acquisition and increased with higher image agreement and familiarity. Importantly, however, findings for this last map have to be treated with caution as this map was cut off as it exceeded the upper boundary of 550ms of the time window included in the analysis.

Discussion

The present study was conducted to investigate electrophysiological correlates of six feature-based semantic variables and to explore whether they influence the neuronal networks involved in word production. We studied item-inherent semantic variables, which were previously suggested to influence behavioural measures of picture naming when (mostly) studied individually: number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness. Most of these variables had not previously been investigated using ERPs (or ERFs) in standard picture naming. We conducted three different analyses, including two waveform analyses (i.e., a *mean amplitude analysis* and a temporally more fine-grained *time course analysis,* replicating the methods of the only previous EEG study investigating item-inherent semantic variables in picture naming; Rabovsky et al., 2021) and a *microstate analysis.*

In the mean amplitude analysis (posterior ROI, 200–550ms), we found that number of semantic features, a variable that, behaviourally, leads to faster and more accurate responses, was the only semantic variable predicting the ERP signal: Higher numbers of semantic features resulted in an enhanced posterior positivity. In the time course analysis, no semantic variable resulted in an effect that consistently survived correction for multiple testing. Therefore, we have no basis on which to discuss the precise time course of the semantic variables, given that the remaining analyses provide insufficient temporal precision to inform understanding of the fine-grained temporal development of the effects.

Moreover, four of the six semantic variables, number of semantic features, intercorrelational density, number of near semantic neighbours, and semantic similarity, significantly affected the number of timeframes associated with different periods of electrophysiological stability (topographic maps spanning processing between 50 and 550ms post picture onset) in the microstate analysis. Interestingly, all these variables affected the duration of one particular stable topographical map (Microstate 5; ~364–440ms), with number of near semantic neighbours additionally affecting Microstate 4 (~280–364ms). In contrast, we did not find electrophysiological correlates of typicality and distinctiveness in any of the analyses. The findings of this study compared to those of Rabovsky et al. (2021) are summarised in Table 4.

Table 4

Summary of EEG results for semantic variables

Semantic variable	Rabovsky et al. (2021)	Waveform analysis	Microstate(s) affected
Number of semantic features	↗	↗	5
Intercorrelational density	↗	Ø	5
Number of near semantic neighbours		Ø	4, 5
Semantic similarity		Ø	5
Typicality		Ø	Ø
Distinctiveness		Ø	Ø

Note. Ø = non-significant effect, ↗ = enhanced posterior positivity with higher values of the variable, blank cells were not investigated.

The fact that the variables studied here aim to measure aspects of the representation of meaning implies that they influence processing when semantic information is encoded (i.e., semantic and/or lexical processing). Thus, our finding of significant ERP effects of number of semantic features, intercorrelational density, semantic similarity, and number of near semantic neighbours suggests that they modulate the semantic and/or lexical networks involved in word production. For number of semantic features and intercorrelational density, this is in line with Rabovsky et al. (2021) and some of the findings by Clarke et al. (2013). However, this is the first time that item-inherent measures of number of near semantic neighbours and semantic similarity have been shown to modulate ERPs in word production, thus extending the evidence base regarding evoked responses to semantic variables to further variables, which had not been studied with ERPs or ERFs in the past.

The posterior positivity in the mean amplitude analysis was enhanced for words with a higher number of semantic features. This directly replicates the findings by Rabovsky et al. (2021) who reported an enhanced posterior positivity for words with higher numbers of semantic features in the mean amplitude analysis (200–550ms; across two presentation rounds) and in the more fine-grained analysis between around 330 and 600ms (first naming round; not corrected for multiple testing). Following Rabovsky et al. (2021), we interpret the enhanced posterior positivity as representing "activation in the lexical semantic system during competitive lexical selection" (p. 515). More semantic features result in stronger semantic activation of the target's semantic representation in a similar way to semantic priming (Rabovsky et al., 2021; Rabovsky & McRae, 2014, Simulation 2 with a neural network model of the word recognition processes). Depending on the semantic organisation assumed by a word production theory, this could be due to a spread of activation through the semantic system via bidirectionally interconnected semantic features (e.g., McRae et al., 1997; Rabovsky & McRae, 2014), via feedback from lexical to semantic representations (e.g., Dell, 1986), or spreading activation between holistic lexical concepts, where the measure 'number of semantic features' could be thought to represent the number of connections to other lexical concepts (e.g., Abdel Rahman & Melinger, 2009; Collins & Loftus, 1975; Levelt et al., 1999; see Lampe, Hameau, Fieder, et al., 2021, for in-depth discussion of the possible mechanisms). The strong semantic activation of a concept with many semantic features is thought to result in stronger activation of its corresponding lexical entry. Hence, the posterior positivity likely reflects this stronger activation of the semantic and lexical representations of target words with more semantic features. In contrast, number of semantic features (as well as proportion of visual features) was non-significant in the MEG study by Clarke et al. (2013). However, Clarke et al. focused on perceptual and semantic processing and only analysed the first 300ms of word planning and it is hence possible that our effect of number of semantic features was outside their analysis window.

Given that all those semantic variables to show an influence on ERPs significantly affected the number of timeframes associated with Microstate 5 (with additional effects of number of near semantic neighbours on Microstate 4), we may conclude that the neuronal network associated with Microstate 5 is engaged in processing of, or influenced by, semantic information. Hence, it likely captured the neuronal network associated with semantic and/or lexical processing. The functional basis of the observed effects of semantic variables on Microstate 5 is not entirely clear. However, it is possible that these variables might affect the strength of activity in the neuronal network associated with semantic and lexical processing during word planning, similar to the argument made for the posterior positivity by Rabovsky et al. (2021), which we adopted above. This network may be more active when processing some words compared to others, depending on the item-inherent values of the semantic variables[6]. Importantly, higher activity in the network could be caused by increased activation of the target representation itself or increased activation distributed across a number of co-activated competitors. For the semantic variables we found to be significant in the microstate analysis, this suggests that activity in the neuronal network associated with semantic and lexical processing may be higher when planning targets with higher numbers of semantic features (many semantic features strongly activate the target's lexical representation, as explained above), words with higher intercorrelational density, a higher number of near semantic neighbours, and words with higher

[6] We do not believe that interpreting the direction of change in the number of timeframes associated with a single microstate induced by the semantic variables (i.e., decrease or increase of a microstate's duration) is meaningful, nor that it is possible to relate that change to a change in processing time needed for naming.

semantic similarity. We have already addressed the proposed mechanism for number of semantic features and discuss the other variables below.

Even though we did not find an effect of intercorrelational density in the mean amplitude analysis, the finding that this variable influenced the activity in the neuronal network associated with semantic and lexical processing during word planning in the microstate analysis is in line with Rabovsky et al. (2021). Intercorrelational density likely measures the size of the co-activated lexical cohort: For words with higher intercorrelational density, mutual activation via intercorrelated features causes an increased number of lexical competitors to be co-activated, which compete for lexical selection with the target word (see also Lampe et al., in press; Rabovsky et al., 2016, 2021). More specifically, highly intercorrelated features strongly mutually co-activate, which increases the activation of these features (and the target lexical representation). However, these highly intercorrelated features (e.g., *has fur, has four legs, has paws, has whiskers*) also characterise groups of closely related concepts (e.g., 'cat', 'dog', 'tiger'), which are therefore also assumed to be strongly co-activated during processing. Hence, the effect of intercorrelational density on Microstate 5 might represent increased activation in the semantic and lexical network, which is mostly associated with the co-activated semantically related cohort. Similarly, in their MEG study, Clarke et al. (2013) found that around 224–260ms, an increased MEG response was associated with decreased values of their *correlational strength of shared features* component (a slightly different measure to our intercorrelational density measure, containing two measures capturing the correlational structure of shared features). This was suggested to be due to additional processing being required to activate and integrate semantic features of words with lower intercorrelational strength in the absence of mutual activation between intercorrelated features, which would benefit the integration of semantic information. Indeed, both Clarke et al.'s semantic and Rabovsky et al.'s predominantly lexical mechanisms might both be at play: Semantic processing may be facilitated by higher intercorrelational density (Clarke et al., 2013) while lexical processing may be inhibited due to increased competition (Rabovsky et al., 2021). Both mechanisms might cause increased activation in the semantic and lexical network for words with higher intercorrelational density, which is likely what we capture with the significant effect for Microstate 5.

Similar to intercorrelational density, number of near semantic neighbours represents the size of a semantically closely related co-activated lexical cohort. No previous ERP or ERF investigations of this semantic variable in word production have been conducted, as far as we are aware. Number of near semantic neighbours significantly affected the number of timeframes associated with Microstates 4 and 5. This indicates that number of semantic neighbours influenced the neuronal network associated with semantic and lexical processing, despite the absence of a significant behavioural effect of this variable (Lampe et al., in press). Based on previous behavioural findings (Fieder et al., 2019; Mirman, 2011), near semantic neighbours may be co-activated during processing via the many semantic features they share with the target. For words with higher numbers of near semantic neighbours, there might thus be increased activity from many co-activated competitors in the semantic and lexical network during processing.

Item-inherent semantic similarity has also not previously been investigated using electrophysiological measures (although see Rose & Abdel Rahman, 2017, for an ERP investigation of effects of between items semantic similarity on cumulative semantic interference). Our measure captures the semantic similarity of the target word and all other words in the mental lexicon. The interpretation of the effect of semantic similarity on Microstate 5 is complicated by the fact that previous behavioural analyses of the effects of semantic similarity are inconclusive: While the behavioural effects were not significant in the behavioural data associated with this study (Lampe et al., in press), Fieder et al. (2019) reported inhibitory effects for words with higher semantic similarity in unimpaired participants in a speeded picture naming task, which were suggested to be due to increased competition during lexical processing caused by co-activated representations. This would indicate stronger activity in the lexical and semantic network for words with higher semantic similarity to be distributed across the many lexical competitors (similar to words with higher intercorrelational density and a higher number of near semantic neighbours). In contrast, Lampe, Hameau, Fieder, et al. (2021) found a facilitatory effect of semantic similarity in people with aphasia: Participants with severe semantic impairments were more likely to respond correctly to words with higher semantic similarity and, correspondingly, more likely to make a semantic error on words with lower semantic similarity. This facilitation of words with higher semantic similarity was proposed to be due to increased spreading activation at the semantic level or due to enhanced feedback from lexical to semantic representations with converging activation on the target's semantic representation. The strongly activated semantic representation of a target with higher semantic similarity might then activate its lexical representation more strongly, making it a strong candidate for lexical selection (similar to the argument made for words with many semantic features above). In this case, stronger activity in the semantic and lexical network when processing words with higher semantic similarity would be related to the target word itself, rather than co-activated competitors. Even though the electrophysiological effect of semantic similarity does not allow us to adjudicate between the two mechanisms proposed to explain previous behavioural data, our finding supports the two behavioural studies in that semantic similarity seems to be an influential variable in word production, affecting the strength of activity in the semantic and lexical network. Further research into this variable is warranted to better understand its effect.

Neither distinctiveness nor typicality showed reliable effects in any of the ERP analyses. Distinctiveness had not previously been investigated using EEG. However, using MEG, Clarke et al. (2013) and Miozzo et al. (2015) found effects of principal components that captured the proportion of distinctive versus shared information related to a concept as well as other measures. However, direct comparison to our analyses is not possible, given that multiple measures were combined to form these "distinctiveness" principal components (including our measure of distinctiveness in Clarke et al. and a similar measure, number of distinctive features, in Miozzo et al.). Further reasons for the discrepancies between our and the previous findings may be the precision of the method employed (MEG vs EEG), the analysis conducted and, importantly, the other variables controlled for in the analyses (e.g., Clarke et al. only controlled for familiarity and image complexity). To our knowledge, there are no previous studies investigating effects of typicality on ERPs or ERFs in word production.

Taken together, the findings of the mean amplitude and microstate analyses indicated that several semantic variables (i.e., number of semantic features, intercorrelational density, semantic similarity, and number of near semantic neighbours) influence brain activity during word production. These variables likely affect the activity of the semantic and lexical network engaged in word production, in line with Rabovsky et al.'s (2021) proposal.

Previous ERP research has often used the word production time course estimates by Indefrey and Levelt (2004) and Indefrey (2011) to interpret the findings and to associate them with certain stages of word production. For example, ERP effects around 200ms have been taken to indicate that the experimental manipulation affects lexical processing (e.g., Aristei et al., 2011; Cheng et al., 2010; Dell'Acqua et al., 2010; Python et al., 2018a). Our findings suggest that semantic information must be required during processing at a stage (~364–440ms) that falls beyond the time window that is usually associated with semantic and lexical processing (i.e., up to around 275ms following Indefrey, 2011; but see Strijkers & Costa, 2011, for a call to assess the reliability of this upper limit of lexical selection). Hence, if we compare our data to this time course for word production, it suggests that semantic and lexical processing are likely not completed within the first 275ms of word planning. If interpreted as reflecting the time course of processing, our data would suggest that semantic information is activated longer or that effects of semantic processing are visible in a longer time-window than originally proposed by Indefrey and Levelt.

The effects of semantic variables might occur relatively late because semantic and lexical processing in this study may have been prolonged, also causing longer naming latencies compared to the time course estimates (i.e., 600ms vs around 900ms), with semantic and lexical processing possibly being associated with networks that are activate as late as 280–440ms (Microstates 4 and 5). Yet, the adaptation of Indefrey and Levelt's (2004) time course estimates to longer response latencies is far from straightforward and it is unknown which word production processes would have to be extended and by how much to accommodate for the difference in word planning duration compared to Indefrey and Levelt. Importantly, any attempt to rescale the time course estimates (e.g., Krott et al., 2019; Piai et al., 2012; Schuhmann et al., 2009; Shao et al., 2014) would be applied uniformly to all items, thus possibly levelling out the fine-grained effects of word-inherent variables that we are investigating in this study. Alternatively, effects of semantic variables might occur relatively late in word planning, if semantic information cascades to later stages of word production (e.g., Python et al., 2018a, 2018b). Ultimately, this study adds to the accumulating evidence (e.g., Miozzo et al., 2015, who found effects of phonological variables much earlier than the suggested time-window for phonological processing) suggesting that the reliability of Indefrey and Levelt (2004) and Indefrey's (2011) time course of word production might require further critical assessment (see also Strijkers & Costa, 2011).

As outlined in the Introduction, facilitatory and inhibitory effects of semantic variables in behavioural research have often been interpreted as reflecting either semantic facilitation or lexical competition (e.g., Abdel Rahman & Melinger, 2019; Fieder et al., 2019; Lampe et al., in press; Rabovsky et al., 2016). However, we found that the semantic variables, despite the differences in their behavioural effects in previous studies, influenced the number of timeframes associated with the same microstate, Microstate 5. One might be tempted to interpret this as suggesting an overlap, or interaction, between semantic and lexical processing (e.g., Abdel Rahman & Melinger, 2009; Dell, 1986; Indefrey, 2011; see Rabovsky et al., 2021, for a similar argument). However, our finding might also be compatible with a sequential architecture of semantic and lexical processing (e.g., Levelt et al., 1999): Words differ in their processing times, which causes different time courses of the underlying processing stages (i.e., jittered processing), possibly making a differentiation between semantic and lexical processing impossible, even if the underlying processing architecture was sequential. Consequently, our ERP data does not provide evidence that would allow association of effects of the semantic variables with either semantic or lexical processing or enable us to conclude whether semantic and lexical processing in word production are sequential, sequential and interactive, or parallel processes. To understand the effects of the semantic variables even better, future studies are needed to disentangle if and how semantic and lexical processing interact.

Conclusion

We have shown that the number of semantic features, intercorrelational density, number of near semantic neighbours, and semantic similarity of a target word influence ERPs during word production. This extends previous research into effects of semantic variables on ERPs or ERFs (Clarke et al., 2013; Miozzo et al., 2015; Rabovsky et al., 2021). Our findings suggest that these semantic variables influence the strength of activity in the semantic and lexical network, with increased activity being associated either with the target word itself or distributed across a co-activated lexical cohort.

The fact that semantic variables influenced processing as late as around 400ms may suggest that semantic information is important at stages later than the time windows traditionally associated with semantic and lexical processing. We are only just beginning to understand which variables influence word planning. However, the novel finding of influences of several semantic variables suggests that they should be studied in more detail and that theories of word production as well as future research should account for these variables.

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*(5), 713–734. https://doi.org/10.1080/01690960802597250

Abdel Rahman, R., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience*, *34*(9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Anders, R., Llorens, A., Dubarry, A.-S., Trébuchon, A., Liégeois-Chauvel, C., & Alario, F.-X. (2019). Cortical dynamics of semantic priming and interference during word production: An intracerebral study. *Journal of Cognitive Neuroscience*, *31*(7), 978–1001. https://doi.org/10.1162/jocn_a_01406

Aristei, S., Melinger, A., & Rahman, R. A. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience*, *23*(7), 1567–1586. https://doi.org/10.1162/jocn.2010.21474

Baayen, H. R., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28. https://doi.org/10.21500/20112084.807

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv*, 1–27. https://doi.org/arXiv:1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, *29*(4), 1165–1188. https://doi.org/10.1214/aos/1013699998

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 1–19. https://doi.org/10.3389/fninf.2015.00016

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Blackford, T., Holcomb, P. J., Grainger, J., & Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition*, *123*(1), 84–99. https://doi.org/10.1016/j.cognition.2011.12.007

Bormann, T. (2011). The role of lexical-semantic neighborhood in object naming: Implications for models of lexical access. *Frontiers in Psychology*, *2*, 1–11. https://doi.org/10.3389/fpsyg.2011.00127

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, *2011*, 1–15. https://doi.org/10.1155/2011/813870

Cheng, X., Schafer, G., & Akyürek, E. G. (2010). Name agreement in picture naming: An ERP study. *International Journal of Psychophysiology*, *76*(3), 130–141. https://doi.org/10.1016/j.ijpsycho.2010.03.003

Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2013). From perception to conception: How meaningful objects are processed over time. *Cerebral Cortex*, *23*(1), 187–197. https://doi.org/10.1093/cercor/bhs002

Cohen, X. M. (2014). *Analyzing neural time series data*. MIT Press. https://doi.org/10.1007/s13398-014-0173-7.2

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

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 of the United States of America*, *106*(50), 21442–21446. https://doi.org/10.1073/pnas.0908921106

Davis, C. J. (2005). N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, *32*(4), 588–615. https://doi.org/10.3758/BF03200832

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*, 1–10. https://doi.org/10.3389/fpsyg.2010.00177

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

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*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

den Hollander, J., Jonkers, R., Mariën, P., & Bastiaanse, R. (2019). Identifying the speech production stages in early and late adulthood by using electroencephalography. *Frontiers in Human Neuroscience*, *13*, 298. https://doi.org/10.3389/fnhum.2019.00298

Ewald, A., Aristei, S., Nolte, G., & Rahman, R. A. (2012). Brain oscillations and functional connectivity during overt language production. *Frontiers in Psychology*, *3*, 1–12. https://doi.org/10.3389/fpsyg.2012.00166

Fieder, N., Wartenburger, I., & Abdel Rahman, R. (2019). A close call: Interference from semantic neighbourhood density and similarity in language production. *Memory & Cognition*, *47*(1), 145–168. https://doi.org/10.3758/s13421-018-0856-y

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*(4), 870–878. https://doi.org/10.1006/nimg.2001.1037

Grossman, M., Robinson, K., Biassou, N., White-Devine, T., & D'Esposito, M. (1998). Semantic memory in Alzheimer's disease: Representativeness, ontologic category, and material. *Neuropsychology*, *12*(1), 34–42. https://doi.org/10.1037/0894-4105.12.1.34

Hameau, S., Nickels, L., & Biedermann, B. (2019). Effects of semantic neighbourhood density on spoken word production. *Quarterly Journal of Experimental Psychology*, *72*(12), 2752–2775. https://doi.org/10.1177/1747021819859850

Hirschfeld, G., Jansma, B., Bölte, J., & Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *NeuroReport*, *19*(12), 1227–1230. https://doi.org/10.1097/WNR.0b013e328309ecd1

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition*, *13*(7–8), 884–910. https://doi.org/10.1080/13506280544000093

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*(1), 67–104. https://doi.org/10.1080/02643298808252927

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 1–16. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Jasper, H. H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology*, *10*, 370–375. https://doi.org/10.1016/0013-4694(58)90053-1

Jia, H. (2019). Microstate analysis. In L. Hu & Z. Zhang (Eds.), *EEG signal processing and feature extraction* (pp. 141–157). https://doi.org/10.1007/978-981-13-9113-2

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*(2), 243–275. https://doi.org/10.1016/0010-0285(84)90009-4

Krott, A., Medaglia, M. T., & Porcaro, C. (2019). Early and late effects of semantic distractors on electroencephalographic responses during overt picture naming. *Frontiers in Psychology*, *10*, 1–

13. https://doi.org/10.3389/fpsyg.2019.00696

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Laganaro, M. (2014). ERP topographic analyses from concept to articulation in word production studies. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.00493

Laganaro, M., & Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography*, *24*(1), 19–29. https://doi.org/10.1007/s10548-010-0162-x

Laganaro, M., Valente, A., & Perret, C. (2012). Time course of word production in fast and slow speakers: A high density ERP topographic study. *NeuroImage*, *59*(4), 3881–3888. https://doi.org/10.1016/j.neuroimage.2011.10.082

Lampe, L. F., Fieder, N., Krajenbrink, T., & Nickels, L. (2017). Semantische Nachbarschaft in der Wortproduktion bei Aphasie [Semantic neighbourhood in word production in aphasia]. In A. Adelt, Ö. Yetim, C. Otto, & T. Fritzsche (Eds.), *Spektrum Patholinguistik - Panorama Patholinguistik: Sprachwissenschaft trifft Sprachtherapie* (Vol. 10, pp. 103–114). Universitätsverlag Potsdam. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/39701

Lampe, L. F., Hameau, S., Bürki, A., Fieder, N., Sowman, P. F., & Nickels, L. (2019). *Investigating effects of semantic variables on word production*. Preregistration on the Open Science Framework. osf.io/yw6ma

Lampe, L. F., Hameau, S., Fieder, N., & Nickels, L. (2021). Effects of semantic variables on word production in aphasia. *Manuscript Submitted for Publication*.

Lampe, L. F., Hameau, S., & Nickels, L. (2021). Are they really stronger? Comparing effects of semantic variables in speeded deadline and standard picture naming. *Manuscript Submitted for Publication*.

Lampe, L. F., Hameau, S., & Nickels, L. (in press). Semantic variables both help and hinder word production: Behavioural evidence from picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Lehmann, D., Ozaki, H., & Pal, I. (1987). EEG alpha map series: Brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and Clinical Neurophysiology*, *67*(3), 271–288. https://doi.org/10.1016/0013-4694(87)90025-3

Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, *10*(5), 553–567. https://doi.org/10.1162/089892998562960

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience*, *14*(3), 455–462. https://doi.org/10.1162/089892902317361967

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*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99–130. https://doi.org/10.1080/10430719008404646

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, *25*(10), 3343–3355. https://doi.org/10.1093/cercor/bhu137

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective and Behavioral Neuroscience*, *11*(1), 32–43. https://doi.org/10.3758/s13415-010-0009-7

Mirman, D., & Graziano, K. M. (2013). The neural basis of inhibitory effects of semantic and phonological neighbors in spoken word production. *Journal of Cognitive Neuroscience*, *25*(9), 1504–1516. https://doi.org/10.1162/jocn_a_00408

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 65–79. https://doi.org/10.1037/0278-7393.34.1.65

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705–714. https://doi.org/10.3758/BF03202720

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, *20*(4), 249–264. https://doi.org/10.1007/s10548-008-0054-5

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 943–acc4716-8. https://doi.org/10.1126/science.aac4716

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*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: Model estimation and validation. *IEEE Transactions on Biomedical Engineering*, *42*(7), 658–665. https://doi.org/10.1109/10.391164

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Piai, V., Roelofs, A., & van der Meij, R. (2012). Event-related potentials and oscillatory brain responses associated with semantic and Stroop-like interference effects in overt naming. *Brain Research*, *1450*, 87–101. https://doi.org/10.1016/j.brainres.2012.02.050

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Poulsen, A. T., Pedroni, A., Langer, N., & Hansen, L. K. (2018). Microstate EEGlab toolbox: An introductory guide. *BioRxiv*, 1–30. https://doi.org/10.1101/289850

Python, G., Fargier, R., & Laganaro, M. (2018a). ERP evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex*, *99*, 1–12. https://doi.org/10.1016/j.cortex.2017.09.008

Python, G., Fargier, R., & Laganaro, M. (2018b). When wine and apple both help the production of grapes: ERP evidence for post-lexical semantic facilitation in picture naming. *Frontiers in Human Neuroscience*, *12*, 1–13. https://doi.org/10.3389/fnhum.2018.00136

R Core Team. (2019). *R: A language and environment for statistical computing* (3.6.1). R Foundation for Statistical Computing. https://www.r-project.org/

Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, *132*(1), 68–89. https://doi.org/10.1016/j.cognition.2014.03.010

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, *146*, 240–244. https://doi.org/10.1016/j.cognition.2015.09.016

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2018). Object naming is facilitated by semantic richness but inhibited by semantic density: Electrophysiological and behavioral evidence. *BioRxiv*. https://doi.org/10.1101/509000

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2021). Semantic richness and density effects on language production: Electrophysiological and behavioral evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(3), 508–517. https://doi.org/10.1037/xlm0000940

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Rogers, T. T., Patterson, K., Jefferies, E., & Lambon Ralph, M. A. (2015). Disorders of representation and control in semantic cognition: Effects of familiarity, typicality, and specificity. *Neuropsychologia*, *76*, 220–239. https://doi.org/10.1016/j.neuropsychologia.2015.04.015

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rose, S. B., & Abdel Rahman, R. (2017). Semantic similarity promotes interference in the continuous naming paradigm: Behavioural and electrophysiological evidence. *Language, Cognition and Neuroscience*, *32*(1), 55–68. https://doi.org/10.1080/23273798.2016.1212081

Rose, S. B., Aristei, S., Melinger, A., & Abdel Rahman, R. (2019). The closer they are, the more they interfere: Semantic similarity of word distractors increases competition in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 753–763. https://doi.org/10.1037/xlm0000592

RStudio Team. (2020). *RStudio: Integrated development for R* (1.3.959). RStudio, PBC. http://www.rstudio.com

Schnur, T. T., Lee, E., Coslett, H. B., Schwartz, M. F., & Thompson-Schill, S. L. (2005). When lexical selection gets tough, the LIFG gets going: A lesion analysis study of interference during word production. Brain and Language, 95, 12–13. https://doi.org/10.1016/j.bandl.2005.07.008

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*(9), 1111–1116. https://doi.org/10.1016/j.cortex.2008.10.013

Shao, Z., Roelofs, A., Acheson, D. J., & Meyer, A. S. (2014). Electrophysiological evidence that inhibition supports lexical selection in picture naming. *Brain Research*, *1586*, 130–142. https://doi.org/10.1016/j.brainres.2014.07.009

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*, 1–16. https://doi.org/10.3389/fpsyg.2011.00356

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928. https://doi.org/10.1093/cercor/bhp153

Taylor, K. I., Devereux, B. J., Acres, K., Randall, B., & Tyler, L. K. (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. *Cognition*, *122*(3), 363–374. https://doi.org/10.1016/j.cognition.2011.11.001

Thompson-Schill, S. L., D'esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. Proceedings of the National Academy of Sciences of the United States of America, 94, 14792–14797.

Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience*, *8*, 1–13. https://doi.org/10.3389/fnins.2014.00390

Valente, A., & Laganaro, M. (2015). Ageing effects on word production processes: An ERP topographic analysis. *Language, Cognition and Neuroscience*, *30*(10), 1259–1272. https://doi.org/10.1080/23273798.2015.1059950

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Woollams, A. M. (2012). Apples are not the only fruit: The effects of concept typicality on semantic representation in the anterior temporal lobe. *Frontiers in Human Neuroscience*, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00085

Appendices

Appendix A: Exact replication and extension of Rabovsky et al. (2021) with relevant psycholinguistic control variables

Analogous to the behavioural analysis Lampe et al. (in press), three separate analyses of the EEG data were conducted: 1) Replication of the waveform EEG analysis performed by Rabovsky et al. (2021), 2) Replication of Rabovsky et al. taking a larger number of psycholinguistic control variables as well as ordinal category position into account, and 3) Extension of Rabovsky et al. to other semantic variables while also controlling for psycholinguistic control variables and ordinal category position. This approach was preregistered on the Open Science Framework (Lampe et al., 2019; https://osf.io/yw6ma/). The third analysis was reported in the main body of the text. Here we want to summarise the findings of the Analyses 1 and 2.

Analysis 1: Replication of Rabovsky et al. (2021) to investigate the temporal dynamics of two semantic variables using EEG

In this analysis, the number of semantic features and intercorrelational density were the semantic variables of interest. As in the original study we included the following control variables: familiarity (rated), number of orthographic neighbours (N-Watch database, Davis, 2005), lexical frequency (SUBTLEX-UK, van Heuven et al., 2014), and visual complexity (rated). Moreover, we used random intercepts for participants and items as well as random slopes for participants for both semantic variables, where appropriate (Bates, Kliegl, et al., 2015).

Analysis 2: Replication of Rabovsky et al. (2021) to investigate the temporal dynamics of two semantic variables using EEG, under sufficient control of psycholinguistic variables

Number of semantic features and intercorrelational density continued to be the only two semantic predictor variables in the model. In this analysis, we also took a wide range of psycholinguistic control variables (name agreement, imageability, age of acquisition, familiarity, frequency, a measure of ordinal position within a category within the list, and trial order) into account. Again, random intercepts for participants and items were included in the model as well as random slopes for the two semantic variables by participants, where supported (Bates, Kliegl, et al., 2015).

For the two extra analyses we each ran the mean amplitude analysis (mean ERP amplitude in posterior ROI between 200 and 550ms) as well as the time course analysis (consecutive 10ms segments between 0 and 550ms). These analyses are crucial when assessing the success to replicate Rabovsky et al. (2021) as they allow analysing any failure to replicate effects reported by Rabovsky et al. in the analysis reported in the paper: If effects are non-significant in the analysis that includes further semantic and/or control variables, but significant in the analyses replicating Rabovsky et al., we know that the difference in findings must be due to the inclusion of further psycholinguistic control or/or semantic variables.

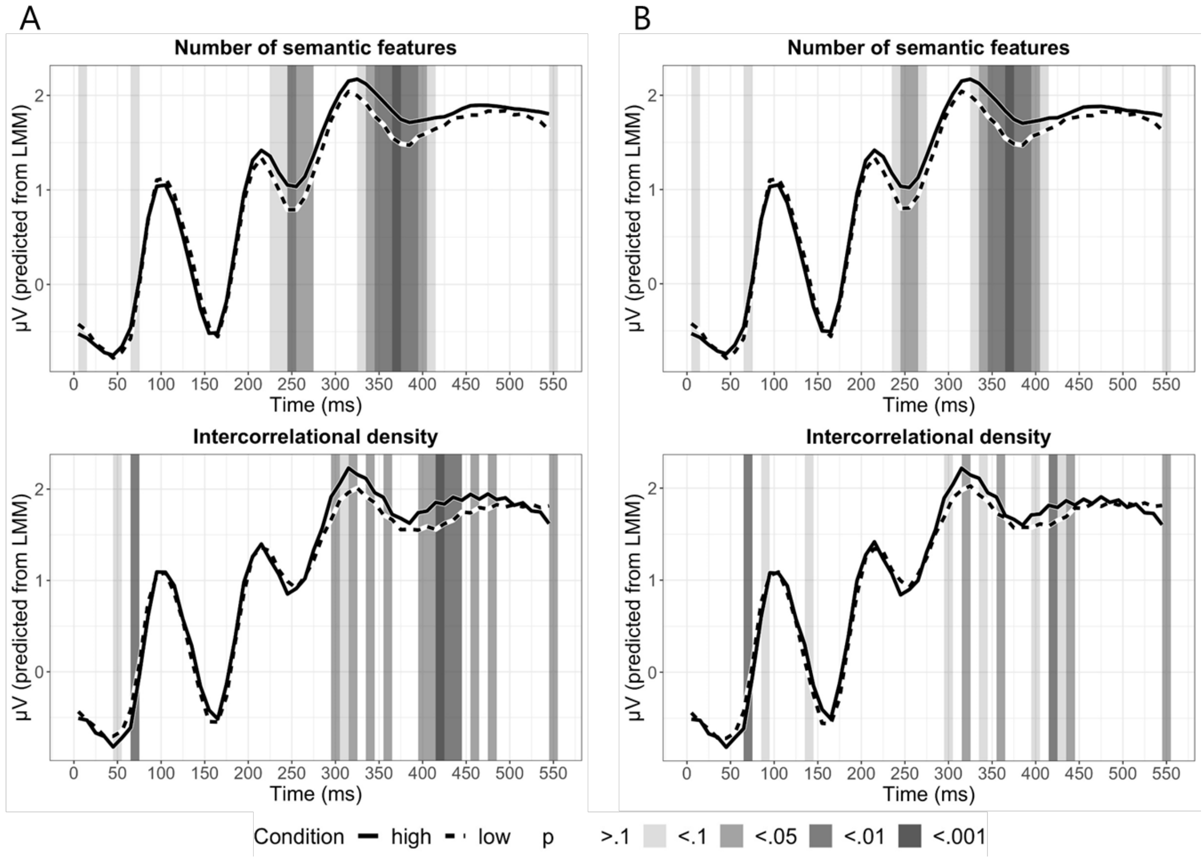
Table A1 summarises the outcome of the two additional analyses. In both analyses we replicated the enhanced posterior positivity for words with a higher number of semantic features. In contrast, the enhanced posterior positivity for words with higher intercorrelational density was non-significant.

Figure A1 plots the time course of the effects of the two semantic variables in 10ms time windows from picture onset to 550ms. In contrast to the analysis reported in the main body of the text, some effects survived correction for multiple comparisons. In the analysis exactly replicating the variables of the analysis by Rabovsky et al. (2021; Analysis 1, Panel A in Figure A1), the posterior positivity was enhanced for words with higher numbers of semantic features in the time-window between 360 and 390ms post picture onset (corrected for multiple testing with the Benjamini-Hochberg correction and adopting a q-level of 0.05, Benjamini & Hochberg, 1995; using the less conservative correction of an effect lasting at least 20ms with $p < 0.1$, a slightly larger time-window between 340 and 390ms survived correction, Laganaro & Perret, 2011). Similarly, the stronger posterior positivity for words with higher intercorrelational density was significant between 410 and 420ms and 430 and 440ms post picture onset after applying the Benjamini-Hochberg false discovery correction (410–440ms using $p < 0.1$).

In the analysis extending Rabovsky et al. (2021) with more control variables (Analysis 2, Panel B in Figure A1), the effect of number of semantic features survived correction in exactly the same time window as in Analysis 1. However, the effect of intercorrelational density did not survive our attempts to correct for multiple testing.

Figure A1

Linear mixed model estimates of the mean voltages ± the effect sizes of each semantic variable at the posterior region of interest in consecutive 10ms segments between 0 and 550ms of the models replicating Rabovsky et al. (2021) (A) and extending their analysis with further psycholinguistic control variables (B)



Note. Grey shading indicates levels of significance. High and low conditions of each semantic variable are based on the mean voltage (model Intercept) ± the effect size of the respective semantic variable.

Table A1

Summarised outputs of linear mixed model analysis replicating and extending Rabovsky et al. (2021) with further control variables

	Replicating Rabovsky et al. (2021)						Extending Rabovsky et al. (2021) with control variables					
Random effect	Variance	*SD*					Variance	*SD*				
Item (Intercept)	0.08	0.29					0.06	0.25				
Subject (Intercept)	0.53	0.73					0.53	0.73				
Residuals	4.55	2.13					4.55	2.13				
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF
(Intercept)	1.65	0.09	1.48 – 1.82	19.28	<.001		1.65	0.09	1.48 – 1.81	19.30	<.001	
Name agreement							0.01	0.02	-0.03 – 0.06	0.55	.583	1.10
Image agreement							0.08	0.02	0.04 – 0.13	**3.49**	**<.001**	1.18
Imageability							-0.02	0.03	-0.08 – 0.04	-0.61	.543	1.82
Age of acquisition							-0.08	0.03	-0.15 – -0.02	**-2.50**	**.012**	2.27
Familiarity	-0.05	0.03	-0.10 – 0.01	-1.72	.086	1.38	-0.05	0.03	-0.10 – 0.00	-1.92	.055	1.51
Frequency	0.05	0.03	-0.01 – 0.11	1.76	.078	1.47	0.05	0.03	-0.01 – 0.10	1.71	.088	1.53
Orthographic neighbourhood	0.00	0.02	-0.05 – 0.05	-0.03	.978	1.18						
Visual complexity	0.01	0.02	-0.04 – 0.06	0.52	.605	1.15						
Ordinal category position							0.10	0.03	0.04 – 0.15	**3.36**	**.001**	2.52
Trial order							-0.11	0.03	-0.16 – -0.05	**-4.01**	**<.001**	2.34
Number of semantic features	0.07	0.03	0.02 – 0.12	**2.80**	**.005**	0.07	0.06	0.02	0.01 – 0.11	**2.38**	**.017**	1.30
Intercorrelational density	0.04	0.03	-0.01 – 0.09	1.75	.080	0.04	0.02	0.02	-0.03 – 0.07	0.88	.379	1.24
Observations	18424											
Marginal R^2 / Conditional R^2	0.006 / 0.120											

Note. VIF = Variance Inflation Factor.

Values of significant effects (p < .05) are printed in bold.

One possible reason for the discrepancies between the findings of our study and Rabovsky et al. (2021) is that they accounted for fewer psycholinguistic control variables in the analyses, which may have distorted their findings (see Lampe et al., in press, for an in-depth discussion). In the two additional analyses reported here (1) exact replication of the model structure used by Rabovsky et al. and 2) a model that improved the control of psycholinguistic variables), we tested this possibility and further examined the reliability of the ERP findings by Rabovsky et al..

When exactly replicating Rabovsky et al.'s (2021) analysis of the EPR data (Table A1, left-hand part, Figure A1 Panel A), we got closer to fully replicating their findings: The effect of number of semantic features was significant with an enhanced posterior positivity for words with higher numbers of features in the mean amplitude analysis and between around 260 and 390ms in the more precise time course analysis. Moreover, the effect of intercorrelational density with an enhanced posterior positivity for words with higher intercorrelational density was significant in the time course analysis around 410 and 440ms. However, when improving control of nuisance variables (Table A1, right-hand part, Figure A1 Panel B), the effect of intercorrelational density did no longer survive correction for multiple testing, while the effect of number of semantic features remained significant (see also Table A2 for a summary).

Hence, our findings resemble Rabovsky et al. (2021) when we exactly replicate the composition of their fixed effects. This suggests that their finding of an effect of intercorrelational density might have been a Type 1 error caused by insufficient control of variables that influence processes during word planning and emphasises, once again, the need for experiments to control for influential variables in the analyses to derive as pure measures of the effects of interest as possible. Importantly though, Rabovsky et al. did not report on the significance of the effect of the semantic variables in the single naming rounds. From visual inspection of Supplementary Figure S2 in Rabovsky et al., the effect of intercorrelational density might have gotten stronger in the repetition compared to the first naming round (but note that the interaction between ‘repetition’ and ‘intercorrelational density’ was non-significant). In contrast, naming was significantly slower for concepts with high intercorrelational density in the repetition compared to the first naming round. Further research is needed to investigate whether repeated naming of the same items may enhance effects of intercorrelational density.

Table A2

Summarised findings for semantic variables of all analyses

Semantic variable	Rabovsky et al. (2021)	Exact replication	Extension with control variables	Extension with control and semantic variables	Microstate affected
Number of semantic features	↗ 330–360ms [a]	↗ 360–390ms	↗ 360–390ms	↗	5
Intercorrelational density	↗ 335–450ms [a] (sporadically)	Ø 420–440ms	Ø	Ø	5
Number of near semantic neighbours				Ø	4, 5
Semantic similarity				Ø	5
Typicality				Ø	Ø
Distinctiveness				Ø	Ø

Note. Ø = non-significant effect, ↗ = enhanced posterior positivity with higher values of the variable, blank cells were not investigated.

[a] From visual inspection of the time course of effects in the first naming round (Supplementary Figure S2, Rabovsky et al., 2021). Not corrected for multiple testing.

Appendix B: Correlation analysis

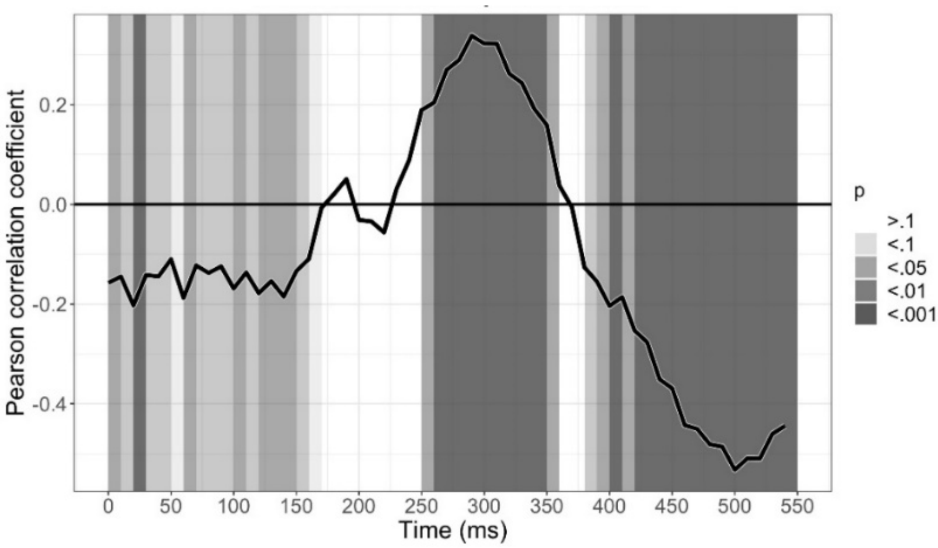
We calculated point-by-point correlations of every 10ms time window and naming latency to test whether the ERP modulations in the posterior ROI were associated with naming latency (e.g., Costa et al., 2009; Rose & Abdel Rahman, 2017). Following Costa et al. we would expect to find a positive correlation between them if both response latency and electrophysiological effects were caused by the same underlying processes in the brain and are related to lexical selection, which was the case in previous investigations (e.g., Costa et al., 2009; Dell'Acqua et al., 2010; Rose & Abdel Rahman, 2017; Strijkers et al., 2010).

Figure B1 shows the time course of the correlation between response latency and the ERP amplitudes at the posterior ROI. Response latency was negatively correlated with EPRs between picture onset and 150ms as well as after 380ms post picture onset, with the earlier time windows being less strongly correlated. In contrast, increased naming latency was associated with more positive amplitudes in the posterior ROI (positive correlation) between 250 and 360ms post picture onset. This may be interpreted as suggesting a duration of lexical processing of about 110ms (compared to about 75ms in Indefrey, 2011, 180ms in Costa et al., 2009, and 145ms in Rose & Abdel Rahman, 2017). This again shows that the duration of processing stages in word production may take longer than suggested by Indefrey (2011) and Indefrey and Levelt (2004) and may vary between studies.

However, these correlations between response latency and ERP amplitude are raw correlations, which do not take any effect of influential variables, be it control or semantic variables, into account. Yet, we believe that these fine-grained influences can alter and shift the durations of all stages of word production, making this coarse correlation a rather unreliable measure.

Figure B1

Correlation between response latency and ERP amplitudes at the posterior ROI



Appendix C: Complete statistical models of microstate analyses

Table C1

Complete statistical models of microstate analyses 1–2

	Microstate 1					Microstate 1				
Random effect	Variance	*SD*				Variance	*SD*			
Subject (intercept)	0.80	0.89				0.44	0.66			
Item (intercept)	0.11	0.34				0.04	0.20			
Residual	6.63	2.57				2.53	1.59			
Predictors	Estimate	*SE*	*CI*	*p*-value	VIF	Estimate	*SE*	*CI*	*p*-value	VIF
Intercept	8.70	0.10	8.49 – 8.90	<.001		3.75	0.08	3.60 – 3.90	<.001	
NameAgr	0.02	0.03	-0.04 – 0.08	.502	1.10	-0.04	0.02	-0.07 – -0.00	.041	1.10
AoA	0.01	0.04	-0.08 – 0.09	.870	2.39	0.01	0.03	-0.04 – 0.07	.588	2.40
Imageability	0.03	0.04	-0.05 – 0.10	.437	1.86	0.00	0.02	-0.04 – 0.05	.838	1.87
ImageAgr	-0.04	0.03	-0.11 – 0.02	.163	1.27	0.00	0.02	-0.04 – 0.04	.980	1.28
Frequency	-0.09	0.04	-0.16 – -0.02	.011	1.57	0.03	0.02	-0.01 – 0.08	.118	1.58
Familiarity	-0.02	0.04	-0.09 – 0.05	.643	1.74	-0.05	0.02	-0.09 – -0.00	.035	1.74
CatPos	-0.05	0.04	-0.12 – 0.02	.190	2.87	0.07	0.02	0.02 – 0.12	**.005**	2.84
Order	0.10	0.03	0.03 – 0.16	**.004**	2.50	-0.10	0.02	-0.14 – -0.06	**<.001**	2.44
NoFeats	-0.09	0.03	-0.15 – -0.02	.013	1.60	-0.03	0.02	-0.07 – 0.02	.212	1.61
IntercorrDens	0.04	0.04	-0.04 – 0.12	.330	2.21	0.01	0.03	-0.04 – 0.06	.624	2.21
SemSim	0.08	0.04	-0.00 – 0.17	.055	2.52	-0.02	0.03	-0.07 – 0.04	.549	2.52
NearSemNeigh	-0.10	0.04	-0.18 – -0.01	.027	2.47	0.01	0.03	-0.05 – 0.06	.848	2.47
Typicality	0.02	0.03	-0.05 – 0.09	.601	1.55	0.01	0.02	-0.03 – 0.05	.636	1.56
Distinctiveness	0.06	0.04	-0.02 – 0.14	.143	2.09	0.01	0.02	-0.04 – 0.06	.783	2.10
Observations	18,140					15,692				

Marginal R^2 / Conditional R^2	0.003 / 0.123	0.003 / 0.160

Note. VIF = Variance Inflation Factor; NameAgr = name agreement; ImageAgr = image agreement; AoA = age of acquisition; OrthNeigh = orthographic neighbourhood density; VisCom = visual complexity; OrdCatPos = ordinal category position; SemSim = Semantic similarity; NrSemNeigh = Number of near semantic neighbours; NoFeats = number of semantic features; IntercorrDensity = intercorrelational density; Distinct = distinctiveness.

Significant effects (Bonferroni corrected for multiple comparisons, significance threshold of $p < .008$) are printed in bold.

Table C2

Complete statistical models of microstate analyses 3–4

	Microstate 3					Microstate 4				
Random effect	Variance	*SD*				Variance	*SD*			
Subject (intercept)	0.54	0.73				0.56	0.75			
Subject | IntercorrDens						0.01	0.09			
Item (intercept)	0.08	0.28				0.07	0.27			
Residual	5.63	2.37				3.85	1.96			
Predictors	Estimate	*SE*	*CI*	*p*-value	VIF	Estimate	*SE*	*CI*	*p*-value	VIF
Intercept	6.99	0.09	6.82 – 7.16	<.001		5.20	0.09	5.03 – 5.37	<.001	
NameAgr	-0.06	0.03	-0.11 – -0.01	.020	1.10	0.04	0.02	-0.01 – 0.08	.112	1.10
AoA	0.10	0.04	0.03 – 0.17	.008	2.39	-0.01	0.03	-0.08 – 0.05	.699	2.37
Imageability	0.01	0.03	-0.05 – 0.08	.676	1.86	-0.01	0.03	-0.07 – 0.05	.793	1.85
ImageAgr	-0.07	0.03	-0.13 – -0.02	.009	1.27	0.00	0.03	-0.05 – 0.05	.874	1.27
Frequency	0.02	0.03	-0.04 – 0.08	.590	1.58	0.05	0.03	-0.01 – 0.10	.085	1.57
Familiarity	-0.05	0.03	-0.11 – 0.02	.139	1.74	-0.08	0.03	-0.14 – -0.03	**.004**	1.74

CatPos	-0.02	0.03	-0.09 – 0.04	.511	2.84	0.06	0.03	0.00 – 0.12	.035	2.87
Order	0.01	0.03	-0.05 – 0.07	.707	2.45	-0.04	0.03	-0.09 – 0.01	.113	2.51
NoFeats	-0.02	0.03	-0.08 – 0.04	.584	1.61	0.05	0.03	-0.00 – 0.10	.075	1.56
IntercorrDens	0.04	0.04	-0.03 – 0.11	.259	2.21	0.01	0.03	-0.05 – 0.08	.668	2.00
SemSim	0.03	0.04	-0.04 – 0.11	.394	2.52	-0.04	0.03	-0.11 – 0.03	.244	2.52
NearSemNeigh	-0.06	0.04	-0.14 – 0.01	.109	2.46	0.13	0.03	0.06 – 0.20	**<.001**	2.40
Typicality	0.03	0.03	-0.03 – 0.09	.356	1.56	-0.01	0.03	-0.06 – 0.04	.678	1.50
Distinctiveness	0.05	0.03	-0.02 – 0.12	.136	2.09	0.02	0.03	-0.04 – 0.09	.438	2.03
Observations	17,829					17,157				
Marginal R^2 / Conditional R^2	0.004 / 0.102					0.005 / 0.147				

Note. VIF = Variance Inflation Factor; NameAgr = name agreement; ImageAgr = image agreement; AoA = age of acquisition; OrthNeigh = orthographic neighbourhood density; VisCom = visual complexity; OrdCatPos = ordinal category position; SemSim = Semantic similarity; NrSemNeigh = Number of near semantic neighbours; NoFeats = number of semantic features; IntercorrDensity = intercorrelational density; Distinct = distinctiveness.

Significant effects (Bonferroni corrected for multiple comparisons, significance threshold of $p < .008$) are printed in bold.

Table C3

Complete statistical models of microstate analyses 5–6

	Microstate 5		Microstate 6	
Random effect	Variance	*SD*	Variance	*SD*
Subject (intercept)	0.27	0.52	0.50	0.70
Item (intercept)	0.02	0.14	0.05	0.23
Residual	2.53	1.59	6.30	2.51

Predictors	Estimate	*SE*	*CI*	*p*-value	VIF	Estimate	*SE*	*CI*	*p*-value	VIF
Intercept	3.57	0.06	3.45 – 3.69	<.001		7.25	0.08	7.08 – 7.41	<.001	
NameAgr	-0.02	0.02	-0.05 – 0.02	.317	1.10	0.03	0.03	-0.02 – 0.08	.187	1.10
AoA	-0.03	0.02	-0.07 – 0.02	.286	2.37	-0.11	0.04	-0.18 – -0.04	**.002**	2.37
Imageability	-0.03	0.02	-0.07 – 0.02	.236	1.85	-0.01	0.03	-0.07 – 0.06	.811	1.85
ImageAgr	0.01	0.02	-0.02 – 0.05	.400	1.27	0.12	0.03	0.07 – 0.17	**<.001**	1.27
Frequency	0.03	0.02	-0.01 – 0.07	.102	1.58	-0.01	0.03	-0.07 – 0.05	.762	1.57
Familiarity	0.03	0.02	-0.01 – 0.07	.132	1.73	0.11	0.03	0.05 – 0.17	**<.001**	1.73
CatPos	0.03	0.02	-0.02 – 0.07	.243	2.80	0.01	0.03	-0.06 – 0.08	.761	2.81
Order	-0.10	0.02	-0.14 – -0.06	**<.001**	2.36	0.00	0.03	-0.06 – 0.06	.963	2.38
NoFeats	0.06	0.02	0.03 – 0.10	**.001**	1.61	0.05	0.03	-0.01 – 0.10	.111	1.61
IntercorrDens	-0.07	0.02	-0.12 – -0.03	**.001**	2.20	-0.05	0.03	-0.12 – 0.01	.121	2.20
SemSim	-0.07	0.02	-0.12 – -0.03	**.002**	2.52	-0.04	0.04	-0.12 – 0.03	.225	2.51
NearSemNeigh	0.07	0.02	0.02 – 0.11	**.006**	2.46	0.01	0.04	-0.06 – 0.08	.761	2.45
Typicality	0.00	0.02	-0.04 – 0.03	.875	1.57	0.00	0.03	-0.06 – 0.06	.981	1.57
Distinctiveness	-0.05	0.02	-0.09 – -0.01	.027	2.10	-0.08	0.03	-0.15 – -0.02	.015	2.09
Observations	15,847					17,790				
Marginal R^2 / Conditional R^2	0.005 / 0.107					0.007 / 0.087				

Note. VIF = Variance Inflation Factor; NameAgr = name agreement; ImageAgr = image agreement; AoA = age of acquisition; OrthNeigh = orthographic neighbourhood density; VisCom = visual complexity; OrdCatPos = ordinal category position; SemSim = Semantic similarity; NrSemNeigh = Number of near semantic neighbours; NoFeats = number of semantic features; IntercorrDensity = intercorrelational density; Distinct = distinctiveness.

Significant effects (Bonferroni corrected for multiple comparisons, significance threshold of $p < .008$) are printed in bold.

CHAPTER 5

Are they really stronger? Comparing effects of semantic variables in speeded deadline and standard picture naming

Leonie F. Lampe
Solène Hameau
Lyndsey Nickels

Submitted for publication

Abstract

Investigations of effects of semantic variables on picture naming have often been inconclusive, with some studies reporting significant and others non-significant effects. One potential explanation may relate to the specific naming tasks used: While most previous studies have used standard picture naming, others have used speeded naming that requires participants to prioritise naming speed over accuracy. Speeded naming has been suggested to cause enhanced effects of item-inherent word characteristics due to disruptions of cognitive control and resulting modulations of responsiveness to input. Consequently, this study investigated whether effects are stronger in speeded compared to standard picture naming, focusing on six feature-based semantic variables: number of semantic variables, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness. The results showed few differences in the variables' effects between the two naming tasks: In the naming latency analysis, the inhibitory effect of distinctiveness was stronger in the speeded naming task, while in the accuracy analysis the effect of number of semantic features was stronger in the standard naming task. These findings cannot, therefore, be accounted for by increased responsiveness to input in speeded naming and we discuss possible underlying mechanisms. We conclude that, while some differences in effects of semantic variables between previous studies may have been caused by the specific naming task used, differences between studies more likely depend on statistical power and control of other influential variables in the experiment.

Introduction

Semantic variables capture aspects of the semantic representation of words and, in spoken word production, they can influence the activation environment of semantic and lexical processing. This influence enables researchers to use semantic variables to study the processes of word production, which has led to increasing interest in their effects in recent years (e.g., Bormann, 2011; Fieder et al., 2019; Hameau et al., 2019; Lampe, Hameau, & Nickels, in press; Mirman, 2011; Rabovsky et al., 2016; Taylor et al., 2012).

Importantly, research examining effects of the same semantic variable in picture naming has not always resulted in converging evidence, even within neurotypical participants. For example, increased semantic similarity, the average featural overlap between a target word and other words in the mental lexicon, has been reported to have both an inhibitory (Fieder et al., 2019) and no significant (Lampe, Hameau, & Nickels, in press) effect. This same pattern has been found for number of near semantic neighbours (words that share semantic information with the target; inhibitory effects: Fieder et al., 2019; Mirman, 2011; no significant effect: Bormann, 2011; Hameau et al., 2019; Lampe et al., 2017; Lampe, Hameau, & Nickels, in press). One systematic difference between the studies that yielded contrasting results is the specific experimental paradigm used. While all studies used standard picture naming tasks (i.e., picture naming of items in the absence of manipulation of the surrounding context or distractors, unlike the Blocked Cyclic Naming or Picture-Word Interference paradigms), those that yielded significant findings (Fieder et al., 2019; Mirman, 2011) used a *speeded picture naming task,* in which word planning is interfered with by enforcing a strict temporal cut-off to response initiation.

Speeded picture naming requires participants to name pictures at an increased rate that is faster than they would normally respond. For example, there may be a deadline of 600ms post picture onset, limiting the time for processing given that non-speeded, standard picture naming latencies are often around 900ms on average (e.g., Lampe, Hameau, & Nickels, in press). This leads to a decrease in naming latencies, which is often accompanied by an increase in naming errors compared to standard picture naming (i.e., speed-accuracy trade-off; e.g., Damian & Dumay, 2007; Kello & Plaut, 2000 in word reading). Naming errors observed in speeded naming are mostly semantically and visually related to the target word and resemble the naming errors made by participants with aphasia (e.g., Hodgson & Lambon Ralph, 2008; Mirman, 2011; Moses et al., 2004).

However, there is debate regarding how exactly the process of word planning is altered when participants are asked to name pictures quickly and disagreement about which processes or level(s) of word production might be affected. Vitkovitch and Humphreys (1991) found that naming errors in speeded participants were mostly semantically related to the target word and more frequently made on low compared to high frequency words, which they interpreted as an indicator of lexical processing being limited by the external time pressure (see Starreveld & La Heij, 1999, for similar findings and reasoning in a speeded Picture-Word Interference task, but note that effects of word frequency also arise in participants in standard naming, e.g., Alario et al., 2004; Lampe, Hameau, & Nickels, in press; see Perret & Bonin, 2019, for a metaanalysis and review). Similarly, Moses et al. (2004) reported predominantly semantically related perseverative and non-perseverative errors in picture naming, which were interpreted as reflecting processing difficulties at the lexical semantic level that were evoked by the speeding of naming. Lloyd-Jones and Nettlemill (2007) and Vitkovitch et al. (1993) also interpreted the different types of errors made in speeded naming (i.e., visual, semantic, and visual-semantic) as reflecting the processes being disrupted by the naming speed: visual and/or semantic processing.

In contrast, Kello and colleagues associated the behaviour observed in speeded naming with changes to cognitive control. They used two different versions of speeded naming: speeded deadline (Kello et al., 2000) and tempo naming (Kello, 2004; Kello & Plaut, 2000, 2003) in reading aloud and Stroop colour naming. In speeded deadline naming participants are instructed to name the pictures before a certain response deadline (e.g., 500ms after picture onset; see also e.g., Damian & Dumay, 2007; Kello et al., 2000; Lloyd-Jones & Nettlemill, 2007; Moses et al., 2004; Vitkovitch et al., 1993; Vitkovitch & Humphreys, 1991). In contrast, in tempo naming, participants are asked to respond at a certain point in time after the onset of the picture on screen (see also e.g., Fieder et al., 2019; Mirman, 2011; Mirman, Kittredge, et al., 2010). Participants hear a series of three beeps, each, for example, 500ms apart, the appearance of the picture coincides with a fourth beep, and participants are instructed to keep the rhythm of the beeps by naming the picture in time with beep number five. Importantly, Kello and colleagues suggest that strategic control allowed for the compression of the processing trajectory for naming in *both* tasks. More specifically, they proposed that changes in cognitive control dynamics evoked by the task can modify the processing parameter *input gain* (e.g., Gotts & Plaut, 2002; Kello, 2004; Kello et al., 2000, 2005; Kello & Plaut, 2000, 2003). Input gain changes the systems' sensitivity to new inputs and thus affects the processing units' responsiveness to their input: When input gain is low, activation of a processing unit is relatively unaffected by any inputs, however, when input gain is high, inputs have much stronger effect on activation. A rise in input gain consequently accelerates processing as the processing unit's activation reaches threshold much faster due to its amplified sensitivity to input. However, at high gain, processing is also less controlled, which results in decreased model precision and provokes processing errors (i.e., speed-accuracy trade-off). Therefore, this input gain proposal entails an alteration of the processing dynamics, such that the task requirements can be met. Alternatively, some authors have proposed a change in selection thresholds in speeded versus standard naming (e.g., Humphreys et al., 1995; Kello, 2004): In order to meet the task requirements, participants lower the threshold of activation required for information selection, which allows them to respond faster, however, at the expense of increased error rates due to incomplete processing (see also Coltheart et al., 2001).

Hodgson and Lambon Ralph (2008) directly compared the speeded deadline and tempo versions of speeded naming and proposed that both reduce the time for controlled semantic processing. Although both tasks resulted in many naming errors, there were significantly more errors in the tempo naming task (20% versus 18% in speeded deadline task). Hodgson and Lambon Ralph suggested that the dual-task requirement of the tempo naming task (i.e., naming pictures quickly while keeping the required tempo) resulted in a diversion of attention and executive resources from speech production and ultimately insufficient executive resources to be assigned to semantic cognition for successful picture naming.

Mirman (2011) also suggested a modulation of cognitive control as an explanation for his findings of significant effects of semantic neighbours on accuracy and the proportion of semantic errors in speeded naming. However, while he also used a tempo naming task (500ms tempo), he did not attribute his findings to dual-task requirements (cf. Hodgson & Lambon Ralph, 2008), but argued more generally for disrupted cognitive control of semantic processing resulting from the forced increase of processing speed in this task. In addition to effects of near semantic neighbours, Mirman also studied effects of distant semantic neighbours (i.e., words sharing only little semantic information with the target word, 0.25 < features vector cosine similarity > 0). While many *near* semantic neighbours had inhibitory effects on naming accuracy and led to an increase of semantic errors, many *distant* semantic neighbours facilitated processing and resulted in more accurate responses with marginally significantly fewer semantic errors[1]. Mirman explained these findings in the context of an attractor model of semantic cognition, in which attractors represent stable states of the model, corresponding to the target's pattern of activation across all semantic features. He argued that while near semantic neighbours interfere with the system successfully settling into the target attractor, many distant semantic neighbours pull towards the target and thus facilitate settling. Mirman argued that any effects of these semantic variables on the attractor landscape are magnified in speeded naming (and in participants with aphasia) due to disruptions of cognitive control mechanisms. He stated that disruption of control related to competition between co-activated lexical alternatives, could not explain the facilitatory effects of number of distant semantic neighbours and that disrupted competitive lexical selection in speeded naming (or aphasia) was consequently unable to explain the opposite effects of near and distant semantic neighbours[2].

---

[1] Please note that, as we have argued elsewhere (Hameau et al., 2019; Lampe, Hameau, Fieder, et al., in press), the effect of number of distant semantic neighbours is far from established in the literature. While Mirman (2011) reported effects in participants with and without aphasia, other work (Fieder et al., 2019; Hameau et al., 2019; Lampe, Hameau, Fieder, et al., in press), partly using the same database of picture naming data of participants with aphasia as Mirman (i.e., MAPP Database, Mirman, Strauss, et al., 2010), failed to replicate the significant effects of number of distant semantic neighbours. However, a replication of this finding would be particularly important as Mirman's item sets were subject to a number of shortcomings (i.e., small item sets that were insufficiently matched for some psycholinguistic variables with a non-unambiguous allocation of items into the sets of high vs low numbers of near/distant semantic neighbours). Given this lack of reliability of any effect of number of distant semantic neighbours on processing, we did not include it as a semantic variable of interest in the analyses conducted here.

[2] Importantly, we (e.g., Lampe, Hameau, & Nickels, in press) would argue that facilitatory and inhibitory effects of semantic variables could indeed be explained by spreading activation dynamics at the semantic level in the context of competitive lexical selection in word production (see also e.g., Abdel Rahman & Melinger, 2019; Rabovsky et al., 2016). This interpretation may be more plausible compared to the cognitive control mechanism

Instead, Mirman (2011) claimed that, under time pressure, the cognitive control mechanism increases the rate of processing such that participants can respond faster, which, however, comes at the cost of precision, leading to increased naming errors. More specifically, following the proposal of Kello and colleagues outlined above (e.g., Kello, 2004; Kello et al., 2000, 2005; Kello & Plaut, 2000, 2003), Mirman suggested that this change of cognitive control dynamics influenced input gain. He specified that the input gain parameter affects the processing units' sensitivity to their input, be it excitatory or inhibitory in nature. Consequently, Mirman proposed that facilitatory and inhibitory effects of distant and near semantic neighbours would be amplified following modulations of the processing units' responsiveness to input, such as in speeded naming.

However, this has never been tested empirically with a direct comparison of effects of semantic variables on performance in standard and speeded picture naming. Indeed, to date, only a limited number of studies (Fieder et al., 2019; Lloyd-Jones & Nettlemill, 2007; Mirman, 2011; Mirman, Kittredge, et al., 2010; Vitkovitch et al., 1993; Vitkovitch & Humphreys, 1991) have investigated effects of any item-inherent variables on performance in either variant of the speeded naming task, and even fewer have compared effects of these variables between speeded and standard naming. In fact, very few studies have directly compared performance in the two tasks at all. In picture naming, as far as we are aware, only Lloyd-Jones and Nettlemill (2007) have looked at effects of psycholinguistic variables (i.e., complexity, decomposability, contour overlap, imageability, age of acquisition, frequency, animacy) in speeded deadline and standard naming. Importantly however, they examined effects on different dependent variables in the two tasks (i.e., naming latencies in standard naming and naming accuracy and error types in speeded deadline naming) and effect sizes were not directly compared between the tasks. To our knowledge, the only direct comparisons of effects of psycholinguistic variables in standard and speeded processing have been conducted in reading aloud: Gerhand and Barry (1999) found stronger effects of age of acquisition on latency in a speeded deadline compared to

---

based interpretation provided by Mirman (2011), as facilitatory and inhibitory effects of semantic variables have also been observed in standard picture naming in neurotypical participants (e.g., Lampe, Hameau, Fieder, et al., in press; Rabovsky et al., 2016), thus in a context where responses were given at a normal pace, rendering involvement of a control mechanism to increase the naming rate unnecessary.

a standard word reading task, while the effect of word frequency was similar in the two tasks. In contrast, Kello and Plaut (2000) reported *attenuated* effects of word frequency and spelling-sound consistency on naming latency in tempo compared to standard reading. However, it is debateable how valid a latency measure is in the tempo task, given that in this task, in contrast to the speeded deadline task, participants aim to produce all responses at the same latency.

Given the paucity of evidence, we aimed to assess Mirman's (2011) proposal that influences of (semantic) variables are stronger in speeded compared to standard picture naming. To facilitate comparability of our findings to previous work, we included the same six feature-based semantic variables that we have previously investigated (Lampe, Hameau, & Nickels, in press; Lampe, Bürki, et al., 2021; Lampe, Hameau, Fieder, et al., in press). These capture aspects of the degree of activation spread at the semantic level (e.g., number of semantic features, Lampe, Hameau, & Nickels, in press; Rabovsky et al., 2016; Taylor et al., 2012) or the size and strength of activation of a co-activated lexical cohort (e.g., number of near semantic neighbours, Fieder et al., 2019; Hameau et al., 2019; Lampe, Hameau, & Nickels, in press; Mirman, 2011). More specifically, these six variables are 1) a count of semantic features associated with a concept (number of semantic features), 2) the degree of intercorrelation of the features of a concept, which characterises clusters of closely related concepts (intercorrelational density), 3) the number of words that share a substantial part of their semantic information with the target (number of near semantic neighbours), 4) the featural similarity of the target with the other words in the mental lexicon (semantic similarity), 5) a concept's representativeness of its semantic category (typicality), and 6) the degree to which the features of a concept are shared with other concepts (distinctiveness).

Importantly, as noted above, while the tempo naming task provides useful data on accuracy, we believe that naming latency data from this task is not appropriate for uncovering processes during word production (cf. Fieder et al., 2019 and Mirman, 2011) as the participants are presented with explicit and precise cues for when to initiate their response. In contrast, in the speeded deadline naming task, participants are required to respond *before* a specific deadline, thus giving them more flexibility in their actual response speed. Consequently, we used a speeded deadline naming task following the procedure suggested by Damian and Dumay (2007), in which the pictures in the speeded deadline task remained on the screen for the same duration (i.e., 2000ms) as in the standard naming task, but participants were informed on their performance with feedback that was presented after the picture offset (see also Kello et al., 2000, for a similar approach).

Methods

Participants

Eighty-three English native speakers were recruited from Macquarie University's Psychology participant pool. They provided written informed consent to participate in this study and received course credit or monetary compensation (AUD15 per hour) for their participation. Participants had to be right-handed, 17–35 years old, and had to have normal or corrected vision. Moreover, they had to be without a history of speech and language, neurological, or cognitive impairments.

All 83 participants first named the pictures in a standard naming task (analysed and reported in Lampe, Hameau, & Nickels, in press; Lampe, Bürki, et al., 2021). Subsequently, all participants named the pictures again: 42 of these participants this time named the pictures in a speeded naming task, while the other 41 participants completed another round of standard naming. Three of the tested participants had to be excluded because they did not fulfil the inclusion criteria, did not follow the task instructions, or were outliers in terms of naming accuracy (more detail provided below). Consequently, the data analysed was from 39 participants (*M* = 20.04 years old, range = 17–26 years, *SD* = 1.98, 35 females) on the speeded naming task and data from 41 participants (*M* = 20.18 years old, range = 17–33 years, *SD* = 2.66, 29 females) on the (second) standard naming task.

Stimuli

Colour photographs of 297 items from the McRae et al. (2005) feature database that had high name agreement in Australian English (> 75%; Lampe, Hameau, & Nickels, in press) were used as stimuli. They were divided into four blocks (Block 1 *n* = 35 items, Blocks 2 and 3 *n* = 87, Block 4 *n* = 88 items) and for each block we created three pseudorandomised orders with a minimum of two items from different semantic categories being presented between two items of the same category. The different pseudorandomisations of the blocks were then arranged in different orders to create 6 experimental lists (in each of the lists, Block 1 always appeared first).

For all 297 items, information on six semantic and further psycholinguistic control variables (Perret & Bonin, 2019) was available. The semantic variables were derived from, or calculated based on, information given in the McRae et al. (2005) feature database. *Number of semantic features* was a count of semantic features generated for a target word (e.g., Rabovsky et al., 2016). *Intercorrelational density* indicated the summed shared variance of a concept's correlated feature pairs (e.g., Rabovsky et al., 2016). *Number of near semantic neighbours* was a count of the concepts in the McRae et al. database that had a cosine feature vector similarity of at least .4 with the target (e.g., Hameau et al., 2019; Mirman, 2011; Mirman & Graziano, 2013). *Semantic similarity* was the average similarity between the feature vectors of the target and all other words in the database (Mirman & Magnuson, 2008). *Typicality* was based on Rosch and Mervis' (1975) family resemblance score and was calculated by summing, for each concept, the values that reflect, for each of the concept's features, the prevalence of that feature in the target's semantic category as well as its production frequency and the number of items in the target's category. Finally, *distinctiveness* was the average of the inverse number of concepts in which the features of a target occurred across the database (e.g., Rabovsky et al., 2016). Lampe, Hameau, and Nickels (in press) and Lampe, Hameau, Fieder, et al. (in press) provide more detailed descriptions of the calculation of the semantic variables.

The psycholinguistic control variables were derived from different sources: *name agreement, image agreement, imageability, age of acquisition*, and *familiarity* were rated (Lampe, Hameau, & Nickels, in press), (spoken word) *frequency* was based on television subtitles (SUBTLEX-UK; van Heuven et al., 2014), and two further control variables were retrieved from the experimental lists: *ordinal category position* indicated the number of previously seen items of the same semantic category (to account for the cumulative semantic interference effect; Howard et al., 2006), and *trial order* indexed the rank-order of an item in the experiment (to control for habituation to the experiment or fatigue; Baayen & Milin, 2017).

Procedure

This study was part of a larger investigation for which all participants completed three picture naming tasks. The first task was standard picture naming, which was analysed in Lampe, Hameau, and Nickels (in press) and Lampe, Bürki, et al. (2021) and will not be considered here. Subsequently, all participants performed the speeded deadline and standard naming tasks analysed here. Participants were not informed on the accuracy of their responses in the first standard naming task. We hoped that this approach where the participants named the stimuli in the first standard naming task would decrease any difficulties related to stimulus identification that could otherwise disproportionally affect performance on the speeded picture naming task due to the reduced processing time available (some of the previous speeded naming studies also provided a 'practice' naming round and/or familiarised their participants with the pictures and target names before the experiment; e.g., Damian & Dumay, 2007).

All participants completed the three naming tasks, and the same 297 stimuli were presented in all three tasks, but participants saw a different pseudorandomisation in each task. The order of administration of the speeded and second standard naming tasks was counterbalanced across participants. We had originally intended to assess the effect of naming task in within-participants analyses, including data of all participants for both the speeded and second standard naming task. However, effects of order of task administration interacted with various control and semantic variables and we therefore only used half of the available data and analysed only the data of the second naming round, which was speeded naming for one half of the participants and standard naming for the other half.

Picture presentation and trial-sequence in all three tasks were controlled by Presentation® (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com), which also recorded verbal responses and naming latencies using a voice trigger with the help of a Behringer preamplifier (Tube Ultragain Mic100) and a Rode NTG1 shotgun microphone. All tasks were presented on an AOC FreeSync LED monitor with a Dell Precision tower 3620 running Windows 10. The keyboard was used to navigate through the tasks.

In addition to naming latency and accuracy, a continuous electroencephalography (EEG) signal was recorded in all three naming tasks using a 64 channels ActiveTwo BioSemi system (BioSemi, Amsterdam, the Netherlands). This data is not reported here. The study procedures, but not the analyses conducted here, were preregistered on the Open Science Framework (Lampe et al., 2019; https://osf.io/yw6ma/). The data and all analysis scripts of this study are available under https://osf.io/5r8fp/.

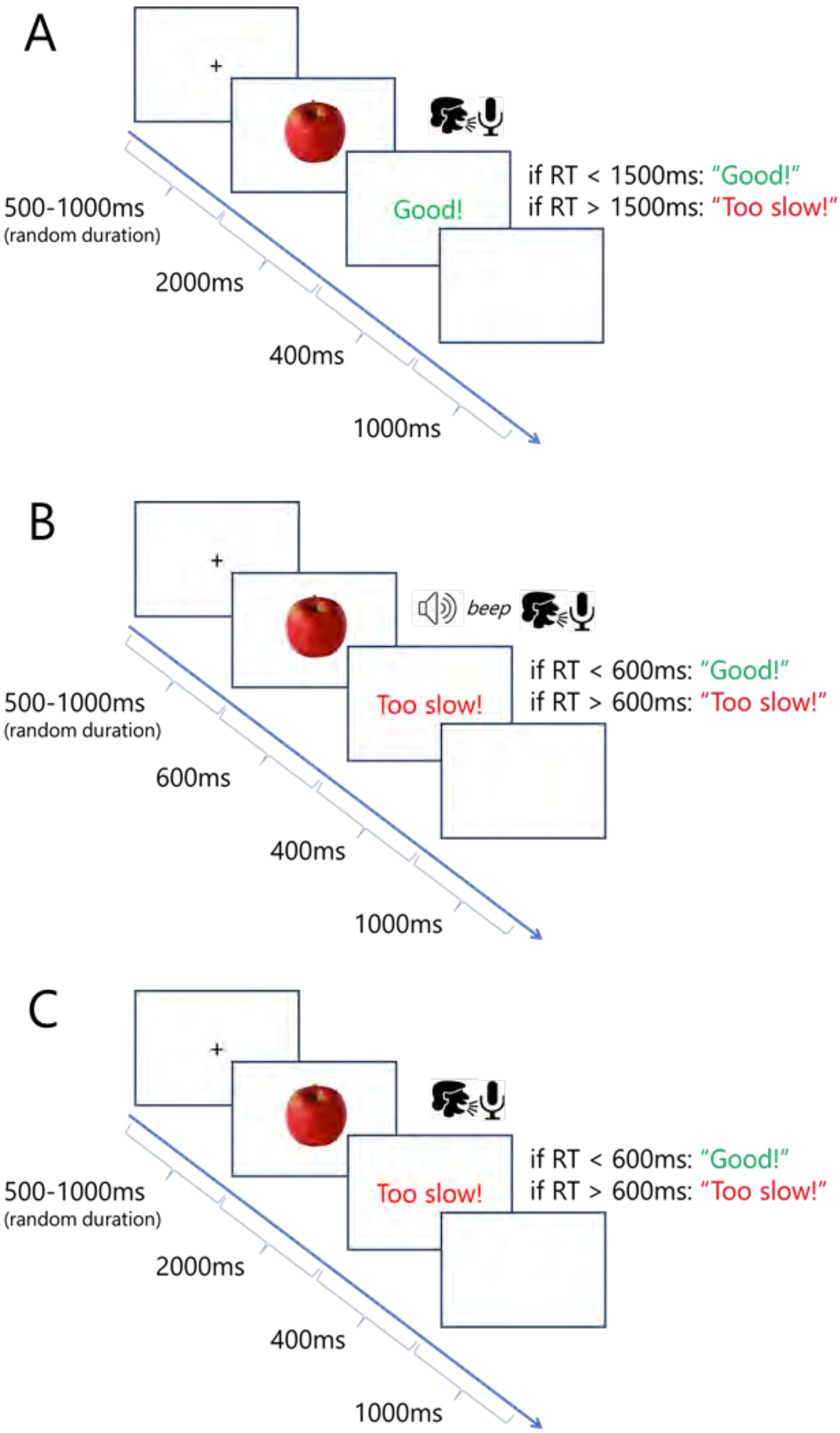
Standard naming task

Trial sequence in the standard naming task was as follows (see Figure 1, Panel A): A fixation cross was presented for a random duration between 500 and 1000ms in the middle of the screen. Then, a picture was displayed for 2000ms on white background. The participants were instructed to name the picture as quickly and accurately as possible, using a single word only, and that they would receive feedback regarding whether they had named the pictures fast enough. After the picture offset, feedback was presented for 400ms: If the participant had responded within 1500ms after onset of the picture on the screen, this feedback was positive, irrespective of response accuracy (i.e., "Good!" in green ink), else the feedback was negative ("Too slow!" in red ink). Then, the screen was blank for 1000ms before the next trial started. The feedback was presented to keep the trial sequence of the standard and speeded naming tasks as comparable as possible (see also Speeded naming task section below) and the 1500ms deadline was chosen expecting that participants would usually not have difficulties adhering to it, given that mean naming latency in the first standard picture naming round was 900ms, and therefore would not put the participants under pressure.

The task began with 4 practice trials in which participants named pictures that were not part of the 297 experimental stimuli and came from different semantic categories to the experimental stimuli. There was a break after the practice phase and after each experimental block for the participants to ask questions and to rest. The first item after each break was another practice item. Altogether, this task lasted about 30 minutes.

Figure 1

Trial sequence in standard (Panel A) and speeded naming task (practice trials in Panel B, experimental trials in Panel C)



Note. RT = Response time

Speeded deadline naming task

In the speeded naming task, the trial sequence differed between the practice (Figure 1, Panel B) and experimental trials (Figure 1, Panel C). The task began with 5 practice trials to familiarise participants with the naming speed expected in this task. In the practice trials, a fixation cross was presented for a random duration between 500 and 1000ms in the middle of the screen. Subsequently, a picture was displayed for 600ms on a white background in the centre the screen. Then, participants heard a beep and the picture disappeared from the screen. They were instructed to name the picture as quickly as possible, ideally while it was still on the screen, and to "beat the beep", prioritising naming speed over accuracy (i.e., speeded deadline naming, see also e.g., Moses et al., 2004; Vitkovitch & Humphreys, 1991). After the picture offset, feedback was presented for 400ms: If the participant had started to respond within 600ms after picture onset, this feedback was positive (i.e., "Good!" in green ink), else the feedback was negative ("Too slow!" in red ink). Subsequently, the screen was blank for 1000ms before the next trial started.

In contrast to the practice trials, in the experimental trials of the speeded naming task, the pictures were displayed for 2000ms and no beep was played (see Damian & Dumay, 2007, for a similar procedure). However, the participants were instructed to name the pictures like in the practice trials and as if there was a beep. Feedback after picture offset informed the participants on whether they had managed to initiate their response within the aspired 600ms (Figure 1, Panel C).

A break separated the experimental blocks. After each break, the participants were presented with 5 different practice items following the trial sequence of the practice items to remind them of the necessary naming speed. Each experimental block started with a sixth practice item that was presented according to the experimental trial sequence. All 24 practice items were not part of the experimental stimuli and came from different semantic categories. This task also lasted about 30 minutes.

Data analysis

After the experiment, responses from both tasks were transcribed and checked for naming accuracy. The first response that consisted of at least one English syllable was coded for accuracy. Correct responses consisted of (only) the correctly named target word and these responses were analysed in the naming latency analysis. For the naming accuracy analysis, responses were also coded as correct if they were correct responses that were preceded by a determiner (e.g., *a box*), a hesitation (e.g., *umm, box*), or an elaboration (e.g., *cardboard box* for *box*), or contained a dysfluency on the initial phoneme of the target (e.g., *b..box*). Abbreviations, synonyms or acceptable alternative responses, and trials where a response from the previous trial intruded into the current trial (but where the target item was subsequently named correctly) were not analysed (NA). All other response types (dysfluencies with production of a non-target phoneme (e.g., *st raspberry*), incomplete (i.e., responses that were not legal English syllables: a single consonant or short vowel, or CV responses of consonants followed by a short vowel), incorrect responses, and omissions) were coded as naming errors.

The recorded naming latency was manually adjusted through visual and auditory inspection of the waveform and the voice recording using Praat (Version 6.0.49; Boersma & Weenink, 2019) to derive maximally accurate naming latencies. All statistical analyses were conducted in R Studio (Version 1.3.959; RStudio Team, 2020). We ran (Generalised) Linear Mixed Effect Models on the naming latency and the naming accuracy data using the lme4-package (Version 1.1.25; Bates, Mächler, et al., 2015), with *p*-values being derived using lmerTest (Version 3.1.3; Kuznetsova et al., 2017). For the random effects structure we considered random by-participant slopes for the six semantic variables and followed Bates, Kliegl, et al. (2015) for the specification of a simplified random effects structure. Model fit was compared using likelihood ratio tests (stats package, Version 4.0.2; R Core Team, 2020). Interactions were plotted using the package sjPlot (Version 2.8.6, Lüdecke, 2020).

All semantic and psycholinguistic control variables were standardised using a *z*-transformation. Based on the output of the boxcox function (EnvStats package, Version 2.4.0; Millard, 2013), naming latency was transformed to approximate a normal distribution (negative reciprocal transformation, which preserves order among values of the same sign).

Naming accuracy outliers for the two tasks separately were identified based on visual inspection of boxplots. In the standard naming task, data of one participant who performed considerably less accurately than the other participants (68% mean naming accuracy versus mean accuracy of 91% in the remaining participants, range = 78–99%, *SD* = 5%) was excluded. No participant had to be removed from the speeded naming data (*M* = 76% naming accuracy, range = 59–91%, *SD* = 7%). Moreover, 18 items were outliers in terms of naming accuracy in the standard naming task with an average accuracy of 60% (range = 49–68%, *SD* = 6%), in contrast to an average naming accuracy of 93% for the remaining items (range = 70–100%, *SD* = 8%). However, the data points associated with these items were part of the tail of the distribution and their removal comes at the cost of statistical power. We consequently tested if their exclusion changed the findings compared to analyses conducted on all items. As the findings were mostly comparable, with no changes in the interactions between semantic variables and task in either of the analyses, we focus on the analyses performed on all 297 items. However, the findings of the analyses conducted after removing these items are reported in Appendix A.

For the naming accuracy analysis, to be confident that any naming errors were not caused by a misidentification of or unfamiliarity with the depicted object, we tested whether the findings differed when including only trials that were named correctly by the participant in the first (standard) naming round that was reported in Lampe, Hameau, and Nickels (in press) (i.e., items where familiarity with and successful identification of the depicted object under normal processing conditions can be guaranteed). However, given the low number of errors in the standard naming task when focusing on trials that were named accurately in the first naming round (*n* = 254 naming errors from 10,461 data points), we were only able to test for such differences in the speeded naming task (*n* = 1,451 naming errors from 9,480 data points for items that were named correctly in the first standard naming task). No differences in the effects of semantic variables were found between Linear Mixed Effect Models on naming latency in speeded picture naming including or excluding trials depending on their naming accuracy in the first standard naming round. Consequently, we performed the analysis of naming accuracy on all trials, irrespective of their accuracy in the first naming round.

The naming latency analysis was conducted on correctly named items, irrespective of their actual naming latency (i.e., we did not exclude trials with naming latencies > 600ms in the speeded naming task or trials with naming latencies > 1500ms in the standard naming task), resulting in 19,519 data points from 297 items and 80 participants, which consisted of 10,864 data points from standard naming and 8,655 data points from speeded naming. The naming accuracy analysis was conducted on all items, except for those coded 'NA' for naming accuracy, resulting in 23,608 data points from 297 items and 80 participants (*n* = 3,820 naming errors), consisting of 12,100 data points from the standard naming task (*n* = 1,097 naming errors) and 11,508 data points from the speeded naming task (*n* = 2,723 naming errors). Following Mirman (2011), we also conducted an analysis that contrasted the proportion of semantic errors and correct responses. The findings were comparable to the naming accuracy analysis and can be found in Appendix B.

To test whether effects of the six semantic and eight psycholinguistic control variables were larger in the speeded than in the standard naming task, we combined the data from the two tasks and tested for significant interactions between task and these variables. Subsequently, the significant interactions were unpacked in planned comparisons using the emtrends function of the R package emmeans (Version 1.5.2-1, Lenth, 2020) to investigate the significance of the effect in the two tasks separately and to determine whether the effects were indeed stronger in the speeded naming task. Task was treatment coded (speeded as 0.5 and standard as -0.5).

Results

There were main effects of task in the naming latency and accuracy analyses (Table 1 and Table 2, respectively), indicating that the speeded deadline naming procedure yielded the desired effect and caused a speed-accuracy trade-off with faster but less accurate naming compared to the standard naming task. More specifically, participant mean naming latency in the speeded naming task was 609ms (range = 506–718ms, SD = 56ms) and 760ms in the standard naming task (range = 650–916ms, SD = 68ms), and naming accuracy in the speeded naming task was 76% (range = 59–91%, SD = 7%) and 91% in the standard naming task (range = 78–99%, SD = 5%). We will now address the findings of the naming latency and accuracy analyses.

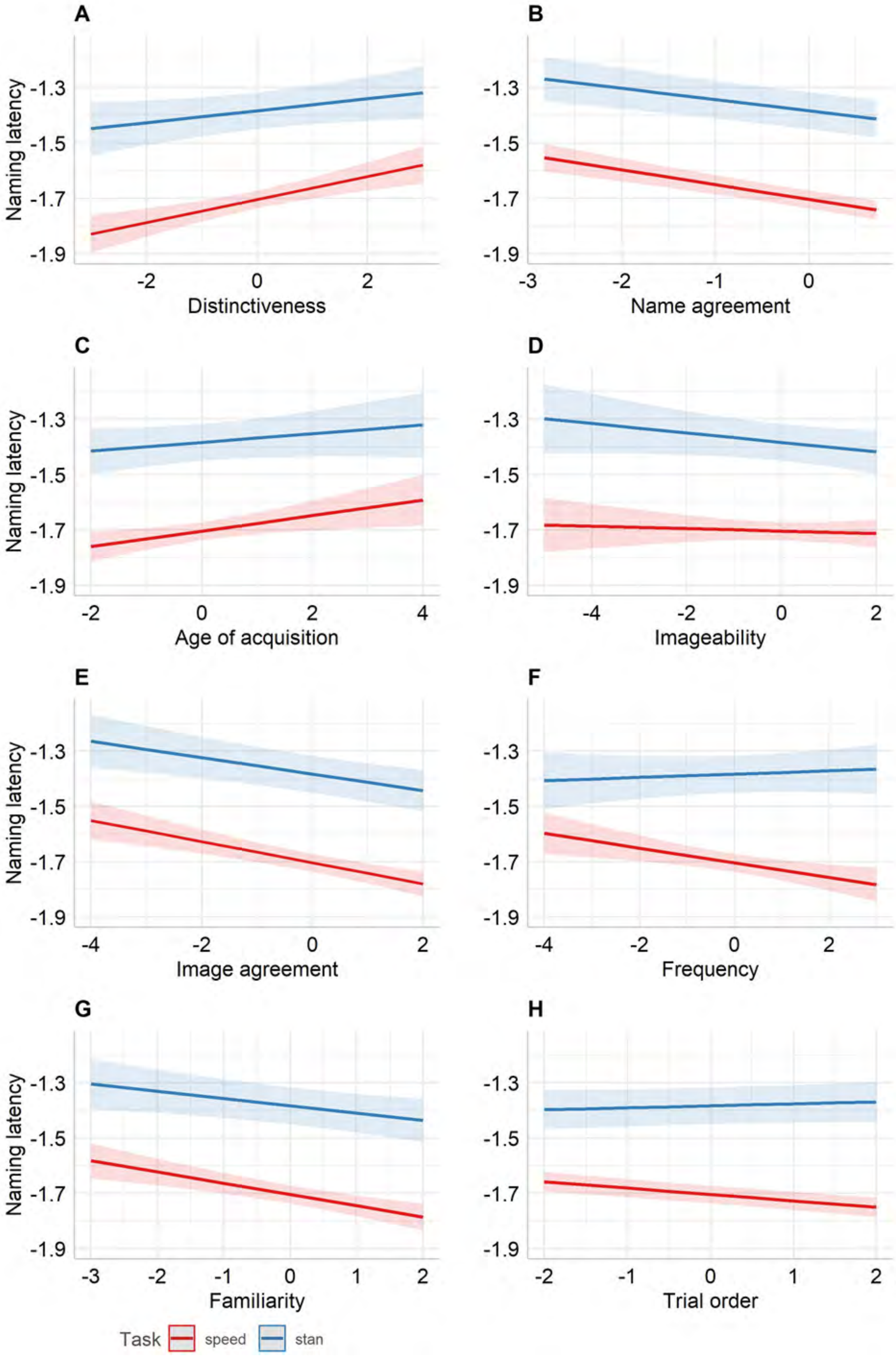
Naming latency

The findings of the naming latency analysis are summarised in Table 1. The only semantic variable that significantly interacted with naming task was distinctiveness. Analysis of this interaction confirmed that the effect was significant and inhibitory in both naming tasks with a stronger effect in speeded naming (Figure 2, Panel A). Moreover, all control variables, except for ordinal category position, significantly interacted with task. The effects were stronger in the speeded naming task than in the standard naming task for name agreement (facilitatory in both tasks, Figure 2, Panel B), age of acquisition (inhibitory only in speeded naming, Figure 2, Panel C), image agreement (facilitatory in both tasks, Figure 2, Panel E), frequency (facilitatory only in speeded naming, Figure 2, Panel F), familiarity (facilitatory in both tasks, Figure 2, Panel G), and trial order (facilitatory only in speeded naming, Figure 2, Panel H). In contrast, the effect of imageability was stronger in the standard naming task (Figure 2, Panel D), however, this facilitatory effect was non-significant in both tasks in the follow-up analysis. This could suggest that the significant interaction was a Type 1 error, a false positive finding, and that there was, in reality, no difference between the two simple effects. Alternatively, this could be indicative of a Type 2 error, a false negative finding, incorrectly stating the absence of simple effects in one or both tasks. Given that we statistically controlled for Type 1 errors reasonably well (i.e., our significance level of .05 allowed for a 5% probability of incorrectly rejecting the null hypothesis that there is no difference in the effect of imageability between the two tasks), the most likely scenario is that we made a Type 2 error due to insufficient statistical power. However, our data is unable to differentiate between these options and further research is needed to do so.

In addition, some of the main effects of semantic and control variables were significant: Combined across tasks, responses were slower the higher the number of near semantic neighbours, the greater the distinctiveness, and the higher the age of acquisition of the target. Moreover, responses were faster the higher the name agreement, image agreement, and familiarity of the target, and the later in the experiment a trial occurred.

Figure 2

Significant interactions between semantic or control variables and task in the naming latency analysis



Note. Naming latency was negative reciprocally transformed; all variables were standardised.

Table 1

Naming latency: summarised output of Linear Mixed Model analysis for task comparison and simple effects

Model structure	lmer(RT ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + OrdCatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness) * Task + (1 | Item) + (NearSemNeigh + SemSim || Participant), data, REML = TRUE)					
Random effect	Variance	*SD*				
Item (Intercept)	0.01	0.12				
Participant (Intercept)	0.02	0.13				
Participant | NrSemNeigh	0.00	0.02				
Participant | SemSim	0.00	0.01				
Residuals	0.06	0.24				
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF
(Intercept)	-1.54	0.02	-1.58 – -1.51	-95.53	<.001	
NameAgr	-0.05	0.01	-0.06 – -0.03	**-6.59**	**<.001**	1.11
AoA	0.02	0.01	0.00 – 0.04	**2.01**	**.045**	2.45
Imageability	-0.01	0.01	-0.03 – 0.01	-1.14	.254	1.88
ImageAgr	-0.03	0.01	-0.05 – -0.02	**-4.34**	**<.001**	1.25
Frequency	-0.01	0.01	-0.03 – 0.01	-1.18	.239	1.61
Familiarity	-0.03	0.01	-0.05 – -0.02	**-3.60**	**<.001**	1.74
Order	-0.01	0.00	-0.02 – -0.00	**-2.15**	**.032**	3.07
OrdCatPos	0.00	0.00	-0.01 – 0.01	0.63	.531	3.14
NoFeats	-0.00	0.01	-0.02 – 0.01	-0.41	.681	1.57
IntercorrDens	0.02	0.01	-0.00 – 0.04	1.64	.101	2.23
NearSemNeigh	0.02	0.01	0.00 – 0.05	**2.26**	**.024**	2.37
SemSim	0.01	0.01	-0.01 – 0.03	0.84	.404	2.40

Typicality	-0.00	0.01	-0.02 – 0.01	-0.37	.711	1.44	Simple effects of variables in significant interactions							
Distinctiveness	0.03	0.01	0.01 – 0.05	**3.11**	**.002**	2.01	Speeded naming				Standard naming			
Task	-0.32	0.03	-0.38 – -0.26	**-10.95**	**<.001**	1.00	Estimate	*SE*	*z*-value	*p*-value	Estimate	*SE*	*z*-value	*p*-value
NameAgr * Task	-0.01	0.00	-0.02 – -0.01	**-3.46**	**.001**	1.11	-0.05	0.01	**-7.18**	**<.001**	-0.04	0.01	**-5.56**	**<.001**
AoA * Task	0.01	0.01	0.00 – 0.02	**2.24**	**.025**	2.38	0.03	0.01	**2.47**	**.013**	0.02	0.01	1.41	.157
Imageability * Task	0.01	0.00	0.00 – 0.02	**2.63**	**.009**	1.85	-0.00	0.01	-0.46	.649	-0.02	0.01	-1.77	.077
ImageAgr * Task	-0.01	0.00	-0.02 – -0.00	**-2.14**	**.032**	1.20	-0.04	0.01	**-4.69**	**<.001**	-0.03	0.01	**-3.71**	**<.001**
Frequency * Task	-0.03	0.00	-0.04 – -0.02	**-7.29**	**<.001**	1.55	-0.03	0.01	**-2.90**	**.004**	0.01	0.01	0.65	.515
Familiarity * Task	-0.01	0.00	-0.02 – -0.01	**-3.08**	**.002**	1.63	-0.04	0.01	**-4.21**	**<.001**	-0.03	0.01	**-2.76**	**.006**
Order * Task	-0.03	0.01	-0.04 – -0.02	**-5.74**	**<.001**	2.24	-0.02	0.00	**-4.85**	**<.001**	0.01	0.00	1.63	.103
OrdCatPos * Task	-0.01	0.01	-0.02 – 0.00	-1.17	.241	2.68								
NoFeats * Task	-0.01	0.00	-0.01 – 0.00	-1.37	.172	1.64								
IntercorrDens * Task	-0.00	0.01	-0.01 – 0.01	-0.20	.845	2.04								
NearSemNeigh * Task	0.01	0.01	-0.01 – 0.02	1.09	.274	1.59								
SemSim * Task	0.01	0.01	-0.00 – 0.02	1.45	.148	1.70								
Typicality * Task	-0.00	0.00	-0.01 – 0.01	-0.38	.701	1.48								
Distinctiveness * Task	0.02	0.01	0.01 – 0.03	**3.92**	**<.001**	1.67	0.04	0.01	**3.95**	**<.001**	0.02	0.01	**2.08**	**.037**
Observations: 19,519														
Marginal R^2 / Conditional R^2: 0.287 / 0.531														

Note. Participant | X = random slope of X by participants, VIF = Variance Inflation Factor, NameAgr = Name agreement, AoA = Age of acquisition, ImageAgr = Image agreement, Order = Trial order, OrdCatPos = Ordinal category position, NoFeats = Number of semantic features, IntercorrDens = Intercorrelational Density, NearSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity.

Values of significant effects (p < .05) are printed in bold; in the simple effects, the variable with the stronger effect within significant interactions is highlighted in grey.

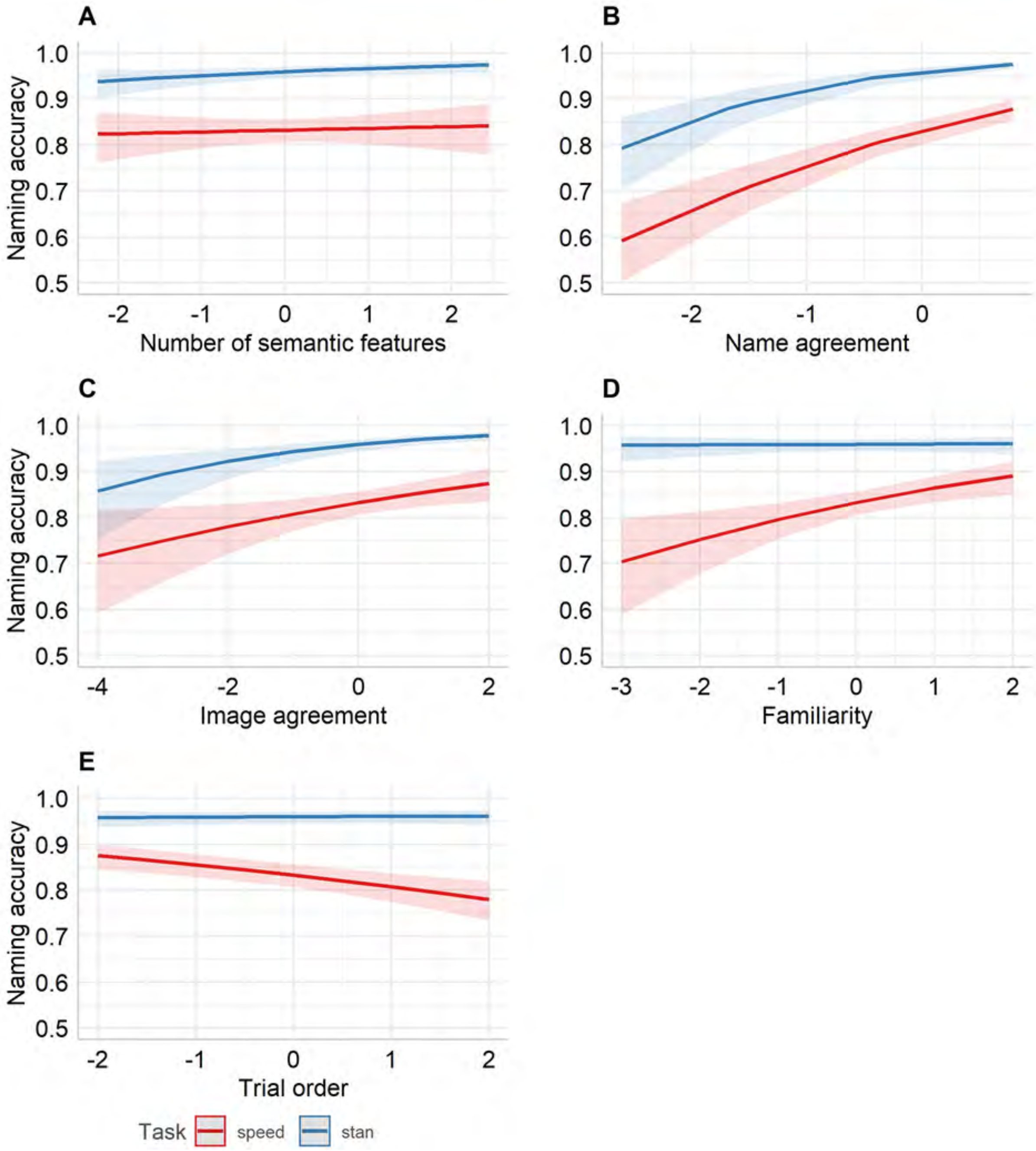
Naming accuracy

In the naming accuracy analyses the only semantic variable that significantly interacted with naming task was number of semantic features (see Table 2). The planned follow-up analysis revealed that the facilitatory effect was stronger in the standard naming task (Figure 3, Panel A) and non-significant in speeded naming. Moreover, some control variables significantly interacted with task: The effects were stronger in the standard naming task for name agreement (facilitatory in both tasks, Figure 3, Panel B) and image agreement (facilitatory in both tasks, Figure 3, Panel C). In contrast, effects were stronger in the speeded naming task for familiarity (facilitatory only in speeded naming, Figure 3, Panel D) and trial order (inhibitory only in speeded naming, Figure 3, Panel E).

When the data were combined across both tasks, responses were less accurate for words with higher intercorrelational density and words that are acquired later in life. In contrast, responses were more accurate for words with higher name agreement, imageability, image agreement, and frequency.

Figure 3

Significant interactions between semantic or control variables and task in the naming accuracy analysis



Note. All variables were standardised.

Table 2

Naming accuracy: summarised output of Generalised Linear Mixed Model analysis for task comparison and simple effects

Model structure	glmer(ACC ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + CatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness)*Task + (1 | Item) + (NearSemNeigh | Participant), data, family = binomial())							
Random effect	Variance	*SD*	Correlation					
Item (Intercept)	0.89	0.94						
Participant (Intercept)	0.34	0.58						
Participant | NearSemNeigh	0.02	0.14	0.76					
Fixed effects	Estimate	*SE*	*CI*	*z*-value	*p*-value	VIF	Simple effects of variables in significant interactions	
							Speeded naming	Standard naming
(Intercept)	2.39	0.09	2.21 – 2.56	26.21	<.001			
NameAgr	0.59	0.06	0.46 – 0.71	**9.41**	**<.001**	1.11		
AoA	-0.21	0.10	-0.40 – -0.03	**-2.24**	**.025**	2.55		
Imageability	0.20	0.08	0.04 – 0.37	**2.46**	**.014**	1.93		
ImageAgr	0.26	0.07	0.12 – 0.39	**3.75**	**<.001**	1.29		
Frequency	0.17	0.08	0.02 – 0.32	**2.22**	**.026**	1.63		
Familiarity	0.13	0.08	-0.03 – 0.29	1.64	.101	1.80		
Order	-0.07	0.04	-0.16 – 0.01	-1.70	.089	3.18		
OrdCatPos	-0.01	0.05	-0.11 – 0.09	-0.25	.803	3.38		
NoFeats	0.11	0.08	-0.04 – 0.26	1.47	.141	1.61		
IntercorrDens	-0.20	0.09	-0.37 – -0.02	**-2.17**	**.030**	2.27		
NearSemNeigh	-0.11	0.10	-0.29 – 0.08	-1.09	.274	2.47		
SemSim	-0.07	0.10	-0.26 – 0.12	-0.77	.443	2.52		
Typicality	0.05	0.08	-0.10 – 0.20	0.69	.491	1.52		
Distinctiveness	-0.08	0.09	-0.25 – 0.09	-0.92	.357	2.07		

							Estimate	*SE*	*z*-value	*p*-value	Estimate	*SE*	*z*-value	*p*-value
Task	-1.56	0.14	-1.84 – -1.29	**-11.07**	**<.001**	1.26								
NameAgr * Task	-0.23	0.04	-0.31 – -0.14	**-5.44**	**<.001**	1.10	0.47	0.06	**7.45**	**<.001**	0.70	0.07	**10.31**	**<.001**
AoA * Task	0.12	0.07	-0.02 – 0.25	1.69	.092	2.72								
Imageability * Task	-0.01	0.06	-0.12 – 0.10	-0.21	.830	1.98								
ImageAgr * Task	-0.17	0.05	-0.27 – -0.08	**-3.67**	**<.001**	1.34	0.17	0.07	**2.44**	**.015**	0.34	0.07	**4.58**	**<.001**
Frequency * Task	0.10	0.05	-0.01 – 0.20	1.77	.078	1.60								
Familiarity * Task	0.23	0.06	0.11 – 0.35	**3.78**	**<.001**	1.90	0.25	0.08	**3.02**	**.003**	0.02	0.09	0.21	.835
Order * Task	-0.19	0.06	-0.32 – -0.07	**-3.01**	**.003**	2.36	-0.17	0.05	**-3.59**	**<.001**	0.02	0.06	0.38	.705
OrdCatPos * Task	0.02	0.07	-0.12 – 0.16	0.22	.825	2.93								
NoFeats * Task	-0.17	0.06	-0.28 – -0.06	**-2.97**	**.003**	1.72	0.03	0.08	0.36	.718	0.20	0.09	**2.31**	**.021**
IntercorrDens * Task	0.10	0.06	-0.03 – 0.22	1.53	.127	2.28								
NearSemNeigh * Task	0.02	0.07	-0.13 – 0.16	0.25	.805	2.60								
SemSim * Task	-0.02	0.07	-0.16 – 0.13	-0.22	.825	2.59								
Typicality * Task	0.01	0.06	-0.11 – 0.12	0.12	.904	1.78								
Distinctiveness * Task	0.06	0.06	0.07 – 0.18	0.89	.374	2.12								
Observations: 23,608														
Marginal R^2 / Conditional R^2: 0.261 / 0.465														

Note. Participant | X = random slope of X by participants, VIF = Variance Inflation Factor, NameAgr = Name agreement, AoA = Age of acquisition, ImageAgr = Image agreement, Order = Trial order, OrdCatPos = Ordinal category position, NoFeats = Number of semantic features, IntercorrDens = Intercorrelational Density, NearSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity.

Values of significant effects (p < .05) are printed in bold; in the simple effects, the variable with the stronger effect within significant interactions is highlighted in grey.

Discussion

This research aimed to test the hypothesis that effects of semantic variables are stronger in a speeded deadline naming task compared to a standard picture naming task. This followed a suggestion by Mirman (2011) that greater responsiveness to input, due to higher input gain in the speeded naming task, may cause stronger inhibitory and facilitatory influences of item-inherent variables. We focused on six feature-based semantic variables (i.e., number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness) and compared their effects in two groups of participants naming pictures either in non-speeded standard picture naming or in a speeded naming task where they had to prioritise naming speed over accuracy and aim to initiate their response within 600ms of picture presentation. Consistent with the aim of this research, in this Discussion, we focus on the interactions between naming tasks and the effects of the semantic variables.

As is clear from Table 3, few interactions between semantic variables and task were significant. In the naming latency analysis, the only significant interaction was with distinctiveness: While the effect of distinctiveness was significant and inhibitory in both tasks, it was stronger in the speeded naming task. In the naming accuracy analysis, the only significant interaction between task and a semantic variable was with number of semantic features. The facilitatory effect of number of semantic features was significant in standard picture naming, with a non-significant effect in speeded naming.

Table 3

Summary of the results of all naming latency and accuracy analyses

	Naming latency			Naming accuracy		
		Simple effects			Simple effects	
Variable	Interaction	Speeded	Standard	Interaction	Speeded	Standard
Name agreement	√	↗	↗	√	↗	↗
Age of acquisition	√	↙	Ø			
Imageability	√	Ø	Ø			
Image agreement	√	↗	↗	√	↗	↗
Frequency	√	↗	Ø			
Familiarity	√	↗	↗	√	↗	Ø
Trial order	√	↗	Ø	√	↙	Ø
Ordinal category position	Ø					
Number sem. features	Ø			√	Ø	↗
Intercorrelational density	Ø					
Number near sem. neighbours	Ø					
Semantic similarity	Ø					
Typicality	Ø					
Distinctiveness	√	↙	↙			

Note. √ = significant interaction of variable and task, the simple effect of the task with the stronger effects is highlighted in grey, Ø = non-significant interaction or simple effect, ↙ = significantly poorer performance (slower responses and decreased accuracy with higher values of the measure), ↗ = significantly improved performance (faster responses with increased accuracy and higher accuracy with higher values of the variable).

Given the small number of significant interactions between semantic variables and task, it is relevant to also consider the interactions between the psycholinguistic control variables and task. In the naming latency analysis, effects of all the control variables that interacted with task (i.e., name agreement, age of acquisition, image agreement, frequency, familiarity, and trial order) were stronger in the speeded naming task, except for imageability (although there was no evidence for a significant effect of imageability in either task separately). In contrast, in the naming accuracy analysis, effects of name agreement and image agreement were stronger in the standard naming task and familiarity and trial order were stronger predictors in the speeded naming task. Interestingly, the effect of trial order was significant only in the speeded naming task in both analyses. It showed that participants got faster, but less accurate throughout the speeded naming experiment, which most likely indicates that participants needed time to get used to the task requirements.

Importantly, in the naming latency analysis, in contrast to accuracy, effects of item-inherent variables (semantic and control variables) were overall stronger in the speeded naming task, thus largely confirming our initial hypothesis. However, while the difference between the tasks were as predicted for latency, can Mirman's (2011) input gain account explain the stronger effect of distinctiveness in the speeded naming task?

We also reported an inhibitory effect of distinctiveness on naming latency in Lampe, Hameau, and Nickels (in press), where we investigated semantic variables using the data from the first standard naming task collected before the two tasks analysed in this study. This finding was in contrast to previous work that had reported a facilitatory effect of distinctiveness on picture naming (Rabovsky et al., 2016; Taylor et al., 2012), which was argued to be due to preferential processing of distinctive features at the semantic level. By definition, words with higher distinctiveness have more unique semantic features and thus share their features with fewer other concepts. This causes *fewer* semantically related lexical representations to be co-activated during processing of such words. Hence, the negative effect of distinctiveness cannot be due to enhanced lexical competition for words with higher distinctiveness (Lampe, Hameau, & Nickels, in press). The inhibitory effect of distinctiveness also cannot be explained in the context of an attractor model of semantic cognition (e.g., Cree et al., 1999; Rogers et al., 2004): A more distinctive concept would be further from other concepts and should therefore not be subject to high levels of interference from other attractors when the model is settling into the target attractor (similar to the argument for distant semantic neighbours provided by Mirman, 2011). Consequently, given that enhanced competition is unlikely to underpin the inhibitory effect of distinctiveness, it is unclear exactly how, under the input gain account, input amplification could cause the stronger effect of distinctiveness in speeded naming. Further research into the effect of this variable is needed to understand the mechanisms underlying effects of distinctiveness before hypotheses can sensibly be drawn as to why the effect may be stronger in speeded naming.

Number of semantic features was the only semantic variable that interacted with task in the naming accuracy analysis, with a stronger effect in the standard naming task and no evidence for a significant effect in speeded naming. The facilitatory effect of more semantic features in standard naming is in line with previous research on naming speed and accuracy (Lampe, Hameau, & Nickels, in press; Rabovsky et al., 2016, 2021; Taylor et al., 2012). It has been attributed to the increased activation of the target's lexical representation following enhanced semantic activation, due to the many active semantic features (see Rabovsky & McRae, 2014, for a simulation). Alternatively, in attractor models, it has been proposed that they can settle faster and more accurately for words with many semantic features because they are represented by stronger attractor basins, which facilitate the system's ability to settle into a stable pattern of activation (Plaut & Shallice, 1993; see also Pexman et al., 2007).

Under the input gain account, it would be expected that the facilitation of a target's lexical selection would be stronger or settlement into its attractor faster in *speeded* naming. However, we found the reverse, and, in fact, there was no evidence that processing in the speeded naming task was affected at all by the number of semantic features in this (accuracy) analysis. In sum, it certainly was not the case that semantic and/or lexical processing exhibited increased sensitivity to the activation of many semantic features in the speeded naming task. An alternative possibility is that, in speeded naming, selection thresholds for processing to continue are lowered in order that the system can generate a response quickly (*threshold account*, e.g., Kello, 2004), which may result in incomplete processing. Thus, due to the time pressure, processing in speeded naming would not benefit from the activation of many semantic features associated with concepts with a higher number of semantic features (resulting in no significant effects). In contrast, in standard naming, where there is ample time to process all the available information, all semantic features are activated, facilitating performance (see also Kello & Plaut, 2000, for an argument that effects of variables would be predicted to be reduced in speeded reading).

Critically, taken together, the data presented here does not provide evidence for systematically stronger effects of semantic variables in speeded compared to standard picture naming. In addition, even when, for distinctiveness, there was a stronger effect in speeded naming, because of the direction of this effect, enhanced responsiveness to inputs following disruptions of cognitive control in the speeded picture naming task does not seem to be a plausible underlying mechanism. Hence, differences in the specific naming task used are unlikely to be the exclusive source of contrasting effects of semantic variables in the literature, even for number of near semantic neighbours and semantic similarity for which previous significant effects were only found using a speeded naming task.

In contrast, differences between studies in the effects of semantic variables are more likely related to other factors. It seems that even when significant, effects of the semantic variables on processing are relatively small. They also show some variability between participants as demonstrated by the significant by-participant random slopes for number of near semantic neighbours and semantic similarity in our analyses (see also Hameau et al., 2019; Lampe, Hameau, & Nickels, in press, for discussion of between participants variability in the effects of semantic variables). It is particularly important to account for such variability when studying relatively small effects. This also highlights the importance of strongly powered investigations and adequate control of effects of other psycholinguistic variables in the statistical analyses. Here, we controlled for all those variables suggested by Perret and Bonin (2019), but many previous studies did not (see Appendix F in Lampe, Hameau, Fieder, et al., in press), which may have caused false positive effects of semantic variables. Most importantly though, the observed failures to replicate effects of semantic variables indicate that further work is necessary to fully establish the robustness of the effects (Bishop, 2018).

Interestingly, while the data for the semantic variables presented here does not support the input gain account, we did find that effects of the control variables on naming latency were mostly stronger in speeded than in standard naming, as noted above. The variables with stronger effects in speeded naming are associated with many different stages of word production (e.g., image agreement with visual processing, name agreement with link between semantic to lexical processing, frequency with lexical processing, and age of acquisition with phonological processing, e.g., Alario et al., 2004), with the exception of the semantic level. This suggests that the requirement to name pictures quickly may modulate processing dynamics at several stages of word production. Hence, it is possible that, while Mirman (2011) proposed the input gain mechanism for semantic processing in word production, the account may (also) be applied to other levels of processing[3]. In speeded naming, processing units

[3] We believe that even Mirman's (2011) findings of facilitatory and inhibitory effects of distant and near semantic neighbours, respectively, could be explained by enhanced sensitivity to input at the lexical level rather than the semantic level, as argued by Mirman. Specifically, the semantic representation of a target with many distant

at all these levels may become more sensitive to their inputs, leading to stronger effects. However, given that the mechanisms causing effects of many of the control variables are not fully understood, further research is needed to understand how exactly the stronger effects of these variables could be a product of enhanced sensitivity to input at any stage of the model (e.g., How *exactly* does age of acquisition affect processing to result in an inhibitory effect and can enhanced sensitivity to input explain the stronger effect in speeded naming?).

However, it is also possible that the enhanced effects of the control variables in speeded naming are caused by a mechanism other than changes in input gain. For the enhanced effect of number of semantic features on standard naming, we suggested (above) that a lowered selection threshold in speeded naming may result in incomplete processing of the available information, causing the observed reduction of the effect. But could the threshold account also explain findings in the opposite direction: stronger effects of the control variables in speeded naming? With a lowered selection threshold (e.g., Coltheart et al., 2001; Humphreys et al., 1995; Kello, 2004; or selection after fewer time steps, e.g., Dell, 1986; Dell et al., 1997), participants would be able to generate a response faster, but at increased error rates, which is the case in speeded naming. However, this mechanism would cause faster processing of *all* experimental items regardless of their item characteristics (e.g., word frequency). Whether or not the size of effects of these item-inherent variables would remain comparable or change in speeded naming would depend on the precise mechanism by which these variables are implemented. For example, if frequency is implemented as a difference in resting levels of activation (e.g., Morton, 1969, 1970) or the equivalent (e.g., Coltheart et al., 2001) the effects would remain comparable across tasks. However, if the effects are multiplicative, for example, through a frequency weighting on connections, the fewer processing cycles that there are before a response, the

semantic neighbours may receive converging activation following activation spread between the target and its many distant semantic neighbours. This may cause its lexical representation to be strongly activated, while, importantly, none of the distant semantic neighbours is semantically close enough to compete with the target for selection at the lexical level. In contrast, while the lexical representation of a target with many near semantic neighbours may also be facilitated following spreading activation at the semantic level, it is also subject to enhanced lexical competition from its co-activated near semantic neighbours. Consequently, if the lexical processing units are more sensitive to their inputs in speeded naming, both facilitatory and inhibitory effects may be localised at the lexical level, even though they originate at the semantic level.

smaller the effect of the variable. A reduction in the size of the effects would also be the case if effects were driven by feedback, like number of semantic features, as described above. It is hard, however, to conceptualise a mechanism that would result in larger effects of a variable within a threshold account. Nevertheless, this is the result we observed: a *bigger* difference in naming latency between, for example, words of lower and higher frequency in speeded than in standard naming. Hence, the threshold account is an unlikely candidate to explain these findings. To better understand the mechanism underlying effects of semantic and other psycholinguistic variables and their modulations in the speeded naming task, further investigations in combination with computational simulations of the findings (following e.g., Humphreys et al.,1995; Kello & Plaut, 2003) are necessary.

Finally, when interpreting the findings of this study, it must be kept in mind that, in contrast to most previous studies investigating effects of semantic variables, we analysed picture naming data that was collected following a previous naming attempt of the pictures in the first standard naming round. While we did not inform the participants on the accuracy of their response, the previous exposure may have nonetheless affected the findings. For example, Rabovsky et al. (2016) familiarised half of their participants with the pictures and their correct names before their experiment and found that the effect of number of semantic features on naming latency was attenuated (though still significant) in the familiarised participants. In contrast, Rabovsky et al. (2021) presented all experimental items twice to increase the number of data points for the analyses. In their analyses, task repetition did not interact with the number of semantic features or intercorrelational density in the accuracy analysis, but they found that the effect of intercorrelational density on naming latency was only significant in the repetition of the items. More research into whether and how a previous attempt at naming the same pictures influences effects of semantic variables is required.

In sum, the results reported here tested the hypothesis that effects of item-inherent semantic variables are amplified in speeded deadline naming compared to standard picture naming due to increased sensitivity to inputs (i.e., higher input gain) following speed-induced disruptions of cognitive control (Mirman, 2011). Few effects of semantic variables differed between the two tasks, and while the effect of distinctiveness (and other psycholinguistic control variables) was stronger in speeded than in standard naming when analysing naming latencies, the effect of number of semantic features was stronger in the standard naming task in the naming accuracy analysis. Therefore, the speeded naming task is unlikely to be the sole reason for significant effects being reported in some previous investigations of semantic variables and not in others that used standard picture naming.

References

Abdel Rahman, R., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience*, *34*(9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Baayen, H. R., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28. https://doi.org/10.21500/20112084.807

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv*, 1–27. https://doi.org/arXiv:1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bishop, D. V. M. (2018). Fallibility in science: Responding to errors in the work of oneself and others. *Advances in Methods and Practices in Psychological Science*, *1*(3), 432–438. https://doi.org/10.1177/2515245918776632

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 6.0.49). https://doi.org/http://www.praat.org/

Bormann, T. (2011). The role of lexical-semantic neighborhood in object naming: Implications for models of lexical access. *Frontiers in Psychology*, *2*, 1–11. https://doi.org/10.3389/fpsyg.2011.00127

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295x.108.1.204

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*(3), 371–414. https://doi.org/10.1207/s15516709cog2303_4

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, *57*(2), 195–209. https://doi.org/10.1016/j.jml.2006.11.001

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, *104*(1), 123–147. https://doi.org/10.1037/0033-295X.104.1.123

Fieder, N., Wartenburger, I., & Abdel Rahman, R. (2019). A close call: Interference from semantic neighbourhood density and similarity in language production. *Memory & Cognition*, *47*(1), 145–168. https://doi.org/10.3758/s13421-018-0856-y

Gerhand, S., & Barry, C. (1999). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, *73*(2), B27–B36. https://doi.org/10.1016/S0010-0277(99)00052-9

Gotts, S. J., & Plaut, D. C. (2002). The impact of synaptic depression following brain damage: A connectionist account of "access/refractory" and "degraded-store" semantic impairments. *Cognitive, Affective and Behavioral Neuroscience*, *2*(3), 187–213. https://doi.org/10.3758/CABN.2.3.187

Hameau, S., Nickels, L., & Biedermann, B. (2019). Effects of semantic neighbourhood density on spoken word production. *Quarterly Journal of Experimental Psychology*, *72*(12), 2752–2775. https://doi.org/10.1177/1747021819859850

Hodgson, C., & Lambon Ralph, M. A. (2008). Mimicking aphasic semantic errors in normal speech production: Evidence from a novel experimental paradigm. *Brain and Language*, *104*(1), 89–101. https://doi.org/10.1016/j.bandl.2007.03.007

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Humphreys, G. W., Lamote, C., & Lloyd-Jones, T. J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory*, *3*(3–4), 535–586. https://doi.org/10.1080/09658219508253164

Kello, C. T. (2004). Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 942–955. https://doi.org/10.1037/0096-1523.30.5.942

Kello, C. T., & Plaut, D. C. (2000). Strategic control in word reading: Evidence from speeded responding in the tempo-naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 719–750. https://doi.org/10.1037/0278-7393.26.3.719

Kello, C. T., & Plaut, D. C. (2003). Strategic control over rate of processing in word reading: A computational investigation. *Journal of Memory and Language*, *48*(1), 207–232. https://doi.org/10.1016/S0749-596X(02)00512-0

Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of stroop interference in speech production. *Journal of Experimental Psychology: General*, *129*(3), 340–360. https://doi.org/10.1037/0096-3445.129.3.340

Kello, C. T., Sibley, D. E., & Plaut, D. C. (2005). Dissociations in performance on novel versus irregular items: Single-route demonstrations with input gain in localist and distributed models. *Cognitive Science*, *29*(4), 627–654. https://doi.org/10.1207/s15516709cog0000_16

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lampe, L. F., Fieder, N., Krajenbrink, T., & Nickels, L. (2017). Semantische Nachbarschaft in der Wortproduktion bei Aphasie [Semantic neighbourhood in word production in aphasia]. In A. Adelt, Ö. Yetim, C. Otto, & T. Fritzsche (Eds.), *Spektrum Patholinguistik - Panorama Patholinguistik: Sprachwissenschaft trifft Sprachtherapie* (Vol. 10, pp. 103–114). Universitätsverlag Potsdam. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/39701

Lampe, L. F., Hameau, S., Bürki, A., Fieder, N., Sowman, P. F., & Nickels, L. (2019). *Investigating effects of semantic variables on word production.* Preregistration on the Open Science Framework.

osf.io/yw6ma

Lampe, L. F., Bürki, A., Sowman, P. F., Hameau, S., & Nickels, L. (2021). Effects of semantic variables on processes during word planning for production: Evidence from electrophysiological data. *Manuscript in Preparation*.

Lampe, L. F., Hameau, S., Fieder, N., & Nickels, L. (in press). Effects of semantic variables on word production in aphasia. *Cortex*.

Lampe, L. F., Hameau, S., & Nickels, L. (in press). Semantic variables both help and hinder word production: Behavioural evidence from picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means*. (R package version 1.5.2-1). https://cran.r-project.org/package=emmeans

Lloyd-Jones, T. J., & Nettlemill, M. (2007). Sources of error in picture naming under time pressure. *Memory & Cognition*, *35*(4), 816–836. https://doi.org/10.3758/BF03193317

Lüdecke, D. (2020). *sjPlot: Data visualization for statistics in social science*. (R package version 2.8.6). https://cran.r-project.org/package=sjPlot

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

Millard, S. P. (2013). *EnvStats: An R package for environmental statistics*. Springer.

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective and Behavioral Neuroscience*, *11*(1), 32–43. https://doi.org/10.3758/s13415-010-0009-7

Mirman, D., & Graziano, K. M. (2013). The neural basis of inhibitory effects of semantic and phonological neighbors in spoken word production. *Journal of Cognitive Neuroscience*, *25*(9), 1504–1516. https://doi.org/10.1162/jocn_a_00408

Mirman, D., Kittredge, A. K., & Dell, G. S. (2010). Effects of near and distant phonological neighbors on picture naming. *Proceedings of the Annual Meeting of the Cognitive Science Society, 32*, 1447–1452.

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 65–79. https://doi.org/10.1037/0278-7393.34.1.65

Mirman, D., Strauss, T. J., Brecher, A., Walker, G. M., Sobel, P., Dell, G. S., & Schwartz, M. F. (2010). A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function. *Cognitive Neuropsychology*, *27*(6), 495–504. https://doi.org/10.1080/02643294.2011.574112

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178. https://doi.org/10.1037/h0027366

Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 203–253). Academic Press Inc.

Moses, M. S., Nickels, L., & Sheard, C. (2004). "I'm sitting here feeling aphasic!" A study of recurrent perseverative errors elicited in unimpaired speakers. *Brain and Language*, *89*(1), 157–173. https://doi.org/10.1016/S0093-934X(03)00364-X

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). The neural consequences of semantic richness: When more comes to mind, less activation is observed. *Psychological Science*, *18*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2007.01913.x

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*(5), 377–500. https://doi.org/10.1080/02643299308253469

R Core Team. (2020). *R: A language and environment for statistical computing* (4.0.2). R Foundation for Statistical Computing. https://www.r-project.org/

Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, *132*(1), 68–89. https://doi.org/10.1016/j.cognition.2014.03.010

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, *146*, 240–244. https://doi.org/10.1016/j.cognition.2015.09.016

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2021). Semantic richness and density effects on language production: Electrophysiological and behavioral evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(3), 508–517. https://doi.org/10.1037/xlm0000940

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*(1), 205–235. https://doi.org/10.1037/0033-295X.111.1.205

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

RStudio Team. (2020). *RStudio: Integrated development for R* (1.3.959). RStudio, PBC. http://www.rstudio.com/

Starreveld, P. A., & La Heij, W. (1999). Word substitution errors in a speeded picture-word task. *The American Journal of Psychology*, *112*(4), 521–553. https://doi.org/https://doi.org/10.2307/1423649

Taylor, K. I., Devereux, B. J., Acres, K., Randall, B., & Tyler, L. K. (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. *Cognition*, *122*(3), 363–374. https://doi.org/10.1016/j.cognition.2011.11.001

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 664–

680. https://doi.org/10.1037/0278-7393.17.4.664

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 243–259. https://doi.org/10.1037/0278-7393.19.2.243

Appendices

Appendix A: Analyses after removal of items with lower average naming accuracy

Removal of 18 items with lower average naming accuracy led to analysis of data from 279 items and 80 participants. The naming latency analysis included 18,831 data points (standard naming: 10,443 data points, speeded naming: 8,388 data points) and the naming accuracy analysis included 22,168 data points (3,104 naming errors) (standard naming: 11,362 data points with 806 naming errors, speeded naming: 10,806 data points with 2,298 naming errors).

The findings were largely comparable to the analyses conducted on all 297 items. Importantly, the removal of items with lower naming accuracy items caused no differences in interactions between semantic variables and naming task compared to the analyses on all items, thus highlighting the appropriateness of including all items in the analyses reported in the main text of this paper. All differences compared to the analyses conducted on all items are reported below.

Naming latency

In the latency analysis excluding items that had relatively low overall naming accuracy (n = 18), the same semantic and control variables as reported in Table 1 (for all items) interacted with task, except for age of acquisition, which was non-significant in this analysis (Table A1).

Moreover, the main effects of age of acquisition and trial order were no longer significant compared to the analysis on all items. The remaining significant main effects were in the same directions as in the analysis including all items.

In the follow-up analysis, the simple effect of distinctiveness was non-significant in standard naming, in contrast to the analysis including all items. All other simple effects were comparable to the analysis on all items.

Table A1

Naming latency: summarised output of Linear Mixed Model analysis for task comparison and simple effects after removal of items with lower naming accuracy

Model structure	lmer(RT ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + OrdCatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness) * Task + (1 | Item) + (NearSemNeigh + SemSim || Participant), data, REML = TRUE)					
Random effect	Variance	*SD*				
Item (Intercept)	0.01	0.12				
Participant (Intercept)	0.02	0.13				
Participant | NrSemNeigh	0.00	0.02				
Participant | SemSim	0.00	0.01				
Residuals	0.06	0.24				
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF
(Intercept)	-1.55	0.02	-1.58 – -1.52	**-95.26**	**<.001**	
NameAgr	-0.04	0.01	-0.05 – -0.03	**-5.37**	**<.001**	1.10
AoA	0.02	0.01	-0.00 – 0.04	1.52	.128	2.45
Imageability	-0.01	0.01	-0.03 – 0.01	-1.12	.263	1.91
ImageAgr	-0.03	0.01	-0.05 – -0.02	**-3.97**	**<.001**	1.26
Frequency	-0.01	0.01	-0.03 – 0.01	-0.95	.344	1.57
Familiarity	-0.04	0.01	-0.05 – -0.02	**-3.78**	**<.001**	1.73
Order	-0.01	0.00	-0.01 – 0.00	-1.88	.060	3.07
OrdCatPos	0.00	0.00	-0.01 – 0.01	0.23	.817	3.14
NoFeats	-0.01	0.01	-0.02 – 0.01	-0.66	.513	1.60
IntercorrDens	0.02	0.01	-0.00 – 0.04	1.71	.087	2.18
NearSemNeigh	0.02	0.01	0.00 – 0.05	**2.22**	**.026**	2.31
SemSim	0.01	0.01	-0.01 – 0.03	0.90	.367	2.40
Typicality	-0.01	0.01	-0.02 – 0.01	-0.69	.490	1.41

Simple effects of variables in significant interactions

Distinctiveness	0.03	0.01	0.01 – 0.05	**2.88**	**.004**	2.02	Speeded naming				Standard naming			
Task	-0.32	0.03	-0.38 – -0.26	**-10.96**	**<.001**	1.00	Estimate	*SE*	*z*-value	*p*-value	Estimate	*SE*	*z*-value	*p*-value
NameAgr * Task	-0.01	0.00	-0.02 – -0.01	**-3.33**	**.001**	1.09	-0.05	0.01	**-5.97**	**<.001**	-0.03	0.01	**-4.42**	**<.001**
AoA * Task	0.01	0.01	-0.00 – 0.02	1.76	.079	2.36								
Imageability * Task	0.01	0.00	0.00 – 0.02	**2.23**	**.025**	1.86	-0.01	0.01	-0.55	.584	-0.02	0.01	-1.64	.102
ImageAgr * Task	-0.01	0.00	-0.02 – -0.00	**-2.32**	**.020**	1.19	-0.04	0.01	**-4.38**	**<.001**	-0.03	0.01	**-3.32**	**.001**
Frequency * Task	-0.03	0.00	-0.04 – -0.02	**-6.67**	**<.001**	1.53	-0.02	0.01	**-2.52**	**.012**	0.01	0.01	0.705	.481
Familiarity * Task	-0.02	0.00	-0.03 – -0.01	**-3.33**	**.001**	1.63	-0.04	0.01	**-4.44**	**<.001**	-0.03	0.01	**-2.88**	**.004**
Order * Task	-0.03	0.01	-0.04 – -0.02	**-5.49**	**<.001**	2.24	-0.02	0.00	**-4.51**	**<.001**	0.01	0.00	1.70	.089
OrdCatPos * Task	-0.01	0.01	-0.02 – 0.00	-1.47	.142	2.70								
NoFeats * Task	-0.01	0.00	-0.02 – 0.00	-1.60	.109	1.65								
IntercorrDens * Task	0.00	0.01	-0.01 – 0.01	0.51	.609	2.02								
NearSemNeigh * Task	0.01	0.01	-0.00 – 0.02	1.49	.137	1.55								
SemSim * Task	0.01	0.01	-0.00 – 0.02	1.86	.062	1.70								
Typicality * Task	-0.00	0.00	-0.01 – 0.01	-0.83	.405	1.46								
Distinctiveness * Task	0.02	0.01	0.01 – 0.03	**4.20**	**<.001**	1.67	0.41	0.01	**3.78**	**<.001**	0.02	0.01	1.80	.072
Observations: 18,831														
Marginal R^2 / Conditional R^2: 0.280 / 0.527														

Note. Participant | X = random slope of X by participants, VIF = Variance Inflation Factor, NameAgr = Name agreement, AoA = Age of acquisition, ImageAgr = Image agreement, Order = Trial order, OrdCatPos = Ordinal category position, NoFeats = Number of semantic features, IntercorrDens = Intercorrelational Density, NearSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity.

Values of significant effects ($p < .05$) are printed in bold; in the simple effects, the variable with the stronger effect within significant interactions is highlighted in grey.

Naming accuracy

The same variables as reported in Table 2 (for all items) interacted with task in the accuracy analysis (Table A2). In addition, the interaction between frequency and task reached significance, with a stronger effect of frequency in the speeded naming task.

Moreover, there were some changes to the main effects: the effects of age of acquisition, frequency, and intercorrelational density were no longer significant compared to the analysis on all items. Instead, effects of familiarity and number of near semantic neighbours reached significance. In the follow-up analysis, all simple effects were comparable to the analysis that included all items.

Table A2

Naming accuracy: summarised output of Generalised Linear Mixed Model analysis for task comparison and simple effects after removal of items with lower naming accuracy

Model structure	glmer(ACC ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + CatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness)*Task + (1 | Item) + (SemSim | Participant), data, family = binomial())					
Random effect	Variance	*SD*	Correlation			
Item (Intercept)	0.81	0.90				
Participant (Intercept)	0.36	0.60				
Participant | SemSim	0.02	0.15	0.73			
Fixed effects	Estimate	*SE*	*CI*	*z*-value	*p*-value	VIF
(Intercept)	2.55	0.09	2.37 – 2.73	**27.56**	**<.001**	
NameAgr	0.47	0.06	0.35 – 0.59	**7.56**	**<.001**	1.11
AoA	-0.14	0.1	-0.33 – 0.04	-1.50	.135	2.59
Imageability	0.22	0.08	0.05 – 0.38	**2.61**	**.009**	2.00
ImageAgr	0.22	0.07	0.09 – 0.36	**3.26**	**.001**	1.31
Frequency	0.14	0.08	-0.01 – 0.29	1.88	.060	1.60
Familiarity	0.18	0.08	0.02 – 0.34	**2.26**	**.024**	1.79
Order	-0.05	0.05	-0.14 – 0.04	-1.10	.274	3.21
OrdCatPos	-0.03	0.05	-0.13 – 0.08	-0.50	.616	3.43
NoFeats	0.12	0.08	-0.03 – 0.27	1.60	.109	1.67
IntercorrDens	-0.17	0.09	-0.34 – 0.01	-1.90	.058	2.26
NearSemNeigh	-0.19	0.09	-0.37 – -0.00	**-2.01**	**.045**	2.47
SemSim	0.02	0.1	-0.17 – 0.22	0.25	.806	2.49
Typicality	0.06	0.07	-0.09 – 0.20	0.77	.439	1.53

Simple effects of variables in significant interactions

							Speeded naming				Standard naming			
Distinctiveness	-0.05	0.09	-0.22 – 0.12	-0.59	.558	2.08								
Task	-1.62	0.15	-1.90 – -1.33	**-11.13**	**<.001**	1.22	Estimate	*SE*	*z*-value	*p*-value	Estimate	*SE*	*z*-value	*p*-value
NameAgr * Task	-0.15	0.04	-0.24 – -0.07	**-3.47**	**.001**	1.11	0.39	0.06	**6.21**	**<.001**	0.542	0.07	**7.96**	**<.001**
AoA * Task	0.09	0.08	-0.06 – 0.24	1.19	.233	2.88								
Imageability * Task	-0.01	0.06	-0.14 – 0.11	-0.18	.855	2.17								
ImageAgr * Task	-0.16	0.05	-0.26 – -0.06	**-3.16**	**.002**	1.37	0.14	0.07	**2.04**	**.042**	0.30	0.08	**3.98**	**<.001**
Frequency * Task	0.12	0.06	0.01 – 0.23	**2.12**	**.034**	1.56	0.20	0.08	**2.63**	**.009**	0.08	0.09	0.97	.331
Familiarity * Task	0.21	0.06	0.08 – 0.33	**3.18**	**.001**	1.87	0.28	0.08	**3.51**	**.001**	0.08	0.09	0.86	.391
Order * Task	-0.18	0.07	-0.32 – -0.04	**-2.59**	**.010**	2.46	-0.14	0.05	**-2.86**	**.004**	0.04	0.07	0.61	.541
OrdCatPos * Task	-0.00	0.08	-0.16 – 0.15	-0.06	.949	3.02								
NoFeats * Task	-0.19	0.06	-0.32 – -0.07	**-3.11**	**.002**	1.84	0.03	0.08	0.35	.730	0.22	0.09	**2.50**	**.012**
IntercorrDens * Task	0.10	0.07	-0.04 – 0.23	1.39	.164	2.36								
NearSemNeigh * Task	0.09	0.07	-0.05 – 0.22	1.24	.216	2.65								
SemSim * Task	-0.09	0.09	-0.26 – 0.08	-1.00	.316	2.71								
Typicality * Task	-0.04	0.06	-0.17 – 0.09	-0.63	.530	1.87								
Distinctiveness * Task	0.02	0.07	-0.11 – 0.16	0.35	.727	2.13								
Observations: 22,168														
Marginal R^2 / Conditional R^2: 0.243 / 0.443														

Note. Participant | X = random slope of X by participants, VIF = Variance Inflation Factor, NameAgr = Name agreement, AoA = Age of acquisition, ImageAgr = Image agreement, Order = Trial order, OrdCatPos = Ordinal category position, NoFeats = Number of semantic features, IntercorrDens = Intercorrelational Density, NearSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity.

Values of significant effects (p < .05) are printed in bold; in the simple effects, the variable with the stronger effect within significant interactions is highlighted in grey.

Appendix B: Semantic error versus correct responses analysis

In addition to the overall naming accuracy analysis, we also conducted an analysis that contrasted the proportions of responses that were semantic errors rather than correct responses. While increased input gain may lead to stronger effects of semantic variables on naming accuracy in speeded naming, one could also expect an increase in their effects on semantic errors more specifically, as semantic variables have been found to affect the proportion of semantic errors in speeded naming (i.e., Fieder et al., 2019; Mirman, 2011), however, have not previously been examined in standard naming in neurotypical participants. Thus, if the influence of semantic variables is increased in speeded naming one could also expect stronger effects of semantic variables on semantic errors.

We coded the following error types as semantic errors: associates, coordinates, subordinates, superordinates, synonyms, incomplete responses that shared at least 50% of their phonemes with a semantically related item (or vice-versa, e.g., *oct* for *squid*), responses with a part-whole relationship to the target, two-step errors involving a semantic and a phonological error (e.g., naming a picture of a lion *riger*, presumably via *tiger*), and other responses that had a clear semantic relationship to the target but were not categorically related (e.g., *butterfly* for *bird*).

The semantic errors analysis included 22,540 data points from 297 items and 80 participants (*n* = 2,752 semantic errors). 11,800 of those data points came from the standard naming task (*n* = 797 semantic errors) and 10,740 data points from the speeded naming task (*n* = 1,955 semantic errors).

There was a main effect of task in this semantic error analysis indicating that there were more semantic errors overall in the speeded naming task than in the standard naming task. Of all responses, on average 17% were semantic errors in the speeded naming task (range = 6–29%, *SD* = 6%) and 7% in the standard naming task (range = 1–15%, *SD* = 4%). The proportion of semantic errors relative to all other error types was, however, relatively stable in the two tasks (speeded naming: 70%, range = 47–88%, *SD* = 10%; standard naming: 69%, range = 35–92%, *SD* = 15%).

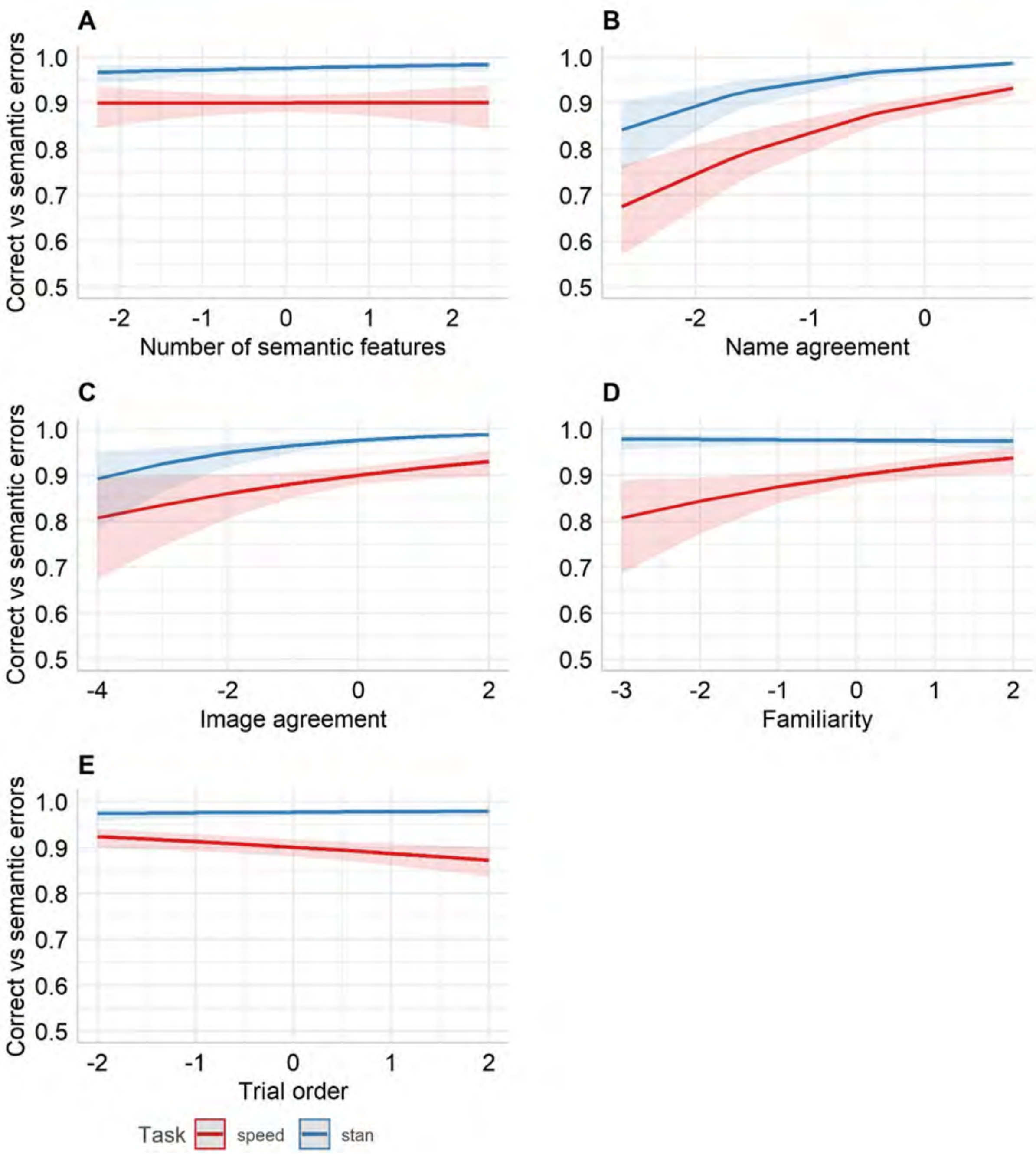
All findings in the semantic errors analysis were largely comparable to the findings of the naming accuracy analysis (which contrasted correct responses with any other error response), presumably because most incorrect responses in both tasks were semantic errors. As in the overall naming accuracy analysis, the only semantic variable that significantly interacted with naming task was number of semantic features (see Table B1). The follow-up analysis suggested that the effect was stronger in the standard naming task than in the speeded naming task (Figure B1, Panel A), however, note that there was in fact no evidence for a simple effect in either task. This indicates that there was either a Type 1 error when finding the significant interaction between number of semantic features and task, or a Type 2 error with non-significant simple effects in one or both naming tasks.

Moreover, the same control variables significantly interacted with task as in the naming accuracy analysis: The effects of name agreement (facilitatory in both tasks, Figure B1, Panel A) and image agreement (facilitatory in both tasks, Figure B1, Panel D) were stronger in the standard naming task, while effects of familiarity (facilitatory in speeded naming, Figure B1, Panel F) and trial order (inhibitory in speeded naming, Figure B1, Panel G) were stronger in the speeded naming task.

With the combined data from both tasks, main effects showed that semantic errors were more likely for words with higher intercorrelational density and responses were more accurate for words with higher name agreement, imageability, or image agreement.

Figure B1

Significant interactions between semantic or control variables and task in the semantic errors analysis



Note. All variables were standardised.

Table B1

Semantic errors: summarised output of Generalised Linear Mixed Model analysis for task comparison and simple effects

Model structure	glmer(SEM ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + CatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness)*Task + (1 | Item) + (NearSemNeigh | Participant), data, family = binomial())					
Random effect	Variance	*SD*	Correlation			
Item (Intercept)	1.51	1.23				
Participant (Intercept)	0.34	0.58				
Participant | NearSemNeigh	0.03	0.19	0.72			
Fixed effects	Estimate	*SE*	*CI*	*z*-value	*p*-value	VIF
(Intercept)	2.97	0.11	2.77 – 3.18	28.01	<.001	
NameAgr	0.67	0.08	0.51 – 0.83	**8.32**	**<.001**	1.10
AoA	-0.22	0.12	-0.46 – 0.03	-1.74	.082	2.54
Imageability	0.22	0.11	0.01 – 0.43	**2.09**	**.037**	1.93
ImageAgr	0.30	0.09	0.13 – 0.47	**3.39**	**.001**	1.29
Frequency	0.12	0.10	-0.08 – 0.31	1.16	.244	1.63
Familiarity	0.11	0.11	-0.10 – 0.32	1.03	.303	1.80
Order	-0.04	0.05	-0.14 – 0.05	-0.88	.381	3.23
OrdCatPos	-0.04	0.06	-0.16 – 0.07	-0.76	.448	3.41
NoFeats	0.08	0.10	-0.11 – 0.27	0.81	.421	1.60
IntercorrDens	-0.26	0.12	-0.49 – -0.03	**-2.25**	**.025**	2.26
NearSemNeigh	-0.14	0.13	-0.38 – 0.11	-1.10	.272	2.45
SemSim	-0.02	0.13	-0.27 – 0.23	-0.18	.860	2.53
Typicality	0.05	0.10	-0.14 – 0.24	0.52	.605	1.50
Distinctiveness	0.00	0.11	-0.22 – 0.22	0.00	.999	2.06

Simple effects of variables in significant interactions	
Speeded naming	Standard naming

Task	-1.54	0.14	-1.82 – -1.25	**-10.65**	**<.001**	1.29	Estimate	*SE*	*z*-value	*p*-value	Estimate	*SE*	*z*-value	*p*-value
NameAgr * Task	-0.22	0.05	-0.31 – -0.13	**-4.74**	**<.001**	1.11	0.56	0.08	**6.86**	**<.001**	0.78	0.09	**9.07**	**<.001**
AoA * Task	0.14	0.08	-0.02 – 0.30	1.77	.078	2.67								
Imageability * Task	0.04	0.07	-0.09 – 0.17	0.56	.575	1.98								
ImageAgr * Task	-0.21	0.05	-0.32 – -0.11	**-3.93**	**<.001**	1.31	0.19	0.09	**2.15**	**.031**	0.41	0.10	**4.27**	**<.001**
Frequency * Task	0.11	0.06	-0.02 – 0.23	1.70	.090	1.59								
Familiarity * Task	0.30	0.07	0.16 – 0.44	**4.23**	**<.001**	1.93	0.26	0.11	**2.43**	**.015**	-0.04	0.12	-0.35	.728
Order * Task	-0.20	0.07	-0.34 – -0.05	**-2.66**	**.008**	2.34	-0.14	0.06	**-2.58**	**.010**	0.05	0.07	0.78	.435
OrdCatPos * Task	0.03	0.08	-0.13 – 0.19	0.41	.680	2.95								
NoFeats * Task	-0.15	0.07	-0.28 – -0.02	**-2.29**	**.022**	1.73	0.00	0.10	0.04	.965	0.16	0.11	1.43	.153
IntercorrDens * Task	0.13	0.07	-0.00 – 0.26	1.94	.053	2.21								
NearSemNeigh * Task	-0.02	0.09	-0.19 – 0.15	-0.23	.819	2.48								
SemSim * Task	-0.06	0.08	-0.22 – 0.11	-0.69	.493	2.53								
Typicality * Task	0.01	0.07	-0.12 – 0.14	0.09	.930	1.69								
Distinctiveness * Task	0.06	0.08	-0.09 – 0.21	0.74	.462	2.08								

Observations: 22,540

Marginal R^2 / Conditional R^2: 0.250 / 0.523

Note. Participant | X = random slope of X by participants, VIF = Variance Inflation Factor, NameAgr = Name agreement, AoA = Age of acquisition, ImageAgr = Image agreement, Order = Trial order, OrdCatPos = Ordinal category position, NoFeats = Number of semantic features, IntercorrDens = Intercorrelational Density, NearSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity.

Values of significant effects (p < .05) are printed in bold; in the simple effects, the variable with the stronger effect within significant interactions is highlighted in grey.

Acknowledgements

The authors would like to thank Serje Robidoux for statistical advice and Maria Zarifyan for her help with data preprocessing.

CHAPTER 6

General Discussion

Knowledge about the meaning of words is central to successful verbal communication and is the starting point of all current models of oral word production. Consequently, the structure of word knowledge, its complexity, and the way activation spreads to semantically related representations may have consequences for processing at other levels of word planning. This may be particularly true for the lexical level where aspects of the semantic representation of the target word can influence both the activation of the target's lexical representation and the size and strength of activation of a cohort of co-activated lexical representations. Some of these aspects of meaning can be operationalised as semantic variables.

However, there is relatively little previous research into effects of semantic variables on word production and that which there is often focused on only a few aspects of semantics thus dismissing the fact that meaning is multidimensional with probable simultaneous effects from different semantic variables (Pexman et al., 2013; Taylor et al., 2012). Yet, a clearer understanding of effects of semantic variables on word production can inform theories of word production as their effects allow us to better understand semantic and lexical processing.

Consequently, the research presented in this thesis was important as it moved beyond the one-dimensional approach to semantic variables often adopted by previous research and contributed well-controlled studies to the evidence base of effects of semantic variables on word production. I focused on, and thoroughly investigated, effects of the feature-based semantic variables, number of semantic features, intercorrelational density, number of near semantic neighbours, semantic similarity, typicality, and distinctiveness. The studies in this thesis aimed to determine *which* feature-based semantic variables reliably influence behaviour and to better understand *how* these variables affect processing in the context of word production models. To achieve these aims, I followed a multifaceted approach that capitalised on different populations (people with aphasia in Paper 1, neurotypical speakers in Papers 2, 3, and 4), different types of data (behavioural data in Papers 1, 2, and 4, electrophysiological data in Paper 3), and different experimental paradigms (standard picture naming in Papers 1, 2, 3, and 4, speeded picture naming in Paper 4). The research presented in this thesis has yielded important findings that add to our understanding of the mechanisms underlying effects of semantic variables in word production. These findings are summarised and discussed in the next section. Subsequently, I address methodological strengths and challenges and theoretical implications of this work. Possible avenues for future research are indicated throughout this General Discussion.

Summary and discussion of the main experimental findings

Table 1 summarises the main findings of the four studies regarding effects of the semantic variables. I will now summarise and discuss the most important findings of the four experimental studies of this thesis, structured around the research questions posed by the different studies. Following this, I will assess why different semantic variables influenced performance in different studies.

Which semantic variables affect word production in people with and without aphasia and what are the underlying mechanisms?

Papers 1, 2, and 4 investigated effects of the semantic variables on behavioural measures (naming speed and/or accuracy and error types) in participants with aphasia (Paper 1) and neurotypical speakers (Papers 2 and 4). As is clear from Table 1, the naming accuracy of the full group of participants with aphasia was unaffected by the semantic variables (but there was a random slope for number of semantic features by participants that significantly improved model fit), while participants with predominantly semantic and/or lexical impairments of the sub-group were more accurate on words with a higher number of semantic features. Additional analyses into effects of single semantic variables in the full group of participants also revealed a significant facilitatory effect of number of semantic features. Importantly, a complementary Bayesian analysis also allowed me to interpret the null findings for the other semantic variables: Positive evidence in favour of the null hypothesis for all semantic variables, except for number of semantic features, indicated that they did not influence naming accuracy of the full group of participants and that the absence of effects was not simply due to low statistical power (e.g., Dienes, 2014). Moreover, across participant groups, words with higher numbers of semantic features, higher semantic similarity, and lower typicality were more likely to be produced correctly rather than resulting in a semantic error, coordinate error, or omission, and some of these effects depended on the integrity of semantic processing.

Table 1

Summary of the findings of all the studies reported in this thesis

	Participants with aphasia			Neurotypical speakers		
	Paper 1: MAPP Database			Paper 2: standard naming [b]	Paper 3: ERPs [b]	Paper 4: naming speed
Semantic variable	Full group	Sub-group	Naming errors [a]	Latency/Accuracy		Latency/Accuracy
Number of semantic features	Ø	↗	↗	↗ / ↗	✓	Ø / ✓
Intercorrelational density	Ø	Ø	Ø	Ø / ↙	✓	Ø / Ø
Semantic similarity	Ø	Ø	↗	Ø / Ø	✓	Ø / Ø
Number of near semantic neighbours	Ø	Ø	Ø	Ø / Ø	✓	Ø / Ø
Typicality	Ø	Ø	↙	Ø / Ø	Ø	Ø / Ø
Distinctiveness	Ø	Ø	Ø	↙ / Ø	Ø	✓ / Ø

Note. ERPs = Event Related Potentials, Ø = non-significant effect, ↙ = poorer performance (slower responses or decreased accuracy with higher values of the semantic variable), ↗ = improved performance (faster responses and higher accuracy with higher values of the semantic variable), ✓ = effect of semantic variable on ERPs (Paper 3) or significant interaction with naming task (Paper 4).

[a] For naming errors, the direction of the arrow indicates the change in correct responses (increased vs decreased) relative to the error type, with higher values of the variable; findings are combined for the full group and the sub-group and across the analyses examining correct responses versus semantic errors, correct responses versus coordinate errors, or correct responses versus omissions; Ø indicates non-significant effects across all comparisons and ↙/↗ a significant effect in at least one of the analyses.

[b] Findings from the most complex analysis including all six semantic variables and all control variables.

In Paper 2, I investigated effects of the same semantic variables on neurotypical speakers. Similar to the participants with aphasia in Paper 1, naming responses of neurotypical participants were faster and more accurate for words with higher numbers of semantic features. In contrast, responses were less accurate for words with higher intercorrelational density and slower for words with higher distinctiveness. Paper 4, in which I compared speeded and standard naming, revealed that the effects of some semantic variables differed between the two tasks: While the effect of number of semantic features on naming accuracy was stronger in standard naming, and non-significant in speeded naming, the effect of distinctiveness on naming latency was stronger in speeded naming, but significant in both tasks.

Taken together, across analyses in Papers 1, 2 and 4, number of semantic features most reliably affected naming performance and did so in both participants with and without aphasia, with additional effects of intercorrelational density, semantic similarity, typicality, and distinctiveness across the populations. While higher values of some of the variables (i.e., number of semantic features and semantic similarity) facilitated processing, leading to faster and/or more accurate responses, higher values of other variables (i.e., intercorrelational density, typicality, and distinctiveness) inhibited performance, leading to slower or more inaccurate responses. Hence, depending on the semantic variable, the aspect of the structure of the semantic representation and its consequences for the activation environment of lexical processing it captured either *supported* or *hindered* processing of the target word. This pattern of semantic facilitation and inhibition needs to be explained by theories of word production. Importantly, most current theories were formulated for experimental data from context manipulation paradigms, however, the mechanisms proposed to account for facilitatory and inhibitory effects from distractor words or previously named words in the experiment may also be able to account for effects of item-inherent variables. Possible cognitive mechanisms for the effects of the significant semantic variables have already been discussed in the experimental chapters of this thesis. The next two sections will briefly explain the repertoire of mechanisms available to current models of word production to explain the findings. To foreshadow the argument, facilitatory effects likely require some mechanism of activation spreading which primes the target word, while inhibitory effects can most parsimoniously be explained when assuming a mechanism of lexical competition.

Mechanisms of semantic facilitation

As detailed in the General Introduction, few of the currently prevalent theories of word production explicitly assume a mechanism that allows for semantic facilitation (i.e., Abdel Rahman & Melinger, 2009; Navarrete et al., 2014). When semantic facilitation has been located in the model, it is placed at the level of semantic processing. For example, the Swinging Lexical Network Hypothesis (Abdel Rahman & Melinger, 2009) proposes holistic lexical concept nodes with activation spreading bidirectionally between them, which causes conceptual priming of the target concept. However, this was first suggested in the context of the Picture-Word Interference paradigm, where participants receive two stimuli: a to-be-named picture and a to-be-ignored distractor. In this paradigm, the target picture (e.g., 'cat') activates related concepts via spreading activation, for example 'animal' and 'pet' via *is a* links, 'fur' and 'tail' via *has a* links, and these pass activation on to related concepts that share (some of) these properties, causing concepts like 'dog', 'rabbit', 'mouse', 'horse', etc. to be co-activated. Simultaneously, a categorically related distractor word (e.g., 'dog') spreads activation that converges on the same concepts. The activated lexical concepts are suggested to enhance each other's activation through their mutual bidirectional links, thus priming the target word. Importantly, even in the absence of a distractor word, such as in standard picture naming, Abdel Rahman & Melinger (2019) proposed that increased spread of activation (e.g., due to many semantic features associated with a target) has the same facilitatory effect on target processing: Bidirectional spread of activation through the semantic network 'primes' the target concept and enhances the activation of its lexical representation. More specifically, when semantic activation spreads between lexical concepts (e.g., from 'cat' to 'animal', 'pet', 'fur', and 'tail', and from those concepts to 'dog', 'rabbit', 'mouse', and 'horse'), each of them sends activation back to the target concept (and other related lexical concepts) via the shared property concept nodes. The more related lexical concepts there are, the more feedback activation converges on the target concept, facilitating its selection.

As demonstrated by the presence of facilitatory effects of semantic variables other than number of semantic features, the extent of the spread of activation cannot exclusively depend on the number of lexical concepts a target concept is connected to. Instead, it is also important whether the target lexical concept is part of a cluster of similar representations (e.g., Collins & Loftus, 1975), which may, for example, be represented as the average number of links a target concept shares with all other concepts in the semantic network (effect of semantic similarity in Paper 1). However, further research is needed that specifically targets the mechanism of spreading activation and establishes the way *feature-based* semantic variables may operate assuming holistic semantic representations.

In addition to the Swinging Lexical Network Hypothesis (Abdel Rahman & Melinger, 2009, 2019), the Ballistic Model (Mahon & Navarrete, 2016; Navarrete et al., 2014) assumes semantic priming via spreading activation. While this theory does not account for standard picture naming (priming via spreading activation was suggested for the Blocked Cyclic Naming Paradigm), the mechanism could presumably work as described for the Swinging Lexical Network Hypothesis. Similarly, this spreading activation account to explain facilitatory effects of semantic variables could work in WEAVER++ (Levelt et al., 1999) and Howard et al. (2006), where a locus and mechanism of semantic facilitation has not been specified. However, when the labelled links between the target's lexical concept and other concepts are assumed to be bidirectional (note that in WEAVER++ they are unidirectional), the mechanism of semantic facilitation could work as outlined above.

Alternatively, if semantic representations are indeed feature-based, clearly the presence of many semantic features may result in strong activation of the target's lexical representation. Additionally, lexical to conceptual feedback may enhance the activation of the target word: If the semantic features of 'cat' co-activate the lexical representations of 'dog' and 'tiger' (see Figure 1 in the General Introduction), feedback from these lexical representations to dog's and tiger's semantic features would converge on the semantic features of 'cat', enhancing its activation and thus facilitating production of 'cat'. This proposal applies to the Interactive Activation model (Dell, 1986, 1988; Dell et al., 1997). On the other hand, or in addition to the described feedback account, bidirectional feature-feature connections (e.g., between correlated features, e.g., Cree et al., 1999; McRae et al., 1997, 1999) may allow for activation to spread between semantic features at the semantic level. Similar to the spreading activation account for holistic lexical concepts, this may ultimately enhance the activation of the target's semantic features. However, none of the current word production theories with feature-based semantic representations proposes such connections.

Finally, the Incremental Learning Model (Oppenheim et al., 2010) does not include a mechanism for semantic facilitation, but incremental strengthening after lexical selection could be understood as long-term facilitation. According to the Incremental Learning Model, each act of lexical retrieval results in *persistent* learning, which causes an adjustment of the weights of the connections between semantic and lexical representations (i.e., repetition priming or incremental weakening of connections). However, this was proposed in the context of the Cumulative Semantic Interference effect, which develops across an experiment. In contrast, it is unclear how long-term strengthening of semantic-to-lexical level connections after successful lexical selection would develop for item-inherent variables. Future research could address how semantic facilitation may work assuming long-term adjustments of semantic-to-lexical connections in this model.

Mechanisms of semantic interference

In addition to facilitatory effects, semantic variables also resulted in inhibitory effects on naming. Hence, even in the absence of contextual manipulations, theories of word production must provide an account of inhibitory effects. The prevalent account is lexical selection by competition, where co-activated semantic representations compete with the target word for selection (different implementations of the competitive selection process, e.g., Luce choice rule (Luce, 1959), were described in the General Introduction). A number of current theories of word production assume lexical selection to be competitive, however, the Swinging Lexical Network Hypothesis (Abdel Rahman & Melinger, 2009, 2019) is the only theory that explicitly accounts for inhibitory effects of item-inherent variables. It proposes that semantic variables can impact the size and strength of activation of a cohort of co-activated lexical representation, which will compete with the target for selection. Importantly, as lexical representations are activated from their respective semantic representations, in a holistic semantic organisation, the co-activated lexical cohort depends on the spread of activation between lexical concepts at the semantic level. Hence, the inhibitory effects of intercorrelational density, typicality, and distinctiveness observed in Papers 1, 2, and 4 likely originate at the semantic level and have their loci at the lexical level.

Even though most other theories of word production that address effects of inhibition (e.g., Howard et al., 2006; Levelt et al., 1999) were proposed to account for inhibitory effects of semantic distractors or previously named items in the experiment, inhibitory effects of semantic variables can also be explained in these theories with only slight modification—specifically, by complementing lexical competition with one of the mechanisms outlined in the previous section (i.e., spreading activation at the semantic level or feedback between the lexical and semantic levels).

Alternatively, long-term adjustments of semantic-to-lexical connections (Navarrete et al., 2014; Oppenheim et al., 2010) may also be able to account for inhibitory effects of item-inherent semantic variables. If the semantic representation of words causes other lexical representations to be repeatedly co-activated but unselected during processing, the semantic-to-lexical connection for these co-activated words would be increasingly weakened. For example, in Paper 2, I described how that might be the case for words with higher intercorrelational density that are co-activated more often over the course of a lifetime, without being named. If the changes in their connections between semantic and lexical representations are persistent, words with long-term weakened connections would be harder to process if they become target words themselves. Consequently, even though the Incremental Learning Model (Oppenheim et al., 2010) and the Ballistic Model of Lexical Access (Navarrete et al., 2014) do not discuss influences from item-inherent variables, they may be able to account for inhibitory effects of the semantic variables when assuming a mechanism that permanently changes the way words are accessed after repeated weaking of semantic-to-lexical connections.

Of the currently available theories of word production reviewed in the General Introduction, only Dell's Interactive Activation Model (Dell, 1986, 1988; Dell et al., 1997) seems unable to account for inhibitory effects of semantic variables as it does not implement any mechanism of lexical competition or learning. While initially activated semantic features and feedback between the lexical and semantic levels may activate multiple lexical candidates in Dell's model architecture, they would not interfere with selection as the highest activated lexical candidate is selected after a certain number of time steps, irrespective of the activation levels of co-activated lexical representations.

In sum, the Swinging Lexical Network Hypothesis (Abdel Rahman & Melinger, 2009, 2019) seems to be the only theory of word production that is able to account for effects of item-inherent semantic variables without further modifications. However, with only minor adjustments, most of the other theories would also be able to explain such effects.

Differences in effects of semantic variables between participants with and without aphasia

In Papers 1, 2, and 4, I discussed effects of semantic variables on behavioural measures of word production separately for participants with aphasia and neurotypical speakers. However, here, I merge the findings from these papers and discuss the differences between effects in participants with and without aphasia.

As becomes evident from Table 1, number of semantic features was the only semantic variable that affected the naming performance of participants with and without aphasia: It was the only semantic variable that affected both naming latency and accuracy of neurotypical participants in standard picture naming (Paper 2), its effect on naming accuracy differed between speeded and standard naming with a stronger facilitatory effect on standard naming (Paper 4), and number of semantic features was also significant across analyses in the Event Related Potential (ERP) data (Paper 3). In addition, this was the only semantic variable that affected naming accuracy in participants with semantic and/or lexical impairments (Paper 1), suggesting that in participants with aphasia this effect of number of semantic features resembled 'normal-like' processing. Both participants with and without language impairments seem to benefit from the availability of increased semantic information for a concept (i.e., many semantic features).

In addition to number of semantic features, the performance of participants with aphasia was affected by typicality and semantic similarity, while neurotypical participants showed effects of intercorrelational density and distinctiveness in standard naming (Paper 2), and the effect of distinctiveness was stronger in speeded than in standard naming (Paper 4). What could be the reasons for such discrepancies in the findings for participants with and without aphasia? In participants with aphasia, variables that influenced performance, with the exception of number of semantic features, showed their effects in the error analyses and some of these effects interacted with the participants' semantic abilities, only becoming apparent for a particular group of participants (e.g., the effect of semantic similarity was only present for participants with severe semantic impairments). In contrast, the effect of distinctiveness was only significant, or differed between tasks, in the response latency analysis for neurotypical participants (Papers 2 and 4). Hence, it is possible that error analyses (participants with aphasia, Paper 1) as well as response latency analyses (neurotypical participants, Papers 2 and 4) have a different sensitivity to semantic variables compared to overall naming accuracy analyses. Response latency data in unimpaired participants may provide a more nuanced representation of the influences of the variables of interest compared to the naming accuracy and error type analyses as it can capture even slight differences in the processing trajectory induced by the semantic variables, which may, in neurotypical participants, not result in a naming error (also note that, surprisingly, not even in the comparison between speeded and standard naming were naming accuracy and production of semantic errors strongly affected by the semantic variables).

In contrast, intercorrelational density was the only semantic variable that significantly affected naming accuracy of neurotypical participants (Paper 2) but not participants with aphasia. One might think that, in participants with aphasia, the fine-grained structure of intercorrelations between semantic features could be impaired, which would reduce the spread of activation between semantic features and lead to fewer lexical representations being co-activated and thus reduced lexical competition. As a result, the performance of participants with aphasia may be unaffected by this variable. However, this account would not apply to people without semantic impairments. It is consequently unclear why the effect of intercorrelational density is present in neurotypical participants but not in people with aphasia. More research into this matter is needed.

Further reasons for the inconsistencies in findings of the participants with and without aphasia may be in the slight differences of control variables included in the analyses: While Paper 1 included control variables that have been shown to influence performance of participants with aphasia, the papers with neurotypical participants (Papers 2–4) controlled for the variables that were identified in a meta-analysis by Perret and Bonin (2019), which found a slightly different set of variables to consistently affect performance of neurotypical than those controlled in Paper 1. Finally, differences in the number of items included in the papers on participants with and without aphasia may have affected the results: While the overall number of data points included in the analyses was high in all studies (Paper 1: 15,573 data points from 89 items and 175 participants (full group), Paper 2: 24,554 data points from 291 items and 85 participants, Paper 4: 23,608 data points from 297 items and 80 participants (combined data)), the smaller number of items included in Paper 1 may have resulted in a decrease of statistical power. Moreover, the range of values of the semantic variables differed slightly between the papers, with wider ranges in Paper 2 than in Paper 1 (Table 2).

Table 2

Descriptive statistics for the semantic variables included in Papers 1 and 2

	Paper 1: Participants with aphasia (*n* = 89 items)			Paper 2: Neurotypical participants (*n* = 291 items)		
Semantic variable	Mean	*SD*	Range	Mean	*SD*	Range
Number of semantic features	13.42	2.52	7–20	12.73	3.01	6–20
Intercorrelational density	130.33	136.66	0.00–739.15	156.07	172.86	0.00–1296.22
Semantic similarity	0.03	0.02	0.00–0.06	0.04	0.02	0.00–0.09
Number of near semantic neighbours	3.94	4.99	0–24	6.19	7.59	0–38
Typicality	31.66	14.82	9.20–76.25	32.72	16.17	4.22–91.25
Distinctiveness	0.42	0.15	0.14–0.80	0.37	0.16	0.04–0.80

Moreover, effects of the semantic variables in the participants with aphasia may have been generally more heterogeneous than in the neurotypical participants. This heterogeneity was suggested by the marked difference in the effects between the full group and the more homogeneous sub-group of participants (but note that there were by-participant random slopes for semantic variables also in Papers 2 and 4, that significantly improved the model fit and indicated some degree of variability in their effects across neurotypical participants). For further discussion on this topic please also refer to the section below “Studying performance of people with aphasia as a group”.

Taken together, the presence of effects of semantic variables shows that the semantic structure of target words matters and that it influences word processing. However, a different combination of semantic variables was significant in each of the papers reported in this thesis and exactly which semantic variables were significant was inconsistent across the different investigations. What this might mean is that while I found significant effects of individual semantic variables in the different analyses, it may not necessarily be the case that *these particular* variables are important, but rather what is important is what they represent *more broadly*: how activation flows between items at the semantic level and how different ways that semantic knowledge overlaps can cause co-activation at the lexical level. While the semantic variables studied here measure different possible aspects of these dynamics, it may be the case that they are not actually perfect measures of the factors underlying activation spread and lexical co-activation and that other measures could be found that might represent aspects of these dynamics even better. The inconsistency in the findings may be taken to suggest that the precise way to measure the factors that influence this spread of activation and the size and activation strength of the co-activated cohort have not yet been identified. Following this idea, it may be possible that new measures that capture either the spread of activation or the size and strength of activation of the lexical cohort may outperform any of the individual measures investigated here. Further research is needed address this idea.

An interesting follow-up to Papers 1, 2, and 4 would be to statistically compare performance of people with and without aphasia and to test whether the effects of semantic variables actually differ statistically in these populations. This, however, was not possible with the data analysed here because the materials, items, and testing conditions were very different in Paper 1 compared to Papers 2 and 4.

Building on the behavioural results of Papers 1, 2, and 4, the findings of Paper 3 provide further insights into the possible cognitive mechanisms of effects of semantic variables on word production. I will first summarise the core findings of Paper 3, which used electrophysiological evidence to explore effects of semantic variables during online word planning, and then assess the contribution of the new knowledge gained to a better understanding of effects of semantic variables.

Which semantic variables affect processes during word planning and what might be the functional basis of their effects?

In Paper 3, to explore the electrophysiological signatures of effects of semantic variables during overt picture naming, ERP data (from electroencephalography (EEG) recordings) was collected in the experiment that also provided the behavioural data analysed in Paper 2. Following the only previous EEG study on semantic variables (Rabovsky et al., 2021), I analysed mean ERP amplitudes at a posterior region of interest (ROI) in a time-window between 200 and 550ms using linear mixed effects models to test for effects of the semantic variables (waveform analysis). In addition, I conducted a microstate analysis, which enabled me to test if the underlying neuronal networks involved were modulated by the semantic variables. The use of two different types of analyses on the same ERP data allowed me to thoroughly test for influences of the semantic variables on online word planning.

Combining across analyses, I found significant effects of number of semantic features, intercorrelational density, semantic similarity, and number of near semantic neighbours. Consequently, there was evidence for effects on brain activity during word planning of semantic variables that had not been tested using evoked responses before (i.e., semantic similarity and number of near semantic neighbours). Number of semantic features was the only semantic variable that was significant in both the waveform and the microstate analysis. The three remaining variables were only significant in the microstate analysis and all four variables affected the number of timeframes associated with Microstate 5.

Following previous literature (Rabovsky et al., 2021), I interpreted the posterior positivity in the mean amplitude analysis as reflecting activity in the lexical and semantic system during processing and speculatively generalised this interpretation to the effects of semantic variables on the number of timeframes associated with Microstate 5. Consequently, I argued that activity in the semantic and lexical network was higher for words with higher numbers of semantic features (activity related to the target representation itself), higher intercorrelational density and higher number of near semantic neighbours (activity distributed across the target representation and a co-activated semantically related cohort), as well as for words with higher semantic similarity (the underlying mechanism is unclear due to the inconsistencies in previous behavioural findings (Paper 1 of this thesis; Fieder et al., 2019) and the activity could be either related to the target itself or distributed across a co-activated cohort).

In the absence of electrophysiological markers of facilitation and interference or semantic and lexical processing in word production (unfortunately the effect of word frequency, which was supposed to act as a marker of lexical processing was non-significant), I can currently only speculate about the interpretation of the effects of semantic variables on the number of timeframes associated with Microstate 5 and alternative interpretations may be possible. Importantly though, given their semantic nature, we can be certain that the semantic variables affect processing where semantic information is required, which is during semantic and/or lexical processing in word production. Thus, even though the precise interpretation of the findings remains rather speculative at this point, Paper 3 revealed that these semantic variables, some of which did not show a significant influence in the analyses of the associated behavioural data (Paper 2), influence semantic and/or lexical processing. Subsequent work can build on this finding and further the understanding of the effects. Importantly, neither of the approaches to the ERP analysis yielded differences between effects that are facilitatory or inhibitory in the behavioural data or allowed discrimination between semantic and lexical contributions to the effects. Future research could attempt to disentangle these further to directly test the association of facilitatory effects with predominantly semantic processes and inhibitory effects with lexical processes (but see Paper 4 for a discussion of some of the difficulties associated with this endeavour).

Do effects of semantic variables depend on the processing requirements of a task?

All the studies of this thesis used overt naming, a task that people naturally engage in in their everyday lives, as an experimental paradigm. Importantly, this task is unlikely to trigger task-specific response strategies or to have major reliance on executive control functions (cf. e.g., Picture-Word Interference paradigm). In addition, in Paper 4, I used a speeded version of the standard picture naming paradigm to elicit a speed-accuracy trade-off in neurotypical participants with the aim of testing whether effects of semantic variables depend, at least to some degree, on the task requirements.

Two of the previous studies investigating semantic variables (Fieder et al., 2019; Mirman, 2011) used a speeded picture naming task and reported effects of semantic variables (i.e., number of near semantic neighbours and semantic similarity) that have been found to be non-significant in standard picture naming (e.g., Paper 2 in this thesis; Bormann, 2011; Hameau et al., 2019; Lampe et al., 2017). Previous research using speeded naming tasks (e.g., Kello, 2004; Mirman, 2011) had suggested that the time pressure in speeded naming may cause a modulation of *input gain*, a mechanism of cognitive control, making processing units more responsive to their inputs, which allows for faster responses at the expense of processing accuracy. In addition, Mirman proposed that this heightened sensitivity to input leads to stronger effects of item-inherent variables (near and distant semantic neighbours in Mirman, 2011).

Motivated by this suggestion and the discrepancies between effects of some semantic variables in the literature, the study reported in Chapter 5 (Paper 4) aimed to investigate whether effects of semantic variables in neurotypical participants depended on the processing speed required by the experimental task and on the resulting processing difficulties. However, in contrast to the predictions of the input gain account, effects of semantic variables were *not* systematically stronger in speeded than in standard picture naming. In fact, of the two semantic variables that interacted with naming task, only the effect of distinctiveness was stronger in the speeded naming task (response latency analysis), while the effect of number of semantic features was stronger in the standard naming task (naming accuracy analysis). Neither of these effects could be explained under the input gain account by suggesting increased responsiveness to inputs in speeded naming.

Consequently, I suggested, as I have throughout this thesis, that while some differences in effects of semantic variables between previous studies may have been caused by the specific naming task used, they more likely depend on differences in statistical power and control of other influential variables in the experiments, particularly as effects of the semantic variables were rather small with between participants variability.

Interestingly, for precisely those two variables that have only been found to be significant in speeded picture naming by previous research (i.e., number of near semantic neighbours and semantic similarity), the model selection process for the statistical analyses in Paper 4 retained random by-participant slopes that indicated that the model with these random slopes fit the data significantly better than a model without them. This suggests that the effects of these variables vary between (neurotypical) participants and may be a reason for the failures to replicate the effects of these semantic variables found in previous research. It is possible that Fieder et al. (2019) and Mirman's (2011) participant samples may have, coincidentally, consisted of (predominantly) participants that were affected by the respective variables in their investigations, while the larger participant sample included in my studies may have been more diverse in that respect, causing an overall non-significant effect of the same variables. However, more research into inter-individual variability of effects of the semantic variables is needed to assess this suggestion.

One aspect of Paper 4 that could be considered a shortcoming is that I did not analyse the simple effects of the semantic variables in the two tasks separately, which made me unable to comment on the effects of the semantic variables in the two tasks separately. However, given that the paper's focus was on a comparison of the two tasks, an analysis of the simple effects was beyond the scope of Paper 4. Moreover, in the context of a task comparison, I may have encountered a statistical dilemma that would have exacerbated the interpretation of the simple effects for the two tasks separately: In the absence of a significant interaction between the two tasks, a significant simple effect in either of the tasks would have been impossible to interpret with confidence. The non-significant interaction would have indicated that the effect of a variable is not significantly different between the two tasks, hence, the simple effect being significant in one of the tasks but not in the other, would have been indicative of a statistical error. This could have been a Type 1 error, false positive, for the significant simple effect, or a Type 2 error, false negative, for the non-significant interaction or the non-significant simple effect. It would have been impossible to disentangle these options without further research.

Importantly, and as mentioned in Paper 4, the findings of this study may have been influenced by an effect of repetition, which may make it impossible to directly compare them to other studies. As described in Paper 4, the participants had already completed the study reported in Papers 2 and 3 before completing the standard or speeded naming tasks reported in Paper 4. While I have argued that this experimental approach has positive effects on the internal comparability of the speeded and standard naming tasks analysed in Paper 4 (i.e., the speeded naming performance might otherwise be disproportionally affected by visual processing difficulties of the stimuli), it may also affect the findings in an undesired way. Previous work on semantic variables (i.e., number of semantic features and intercorrelational density more specifically) that used a familiarisation procedure (Rabovsky et al., 2016) or presented the experimental items twice to increase the number of items for the analyses (Rabovsky et al., 2021) has reported some influence of these practices on effects of semantic variables: The effect of number of semantic features on naming latency was weaker for participants that had been familiarised and the effect of intercorrelational density on naming latency was only significant in the repetition.

To assess the magnitude of any effects of repetition in my own investigations, I conducted a follow-up analysis where I compared the performance on the first and second standard naming tasks of the 41 participants that were included in the standard naming analysis of Paper 4 (i.e., participants that had performed tasks in the order: 1) standard naming, 2) standard naming, 3) speeded naming) with the 291 items included in Paper 2. I ran (Generalised) Linear Mixed Effect Models (lme4; Bates et al., 2015) on their naming latency and naming accuracy. The models included the same fixed effects that were considered in the analyses of Paper 4 and additional interactions between repetition and all semantic and control variables. Repetition was treatment coded (first presentation as -0.5 and second as 0.5).

The findings of these analyses are reported in Table 3. Unsurprisingly, responses were faster and more accurate in the second standard naming round. In the naming latency analysis, I found a significant interaction between repetition and number of semantic features with an attenuated effect in the second standard naming round compared to the first exposure to the items (Figure 1). A follow-up analysis with the emtrends function (Lenth, 2020) revealed that, in this subset of participants, the facilitatory effect of number of semantic features was indeed non-significant in the first naming round (Estimate = -0.01, *SE* = 0.01 , *t* = -1.26, *p* = .207), in contrast to what I had reported for all 87 participants in Paper 2, and also non-significant in the second standard naming round (Estimate = -0.00, SE = 0.01, t = -0.07, p = .942), as reported in Paper 4. However, as described for the interaction of imageability and task in Paper 4, in the presence of a significant interaction between number of semantic features and repetition, at least one of the non-significant simple effects is likely subject to a Type 2 error and is a false negative (presumably the non-significant effect in the first naming round), due to insufficient statistical power. Nonetheless, the weaker effect of number of semantic features in the second naming round resembles the attenuated effect of number of semantic features on naming latency in participants that had been familiarised with the materials in Rabovsky et al. (2016), though their effect of number of semantic features was significant in participants that both had and had not been familiarised (note that Rabovsky et al. (2021) did not report the outcome of the interaction between number of semantic features and task repetition in the naming latency analysis).

In contrast, in the naming accuracy analysis no semantic variable interacted with task repetition, which resembles the non-significant interactions of number of semantic neighbours and intercorrelational density and familiarisation or task repetition in Rabovsky et al. (2016) and Rabovsky et al. (2021), respectively.

These findings indicate that the effects of semantic variables, particularly on naming accuracy, were largely robust over time and item repetitions in standard naming. However, for response latencies, the practice of familiarising participants with the materials for a study, which is common in some research laboratories, may influence behaviour (even without feedback on target names; see Table 3, also for significant effects of repetition on control variables). Importantly, the effect of this practice on speeded naming cannot be assessed due to the experimental design used. Finally, the disappearance of the significant effect of number of semantic features on naming accuracy that was reported in Paper 2 after removing the half of the participants who subsequently performed speeded naming as the next task, highlights the fragility of the effects of semantic variables and the importance of highly powered investigations.

Table 3

Summarised output of picture naming latency and naming accuracy analyses comparing first and second rounds of standard naming in 41 participants.

	Naming latency						Naming accuracy					
Model structure	lmer(RT ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + OrdCatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness) * Repetition + (1 | Item) + (NearSemNeigh + SemSim || Participant), data, REML = TRUE)						glmer(ACC ~ (NameAgr + AoA + Imageability + ImageAgr + Frequency + Familiarity + Order + CatPos + NoFeats + IntercorrDens + NearSemNeigh + SemSim + Typicality + Distinctiveness) * Repetition + (1 | Item) + (NearSemNeigh + Typicality + IntercorrDens | Participant), data, family = binomial()					
Random effect	Variance	*SD*					Variance	*SD*	Correlation			
Item (Intercept)	0.01	0.11					1.44	1.20				
Participant (Intercept)	0.01	0.11					0.41	0.64				
Participant | NrSemNeigh	0.00	0.02					0.03	0.18	0.63			
Participant | SemSim	0.00	0.01										
Participant | Typicality							0.04	0.20	-0.02	-0.41		
Participant | IntercorrDens							0.03	0.16	-0.24	0.09	-0.61	
Residuals	0.05	0.23										
Fixed effects	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF	Estimate	*SE*	*CI*	*t*-value	*p*-value	VIF
(Intercept)	-1.29	0.02	-1.32 – -1.25	-69.76	<.001		3.13	0.13	2.88 – 3.39	24.03	<.001	
NameAgr	-0.05	0.01	-0.06 – -0.03	**-6.61**	**<.001**	1.10	0.74	0.08	0.58 – 0.90	**9.20**	**<.001**	1.08
AoA	0.02	0.01	-0.00 – 0.04	1.62	.106	2.42	-0.18	0.12	-0.42 – 0.06	-1.46	.146	2.47
Imageability	-0.02	0.01	-0.04 – -0.00	**-2.18**	**.029**	1.86	0.36	0.11	0.15 – 0.57	**3.32**	**.001**	1.84
ImageAgr	-0.04	0.01	-0.05 – -0.02	**-4.98**	**<.001**	1.25	0.29	0.09	0.12 – 0.47	**3.30**	**.001**	1.28
Frequency	-0.01	0.01	-0.02 – 0.01	-0.78	.433	1.57	0.21	0.10	0.02 – 0.40	**2.14**	**.033**	1.57
Familiarity	-0.03	0.01	-0.05 – -0.02	**-3.70**	**<.001**	1.72	0.07	0.11	-0.14 – 0.28	0.68	.500	1.76
Order	0.00	0.00	-0.00 – 0.01	0.72	.470	3.11	-0.05	0.05	-0.15 – 0.05	-1.04	.300	3.07
OrdCatPos	0.00	0.00	-0.01 – 0.01	-0.53	.597	3.18	0.06	0.06	-0.06 – 0.17	0.93	.355	3.16
NoFeats	-0.01	0.01	-0.02 – 0.01	-0.69	.492	1.57	0.18	0.10	-0.02 – 0.37	1.77	.078	1.55

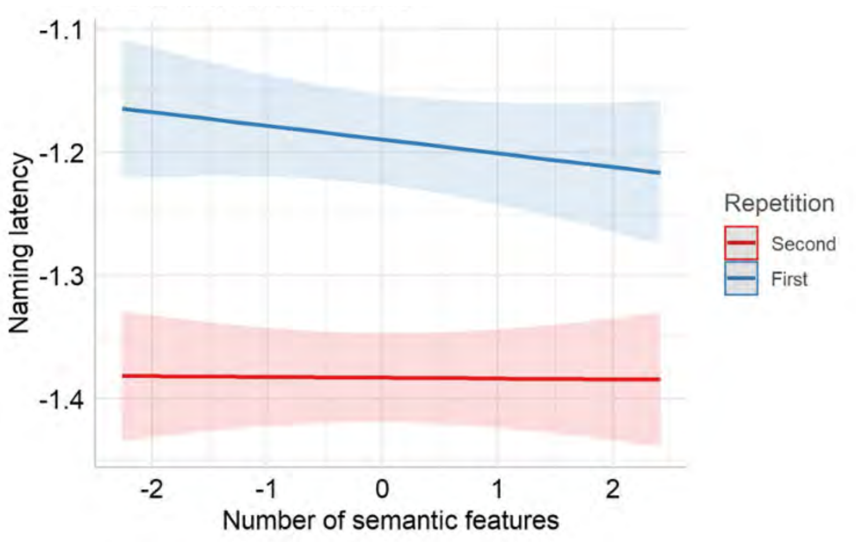
IntercorrDens	0.02	0.01	0.00 – 0.04	**2.13**	**.033**	2.18	-0.26	0.12	-0.50 – -0.03	**-2.23**	**.026**	2.19
NearSemNeigh	0.02	0.01	-0.00 – 0.04	1.57	.117	2.27	-0.09	0.13	-0.34 – 0.16	-0.71	.479	2.40
SemSim	0.01	0.01	-0.01 – 0.03	0.65	.515	2.33	-0.06	0.13	-0.30 – 0.19	-0.45	.654	2.46
Typicality	0.00	0.01	-0.02 – 0.01	-0.52	.601	1.43	0.03	0.10	-0.17 – 0.23	0.31	.759	1.53
Distinctiveness	0.02	0.01	0.00 – 0.04	**2.28**	**.023**	1.97	-0.14	0.11	-0.37 – 0.08	-1.26	.207	2.00
Repetition	-0.19	0.00	-0.20 – -0.19	**-60.61**	**<.001**	1.00	0.52	0.06	0.41 – 0.63	**9.28**	**<.001**	1.40
NameAgr * Repetition	0.01	0.00	0.00 – 0.02	**2.67**	**.008**	1.10	-0.02	0.04	-0.11 – 0.06	-0.53	.594	1.30
AoA * Repetition	-0.01	0.00	-0.02 – 0.00	-1.77	.077	2.39	-0.01	0.08	-0.16 – 0.13	-0.18	.855	2.91
Imageability * Repetition	0.00	0.00	-0.01 – 0.01	0.59	.555	1.86	-0.02	0.06	-0.14 – 0.10	-0.32	.747	2.06
ImageAgr * Repetition	0.02	0.00	0.02 – 0.03	**6.79**	**<.001**	1.26	-0.01	0.05	-0.11 – 0.09	-0.17	.867	1.41
Frequency * Repetition	0.02	0.00	0.02 – 0.03	**5.78**	**<.001**	1.57	-0.06	0.06	-0.17 – 0.05	-1.02	.306	1.65
Familiarity * Repetition	0.01	0.00	0.01 – 0.02	**3.23**	**.001**	1.76	-0.06	0.07	-0.19 – 0.07	-0.97	.330	2.08
Order * Repetition	0.01	0.00	-0.00 – 0.02	1.36	.175	2.21	-0.02	0.07	-0.16 – 0.11	-0.34	.732	2.19
OrdCatPos * Repetition	0.01	0.01	0.00 – 0.02	**2.68**	**.007**	2.74	0.00	0.08	-0.16 – 0.15	-0.04	.972	2.73
NoFeats * Repetition	0.01	0.00	0.00 – 0.02	**2.57**	**.010**	1.64	0.03	0.06	-0.09 – 0.14	0.42	.673	1.73
IntercorrDens * Repetition	-0.01	0.00	-0.01 – 0.00	-1.07	.287	2.21	-0.05	0.07	-0.18 – 0.09	-0.68	.494	2.43
NearSemNeigh * Repetition	0.01	0.00	-0.00 – 0.02	1.86	.063	2.44	-0.04	0.07	-0.18 – 0.10	-0.52	.602	2.94
SemSim * Repetition	-0.01	0.01	-0.02 – 0.00	-1.85	.065	2.50	0.00	0.08	-0.16 – 0.15	-0.06	.956	2.91
Typicality * Repetition	0.00	0.00	-0.01 – 0.01	0.64	.520	1.60	0.05	0.06	-0.07 – 0.18	0.83	.409	1.91
Distinctiveness * Repetition	-0.01	0.00	-0.01 – 0.00	-1.11	.265	2.10	0.01	0.07	-0.13 – 0.14	0.08	.939	2.25
Observations	20,932						23,706					
Marginal R^2 / Conditional R^2	0.189 / 0.449						0.216 / 0.505					

Note. VIF = Variance Inflation Factor, NameAgr = name agreement, ImageAgr = image agreement, AoA = age of acquisition, OrdCatPos = ordinal category position, NoFeats = number of semantic features, IntercorrDensity = intercorrelational density, NrSemNeigh = Number of near semantic neighbours, SemSim = Semantic similarity, Distinct = distinctiveness, Participant | X = random slope of X by participants.

Values of significant effects ($p < .05$) are printed in bold.

Figure 1

Significant interaction between repetition and number of semantic features in naming latency analysis



Note. Number of semantic features was standardised; Naming latency was negatively reciprocally transformed.

Methodological strengths and challenges

The biggest strengths of the papers presented in this thesis are that effects of semantic variables were studied in well-controlled experimental investigations and that converging evidence from different populations and methodologies was used. Moreover, the data was analysed with robust statistical methods that accounted for variability between participants and items in all studies (i.e., (Generalised) Linear Mixed Effect Models).

Nonetheless, across the studies, several methodological considerations came to light, which may be important for future research into effects of semantic variables on word production and their consideration by subsequent studies may further advance our understanding of effects of semantic variables. Some of these limitations were common to all papers of this thesis and are more general in nature, while others related to individual papers.

Multicollinearity

In my study of semantic variables, to decrease the risk of Type 1 errors, false positive findings that may have arisen by chance, and to really understand the contributions of individual semantic variables to behaviour, I needed to study these variables under conditions of maximal experimental control for effects of other variables that can influence word production. This included other semantic and psycholinguistic control variables that each may account for some variability in the data (even if there might not have been main effects of some of these variables) and therefore, through their inclusion, I increased the reliability of the measurement of the effect of the semantic variables of interest. Yet, given the natural correlation among these item-inherent variables, a worry was that including all six semantic variables in the models at the same time, in addition to other control variables, might lead to multicollinearity, which causes unreliable and unstable estimates of regression coefficients (Allison, 2012).

One way to identify harmful levels of multicollinearity is through the use of variance inflation factors (VIFs), which I provided for all the statistical analyses in this thesis. VIF values above a certain cut-off (e.g., 2.5, Allison, 2012; 5.0, Hair et al., 2014; Rogerson, 2001) indicate compromising effects of multicollinearity. Importantly, VIF of the studies reported in this thesis were generally low and in most cases under the more conservative cut-off value formulated by Allison. Therefore, including the six semantic as well as the control variables in the linear mixed effects models for the statistical analyses was acceptable and did not lead to harmful multicollinearity.

Another approach to this issue could have been the formulation of composite scores of the semantic and control variables using Principal Component Analysis (PCA). However, in this thesis I wanted to identify every individual variables' contribution to behaviour, which would not have been possible using principal components. For example Clarke et al. (2013) and Hameau et al. (2019) combined different semantic variables (e.g., number of feature-based near semantic neighbours and rated number of competitors as the feature-based neighbourhood measure in Hameau et al., and the relative degree of shared and distinctive features associated with the concept and the correlation of the distinctive features as the relative distinctiveness measure in Clarke et al.). While they were able to comment on the overall significance of the contribution of these components to behaviour, they were unable to speak to the contribution of the individual semantic variables that weighted on the component. Hence, following such an approach makes it impossible to know which of the variables contained in a principal component drives performance and whether they are predictors of behaviour on their own. Consequently, this approach was unsuited for this thesis, but for future research Factor or Principal Component Analyses may be a way forward. However, for the data presented in Paper 1, I did conduct an exploratory, and unreported, Principal Component Analysis. This did not result in sensible components for the semantic variables, suggesting that caution may be required: It was, for example, not the case that those variables that showed facilitatory effects weighted on one component and those with inhibitory effects weighted on another.

Including only a subset of McRae et al.'s (2005) items

The feature database by McRae et al. (2005) contains 541 concepts, however, I only used a subset of these items. This was necessary given the reduced number of McRae et al. items that are also part of the Philadelphia Naming Test (Roach et al., 1996; Paper 1) or that had sufficiently high name agreement in Australian English (> 75%, Papers 2–4) (albeit it reduced statistical power). However, importantly, and as shown in Table C1 in Appendix C of Paper 2, the items selected for Papers 2–4 were a good representation of the full McRae et al. set of items as the values of semantic variables of the selected items were largely comparable to the whole database. This suggests that the reduction of stimuli did most likely *not* affect the comparability of the findings of the papers presented here to previous work using the full McRae et al. (2005) database (i.e., Rabovsky et al., 2016, but note that this study did not account for name agreement in German for the McRae et al. items).

Moreover, by removing items from McRae et al.'s (2005) database that had lower name agreement in Australian English, I believe that cultural differences between American and Australian English speakers have been minimised. As argued in Paper 2, one can probably assume that cultural differences would not have a dramatic impact on conceptual representations (and features) of the items in the database. However, the lower name agreement on the removed items suggests that they might have been unknown to the Australian participants (e.g., 'gopher') and that Australian undergraduate students were not sufficiently familiar with the concepts to name them accurately. By removing such items, I ensured that the experimental findings were minimally affected by them.

Finally, the 541 items of the McRae et al. (2005) database themselves are just a subset of the words in our mental lexicons and consequently represent only a segment of our word knowledge; for example, they do not contain abstract words. However, it is assumed that the items in the database are a representative sample of the concrete nouns in our lexicons such that any findings based on these items can be generalised to processing for noun production in general. Moreover, the semantic variables that I focused on are calculated either based on information on only the target concept (i.e., number of semantic features) or by taking information on all 541 items of the database into account (e.g., number of near semantic neighbours, distinctiveness). Thus, the use of a subset of items for my experiments does not impact on these calculations and, therefore, will not have compromised the generalisability of the findings to other words in the database nor other (concrete) words in our mental lexicons.

Studying performance of people with aphasia as a group

Aphasia is a heterogenous disorder with varying impairment profiles and consequently the 'average' performance of a person with aphasia is not meaningful (e.g., Nickels & Howard, 1995). Nonetheless, in Paper 1, I studied the participants of the MAPP Database (Mirman et al., 2010) with the only premise being that they had word retrieval difficulties. However, these difficulties may have been caused by a breakdown of processing at different levels. Importantly though, as noted by Howard (2003), it may only be meaningful to combine the results of a number of participants if they suffer from the same underlying impairment(s).

In the study on effects of semantic variables in participants with aphasia (Paper 1, Chapter 2), I investigated a mixed group of participants with aphasia, but rather than grouping these by underlying impairment, which is only possible with extensive background information that was not available here, I used statistical techniques to account for individual differences. My statistical approach using Generalised Linear Mixed Effects Models and including random slopes by participants accounted for differences between participants and allowed for the effects of semantic variables to differ between participants. Furthermore, I included the participants' performance on the Pyramids and Palm Trees test (Howard & Patterson, 1992) in the analyses, as a measure of their semantic abilities. However, the

Pyramids and Palm Trees test has a skewed distribution with most participants performing relatively well, which suggests that it may have lacked the sensitivity to fully capture the participants' semantic abilities. Consequently, it might not have been an ideal measure of semantic abilities, but it was the only measure of (input) semantic processing that was available for most of the participants of the MAPP Database.

To further account for differences between participants, I attempted to create a more homogeneous group of participants with a common level of impairment by only considering the data of participants that made few phonological errors. The participants of this sub-group therefore presumably had a predominantly semantic and/or lexical impairment. It is possible that another approach to define the subgroup may have been better: Instead of selecting participants based on the absence of many phonological errors it may have been useful to test for the presence of semantic and/or lexical impairments. Consequently, more detail on the exact location of the participants' processing breakdown would have been informative, yet this information was unavailable, and the additional administration of further assessment was impossible given that these were participants recruited by another research laboratory (in another country) over several years in the context of the MAPPD project.

In addition, with the error type analyses, I attempted to further control for influences of the heterogeneity of the participant group, given that an increase in certain types of errors can point to a certain level of impairment (e.g., phonological errors can point to a problem of the phonological output buffer). However, naming errors can be multiply determined, with, for example, semantic errors arising from breakdown at the semantic or the lexical level or the link between the two levels, thus impeding association of a specific error type with a clear-cut locus of impairment.

It may also have been beneficial for the interpretation of the findings of the group analyses to examine effects of semantic variables on individual participants. This would be in line with another approach to studying performance of participants with aphasia: single-case or case-series analyses (Howard, 2003; Nickels, 2002; Nickels et al., 2011; Schwartz & Dell, 2010; for examples see e.g., Hameau et al., 2019; Rossiter & Best, 2013). In contrast to analyses at the group level, these enable a more detailed assessment of differences and similarities between individual participants, to take into consideration the level(s) of their underlying impairments, and to link them to the observed behaviour to inform theory (Nickels, 2002; Nickels et al., 2011).

Subsequent research may follow this approach and examine the way individual participants with aphasia are affected by the semantic variables in more detail. This would allow identification of the characteristics of individuals with aphasia who are most strongly affected by semantic variables. Careful selection of participants with either 'pure' semantic (pre-lexical; or testing participants with Semantic Dementia) or lexical impairments may allow the researchers to further test the association of facilitatory and inhibitory effects of semantic variables with different levels of processing (i.e., semantic and lexical).

Using electrophysiological evidence to study effects of semantic variables during online word processing and to inform theory

In Paper 3 (Chapter 4) of this thesis I tested effects of the semantic variables during online word processing with the help of electrophysiological evidence to determine how they affect online word planning rather than the net outcome of processing in terms of response latencies or naming accuracy. This study had both strengths and limitations related to methodology and interpretation of the findings. In contrast to some previous work, including the study I attempted to replicate, Rabovsky et al. (2021), I controlled for multiple comparisons in the time course analysis, which was meant to provide a clearer idea of the temporal development of effects of the semantic variables. Interestingly, none of the effects survived correction for multiple comparisons, and I was consequently unable to discuss the precise temporal development of the effects. However, this highlights the necessity to treat claims about the time course of effects with caution, unless the respective analyses controlled for multiple comparisons, whenever appropriate.

While precise timing information may have furthered our understanding of effects of the semantic variables, we know that they influence semantic and lexical processing as they can be assumed to affect processing where semantic information is required. Importantly though, a differentiation between effects of semantic variables on semantic and lexical processing may be impossible if they, as argued before, affect both levels of processing (i.e., while they originate at the semantic level, they also affect lexical processing, e.g., stronger semantic activation due to spreading activation leads to stronger activation of the target's lexical representation). Additionally, semantic and lexical processing might interact (interactive or parallel processing), impeding a differentiation between these two levels using ERPs or even rendering it impossible. Nevertheless, the data showed that effects of semantic variables occurred relatively late compared to the established time course of semantic and/or lexical processing (Indefrey, 2011; Indefrey & Levelt, 2004). Possible reasons for this were discussed in Paper 3 and, the results of this study, therefore, call for further critical assessment of the time course estimates provided by Indefrey and Levelt.

Another strength of Paper 3 is that it includes a replication of a previous study by Rabovsky et al. (2021; similar to Paper 2 for behavioural effects). Replication is an important process in assessing the reliability of a cognitive phenomenon of interest and should be more widely embraced as there are findings in speech production that may be used as a basis of subsequent research or theory building that have failed to replicate (e.g., Jescheniak et al., 2009; Lee & de Zubicaray, 2010). With the waveform analyses, I replicated the analysis approach used by the only previous EEG study investigating effects of semantic variables on word production (Rabovsky et al., 2021) and used the same time-window included in the analysis (200–550ms in mean amplitude analysis) as well as the posterior region of interest (ROI). However, following the estimates of word production (Indefrey, 2011) this time-window did not include semantic processing, which might be disadvantageous when studying effects of variables that likely originate at the semantic level (please note that the time course analysis included the time-window traditionally associated with semantic processing (i.e., 0–200ms post picture onset), but, unfortunately, findings from the time course analysis could not be interpreted as they did not survive correction for multiple comparisons). In addition, the posterior ROI was chosen by Rabovsky et al. based on previous research and was assumed to reflect the sites of competition during lexical selection. However, this approach made it impossible to detect effects of semantic variables on processing that may be apparent at other scalp locations (see e.g., Binder et al., 2009; Indefrey, 2011, for reviews). A different approach to the waveform analysis that takes data from the entire time course of word planning as well as ERP data from the whole brain into consideration may overcome some of these shortcomings. Importantly, these sources of information were considered in my microstate analysis.

A possible way to improve the ERP analyses conducted in Paper 3 would be to refine the control of articulation artifacts before the statistical analysis. I used a fixed time-window approach and removed fast responses (< 550ms) to not include any trials where responses were given within the time-window considered for the analyses and to obtain a signal that was uncontaminated by articulatory artifacts (note that only relatively few trials were actually removed using this approach as the average naming latency for this study was 900ms). Given that I was interested in "earlier" encoding processes (i.e., semantic and lexical processing), I assumed the analysed time-window would be largely unaffected by articulatory influences for the remaining items. In addition, an ICA analysis was used to identify and remove components from the data that were related to eye, muscle, and heart activity. However, another approach could have been to use an algorithm which selectively removes articulation artifacts (e.g., residue iteration decomposition, RIDE; Ouyang et al., 2016). Similarly, I considered another popular approach to analysis, the combined stimulus- and response-aligned ERP analysis (Laganaro & Perret, 2011), which covers the word planning process form picture onset to just before articulation while accounting for individual differences in processing time and thus reducing the influence of jittered processing. However, I decided not to use this approach as it is of particular use when investigating late encoding processes (phonological and phonetic encoding, Laganaro, 2014), which was not the case here. Finally, when replicating an ERP analysis, the data cleaning procedure (e.g., application of filters) should ideally be identical to the original study (G. Vigliocco, personal communication, October 11, 2019). However, the use of an ICA analysis prevented me from applying exactly the same pre-processing procedure as in Rabovsky et al. (2021).

Operationalising semantic knowledge

For this thesis, I operationalised semantic variables based on a semantic feature database (McRae et al., 2005) and provided a thorough investigation of feature-based semantic variables. Consequently, the findings of significant effects of these variables could be interpreted as providing evidence for, and a proof of concept of, feature-based models of meaning (McRae et al., 1997; Tyler & Moss, 2001; Vigliocco et al., 2004). However, as I have argued in the different experimental papers, it may be the case that the feature-based semantic variables also represent aspects of holistic semantic representations. For example, despite assuming holistic lexical concepts, Abdel Rahman and Melinger (2009, 2019) talk about related concepts (e.g., 'cat', 'dog', 'tiger' in Figure 1 in the General Introduction) sharing semantic *features* (such as *is a pet, has a tail,* and *is feline*) that are holistic lexical concepts themselves. Hence, and as I have argued in the experimental chapters of this thesis, these connections to other concepts may capture the semantic variables that I operationalised based on decomposed semantics. For example, the number of connections of a target lexical concept to other lexical concepts may be equivalent to the number of semantic features. Similarly, the feature-based measure of intercorrelational density may be representative of the strength of the labelled links between a target and other holistic lexical concepts and distinctiveness could capture the average number of links to other concepts of the lexical concepts that the target is directly connected to.

Importantly, however, and as noted in the General Introduction, this thesis did not set out to differentiate between different ways to operationalise semantic knowledge. However, future work may explore further how feature-based semantic variables may be represented in holistic architectures of semantics.

Other forms of semantic organisation besides the feature-based approach also provide different ways to operationalise semantic knowledge as semantic variables (e.g., association-based, context-based). Future work should also consider these in similarly well-controlled experimental investigations and integrate them with feature-based variables in exploration and interpretation of effects of semantic variables on word production. Association-based semantic variables may be of particular interest as some theories of word production (e.g., Abdel Rahman & Melinger, 2009) explicitly propose that not only categorical but also associative relations (via associative links) are encoded during processing. In addition, other researchers (e.g., De Deyne et al., 2017) have proposed that word knowledge is entirely represented by a network of associatively connected word nodes, and association norms (e.g., De Deyne et al., 2019; Nelson et al., 2004) allow the easy calculation of semantic variables just as was the case for the feature-based semantic variables investigated here. Association-based semantic variables that have been investigated by previous research include, for example, the number of associates (Hameau et al., 2019; Pexman et al., 2007; Rabovsky et al., 2012) and the strength of the first associate (Griffiths & Steyvers, 2003).

In addition, there is previous work on effects of different types of semantic features on processing, which I have not included in this thesis. For example, Clarke et al. (2013) studied influences of the proportion of a concept's features that were visual features, Rico Duarte and Robert (2014) investigated effects of the number of perceptual or functional features on naming accuracy of neurotypical participants and people with Alzheimer's Disease, and Miozzo et al. (2015) studied the impact of the number of encyclopaedic and action features. Future work into effects of semantic variables on picture naming should also account for such feature-specific information, which could provide further information on whether and how certain types of features are more influential than others.

Moreover, interactions between semantic variables were not considered here. In the studies of this thesis, I only looked at main effects (and interactions of semantic variables with the semantic abilities of people with aphasia in Paper 1, and with task in Paper 4). However, it is likely that interactions between semantic variables may also be of importance. Such interactions may help to further understand how the semantic variables relate to one another. For example, as discussed in Paper 1, higher typicality has sometimes been described to be due to more shared, more frequent, or more strongly intercorrelated features within a semantic category (e.g., McRae et al., 1999; Pexman et al., 2003; Rogers et al., 2004; Woollams et al., 2008). This suggests that it may be impossible to fully tease apart effects of some semantic variables and highlights the difficulty understanding their effects when variance that is shared with other semantic variables is accounted for. Computational modelling might aid the interpretation of such effects.

Further caveats relate to the operationalisations of semantic variables that were included in this thesis. For example, the measure of number of near semantic neighbours used an arbitrary cut-off to define near semantic neighbours. When calculating this measure, I followed the approach suggested by Mirman (2011) and defined a near semantic neighbour as a word with feature vector cosine similarity of at least 0.4 with the target word. In contrast, Fieder et al. (2019) used a cosine similarity of 0.5 as a cut-off for near semantic neighbours. This decision drastically changes the number of near semantic neighbours a target has: For the 297 items of Papers 2, 3, and 4, using a cut-off of cosine similarity > 0.4 to define a near semantic neighbour, the items had an average of 6.14 near semantic neighbours (range = 0–38 items, *SD* = 7.56). In contrast, for the same items, the use of cosine similarity > 0.5 to define a near semantic neighbour reduced the average number of near semantic neighbours to 2.67 (range = 0–29 items, *SD* = 4.65). It is possible that such a difference in the measure (paired-samples *t*-test: $t(296) = 15.13$, $p = < .001$) causes a difference in the effect of this semantic variable depending on exactly which of these operationalisations is chosen. This severely impedes the ability to compare and generalise findings between studies. Future research into effects of semantic variables needs to be mindful of such seemingly negligible differences in the way semantic variables are operationalised and should systematically investigate these differences and their effects.

Moreover, some of the semantic variables investigated in this thesis were confounded with other semantic measures due to the way they were operationalised. More specifically, intercorrelational density was calculated in McRae et al. (2005) as the summed percentage of shared variance across all significantly correlated feature pairs of a concept. Consequently, concepts with more correlated feature pairs would have a higher intercorrelational density, even if the individual correlations may have been just above the threshold of 6.5% shared variance. Hence, words with higher numbers of semantic features had an increased chance of obtaining higher intercorrelational density scores simply because they had more combinations of feature pairs, which could have been correlated above threshold (note that e.g., Taylor et al. (2012) anticipated this issue and calculated their intercorrelations measure as the *average* correlation strength instead of a sum). In a similar vein, typicality was calculated by summing the weighted feature values capturing a feature’s prevalence in its semantic category as well as its production frequency. Hence, this measure was also confounded with the number of semantic features of a concept (see e.g., Table 2 in Paper 1 for positive correlations between number of semantic features and both intercorrelational density and typicality).

In addition, subsequent research may investigate the relation of the feature-based semantic variables that I focused on and other semantic variables that were treated as covariates in this thesis. For instance, imageability facilitated naming accuracy of the participants with predominantly semantic and/or lexical impairments (Paper 1) and of the neurotypical participants of Papers 2 and 4. In contrast, the effect was non-significant on picture naming latency (similar to Morrison et al., 1992) in neurotypical participants, although it interacted with task in Paper 4, and in the ERP analyses (Paper 3).

Finally, some of the control variables that were included in the analyses (i.e., variables that were rated by an independent group of participants in Study 1 of Paper 2), due to the sheer number of items and variables that had to be rated, were based on ratings of only relatively few participants (n = 11–12 per item). Further norms should be collected to improve the reliability of these measures.

Theoretical implications of the findings of this thesis

In this thesis, I have demonstrated that the structure of the semantic representations of the words we say matters for word production. With the exception of investigations of rated imageability and concreteness, most of the previous research into this topic was conducted in word comprehension. However, as outlined in the General Introduction, spoken word production is a particularly interesting modality to study these effects because of the critical importance of semantic information for processing—it influences processing at both the semantic and the lexical level.

In line with the limited previous research that is available, I have shown that some of the feature-based semantic variables facilitate performance in picture naming, while others inhibit performance. Consequently, theories of word production need to go beyond accounting for data from context manipulation paradigms and embrace the evidence of facilitatory and inhibitory influences arising from item-inherent variables. In the section "Which semantic variables affect word production in people with and without aphasia and what are the underlying mechanisms?" above, I detailed the repertoire of possible mechanisms of semantic facilitation and inhibition that might be at play. While facilitatory effects seem to require some variant of activation spread at the semantic level (e.g., spreading activation between lexical concepts, feedback from (co-activated) lexical to semantic representations), inhibitory effects of semantic variables are easily explained by theories assuming lexical selection to be competitive and possibly even by non-competitive accounts with long-term adjustments of connections between the semantic and lexical representations (learning mechanism, Oppenheim, 2010; as discussed in Paper 2, Chapter 3), however, this hypothesis has to be tested in further research.

Consequently, given their current model architectures, WEAVER++ (Levelt et al., 1999) and Howard et al. (2006) are able to account for the inhibitory effects and need further specifications of the spread of semantic activation to explain the facilitatory findings of this thesis. In contrast, the Interactive Activation Model (Dell, 1986; Dell et al., 1997) may be able to explain the facilitatory, but not the inhibitory findings presented in this thesis and would, for example, need to adopt competition between co-activated lexical representations to account for these findings. The Incremental Learning Model (Oppenheim et al., 2010) and the Ballistic Model of Lexical Access (Mahon & Navarrete, 2016; Navarrete et al., 2014) need further specification of the mechanism underlying weakening of semantic-to-lexical connections but may be able to account for inhibitory effects of semantic variables, while the Ballistic Model seems able to explain facilitatory effects via spreading activation and the Incremental Learning Model explicitly moved facilitation out of the model. Of the currently available theories of word production, solely the Swinging Lexical Network (Abdel Rahman & Melinger, 2009, 2019) explicitly addressed word-inherent variables and seems to comprise the architectural elements (i.e., clearly stated spreading activation and lexical competition) to parsimoniously explain both facilitatory and inhibitory findings.

Based on these mechanisms hypothesised on the basis of behavioural research, facilitatory effects are associated with semantic processing and inhibitory with lexical processing. However, the ERP investigations (Paper 3) did not yield evidence in favour of a separable two-step process, with no distinct electrophysiological signatures for behaviourally faciliatory and inhibitory variables. Moreover, irrespective of their behavioural effects, semantic variables influenced the same underlying neuronal network, which may be taken to suggest that semantic and lexical processing are interactive or parallel processes, however, as argued in Paper 3, the findings could also be in line with separate and sequential semantic and lexical processing mechanisms.

Importantly, irrespective of the precise mechanism underlying their effects, semantic variables do not seem to be restricted to influencing only one of the levels of processing. As I argued throughout this thesis, they likely operate at both the semantic and the lexical level of processing. Following Abdel Rahman and Melinger (2009, 2019), spread of activation at the semantic level may facilitate semantic processing by increasing the activation of the target's semantic and, consequently, lexical representations. However, it also leads to the activation of lexical representations of those co-activated lexical concepts and thereby interferes with lexical processing. Consequently, the way activation spreads through the semantic system determines the size and activation strength of the co-activated lexical cohort. This ultimately determines the direction of the behavioural effect: If the size and strength of activation of the co-activated lexical cohort yields substantial lexical competition, the effect of a variable will be inhibitory, else it is facilitatory. Crucially, it has not yet been established exactly how *close* to the target or how *numerous* lexical competitors must be to result in interference.

In summary, effects of semantic variables seem to originate at the semantic level with the locus of the effect being the lexical level where selection is either facilitated due increased activation of the target's lexical representation relative to any competitors or it is inhibited from a lexical cohort.

This thesis has shown that it is of primary importance that theories of word production are more explicit about mechanisms causing semantically related lexical representations to be co-activated and how exactly semantic facilitation may occur in standard picture naming without context manipulations. I have demonstrated that effects of semantic facilitation and interference go beyond context manipulations and this new data must be integrated in word production theories to make them more accurate representations of the way we say words.

Concluding remarks

Throughout this thesis I have argued that semantic variables can inform models of word production and that it is necessary to determine exactly which semantic variables affect processing and to explain the mechanism underlying their effects. The four experimental studies of this thesis have moved us closer to better understand *which* feature-based semantic variables influence behaviour and *how* they affect processing in word production.

Firstly, as clearly shown in Table 1 of this General Discussion, *all* of the six feature-based semantic variables were important in at least one of the studies: number of semantic features, semantic similarity, and typicality in people with aphasia (Paper 1, Chapter 2), number of semantic features, intercorrelational density, and distinctiveness in neurotypical speakers (Paper 2, Chapter 3), number of semantic features, intercorrelational density, number of near semantic neighbours, and semantic similarity in ERPs (Paper 3, Chapter 4), as well as stronger effects of distinctiveness in the speeded naming task and stronger effects of number of semantic features in the standard naming task (Paper 4, Chapter 5). While number of semantic features seemed to affect performance most reliably across all papers, the picture is far from clear, with no set of variables consistently affecting performance. As I have argued earlier in this General Discussion, this might be indicative that what is important is what the semantic variables represent more broadly (i.e., degree of spread of activation at the semantic level and co-activation at the lexical level) rather than individual semantic variables themselves.

Secondly, for each effect of a semantic variable I have formulated hypotheses about the possible underlying cognitive mechanism(s) by assessing the explanatory power of the range of processing mechanisms proposed by current theories of word production and by using neuropsychological evidence from a large group of participants with aphasia as well as behavioural and electrophysiological data from neurotypical speakers. I found that semantic variables can both facilitate and inhibit performance and that these effects presumably originate at the semantic level and have their locus at the lexical level of processing. Consequently, I argued, that theories of word production must include processing components that allow for facilitation and inhibition of performance, which could be mechanisms of spreading activation and lexical competition. Crucially, current theories of word production should use the rich evidence provided by item-inherent (semantic) variables and integrate findings from standard picture naming studies, like the ones of this thesis, into their models.

Bringing together the spectrum of evidence generated by this thesis, the findings clearly highlight the importance of the role that the structure of word knowledge has for spoken word production. However, they also emphasise the complexity of the underlying theoretical mechanisms and the dramatic lack of explicit and specific statements on effects of item-inherent variables by theories of word production. Importantly, the findings presented here move us closer to an understanding of the elements necessary to explain word-inherent effects of semantic variables and to understand the mechanisms underlying word production.

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*(5), 713–734. https://doi.org/10.1080/01690960802597250

Abdel Rahman, R., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience*, *34*(9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Allison, P. (2012). *When can you safely ignore multicollinearity?* Statistical Horizons. https://statisticalhorizons.com/multicollinearity

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bormann, T. (2011). The role of lexical-semantic neighborhood in object naming: Implications for models of lexical access. *Frontiers in Psychology*, *2*, 1–11. https://doi.org/10.3389/fpsyg.2011.00127

Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2013). From perception to conception: How meaningful objects are processed over time. *Cerebral Cortex*, *23*(1), 187–197. https://doi.org/10.1093/cercor/bhs002

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*(3), 371–414. https://doi.org/10.1207/s15516709cog2303_4

De Deyne, S., Kenett, Y. N., Anaki, D., Faust, M., & Navarro, D. (2017). Large-scale network representations of semantics in the mental lexicon. In M. N. Jones (Ed.), *Big data in cognitive* *science - Frontiers of cognitive psychology* (pp. 174–202). Routledge/Taylor & Francis Group.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124–142. https://doi.org/10.1016/0749-596X(88)90070-8

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*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 1–17. https://doi.org/10.3389/fpsyg.2014.00781

Fieder, N., Wartenburger, I., & Abdel Rahman, R. (2019). A close call: Interference from semantic neighbourhood density and similarity in language production. *Memory & Cognition*, *47*(1), 145–168. https://doi.org/10.3758/s13421-018-0856-y

Griffiths, T. L., & Steyvers, M. (2003). Prediction and semantic association. *Advances in Neural Information Processing Systems 15 - Proceedings of the 2002 Conference, NIPS 2002*, *15*.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson.

Hameau, S., Nickels, L., & Biedermann, B. (2019). Effects of semantic neighbourhood density on spoken word production. *Quarterly Journal of Experimental Psychology*, *72*(12), 2752–2775. https://doi.org/10.1177/1747021819859850

Howard, D. (2003). Single cases, group studies and case series in aphasia therapy. In I. Papathanasiou & R. De Bleser (Eds.), *The sciences of aphasia* (pp. 245–258). Elsevier. https://doi.org/10.1016/b978-008044073-6/50017-1

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures.* Thames Valley Test Company.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 1–16. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Jescheniak, J. D., Oppermann, F., Hantsch, A., Wagner, V., Mädebach, A., & Schriefers, H. (2009). Do perceived context pictures automatically activate their phonological code? *Experimental Psychology*, *56*(1), 56–65. https://doi.org/10.1027/1618-3169.56.1.56

Kello, C. T. (2004). Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 942–955. https://doi.org/10.1037/0096-1523.30.5.942

Laganaro, M. (2014). ERP topographic analyses from concept to articulation in word production studies. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.00493

Laganaro, M., & Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography*, *24*(1), 19–29. https://doi.org/10.1007/s10548-010-0162-x

Lampe, L. F., Fieder, N., Krajenbrink, T., & Nickels, L. (2017). Semantische Nachbarschaft in der Wortproduktion bei Aphasie [Semantic neighbourhood in word production in aphasia]. In A. Adelt, Ö. Yetim, C. Otto, & T. Fritzsche (Eds.), *Spektrum Patholinguistik - Panorama Patholinguistik: Sprachwissenschaft trifft Sprachtherapie* (Vol. 10, pp. 103–114). Universitätsverlag Potsdam. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/39701

Lee, M. M., & de Zubicaray, G. (2010). Lexical selection is by competition: A failure to replicate Mahon et al.’ s (2007) Experiment 7. *Poster Presented at the 6th International Workshhop of Language*

Production, Edinburgh.

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means* (1.5.2-1). https://cran.r-project.org/package=emmeans

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.

Mahon, B. Z., & Navarrete, E. (2016). Modelling lexical access in speech production as a ballistic process. *Language, Cognition and Neuroscience*, *31*(4), 521–523. https://doi.org/10.1080/23273798.2015.1129060

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

McRae, K., Cree, G. S., Westmacott, R., & Sa, V. R. De. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology*, *53*(4), 360–373. https://doi.org/10.1037/h0087323

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99–130. https://doi.org/10.1080/10430719008404646

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, *25*(10), 3343–3355. https://doi.org/10.1093/cercor/bhu137

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective and Behavioral Neuroscience*, *11*(1), 32–43. https://doi.org/10.3758/s13415-010-0009-7

Mirman, D., Strauss, T. J., Brecher, A., Walker, G. M., Sobel, P., Dell, G. S., & Schwartz, M. F. (2010). A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function. *Cognitive Neuropsychology*, *27*(6), 495–504. https://doi.org/10.1080/02643294.2011.574112

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705–714. https://doi.org/10.3758/BF03202720

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. https://doi.org/10.1016/j.jml.2014.05.003

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, *16*(10–11), 935–979. https://doi.org/10.1080/02687030244000563

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, *33*(10), 1281–1303. https://doi.org/10.1016/0028-3932(95)00102-9

Nickels, L., Howard, D., & Best, W. (2011). On the use of different methodologies in cognitive neuropsychology: Drink deep and from several sources. *Cognitive Neuropsychology*, *28*(7), 475–485. https://doi.org/10.1080/02643294.2012.672406

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*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Ouyang, G., Sommer, W., Zhou, C., Aristei, S., Pinkpank, T., & Abdel Rahman, R. (2016). Articulation artifacts during overt language production in event-related brain potentials: Description and correction. *Brain Topography*, *29*(6), 791–813. https://doi.org/10.1007/s10548-016-0515-1

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). The neural consequences of semantic richness: When more comes to mind, less activation is observed.

Psychological Science, *18*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2007.01913.x

Pexman, P. M., Holyk, G. G., & Monfils, M.-H. (2003). Number-of-features effects and semantic processing. *Memory & Cognition*, *31*(6), 842–855. https://doi.org/10.3758/BF03196439

Pexman, P. M., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience*, *7*, 1–3. https://doi.org/10.3389/fnhum.2013.00723

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, *146*, 240–244. https://doi.org/10.1016/j.cognition.2015.09.016

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2021). Semantic richness and density effects on language production: Electrophysiological and behavioral evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(3), 508–517. https://doi.org/10.1037/xlm0000940v

Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00011

Rico Duarte, L., & Robert, C. (2014). Semantic richness influences naming pictures in old age and in Alzheimer's disease. *Journal of Neurolinguistics*, *31*, 55–68. https://doi.org/10.1016/j.jneuroling.2014.06.004

Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, *24*, 121–133.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*(1), 205–235. https://doi.org/10.1037/0033-295X.111.1.205

Rogerson, P. (2011). *Statistical methods for geography*. SAGE Publications, Ltd. https://doi.org/10.4135/9781849209953

Rossiter, C., & Best, W. (2013). “Penguins don’t fly”: An investigation into the effect of typicality on picture naming in people with aphasia. *Aphasiology*, *27*(7), 784–798. https://doi.org/10.1080/02687038.2012.751579

Schwartz, M. F., & Dell, G. S. (2010). Case series investigations in cognitive neuropsychology. *Cognitive Neuropsychology*, *27*(6), 477–494. https://doi.org/10.1080/02643294.2011.574111

Taylor, K. I., Devereux, B. J., Acres, K., Randall, B., & Tyler, L. K. (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. *Cognition*, *122*(3), 363–374. https://doi.org/10.1016/j.cognition.2011.11.001

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(6), 244–252. https://doi.org/10.1016/S1364-6613(00)01651-X

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Woollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia*, *46*(10), 2503–2514. https://doi.org/10.1016/j.neuropsychologia.2008.04.005

Ethics approvals are removed from Open Access version as they may contain sensitive/confidential content.