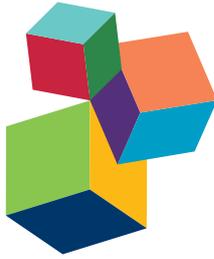


# ACCESSING CONCEPTUAL REPRESENTATIONS FOR SPEAKING

EDITED BY: Peter Indefrey and Ian FitzPatrick  
PUBLISHED IN: Frontiers in Psychology





# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-011-4

DOI 10.3389/978-2-88945-011-4

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

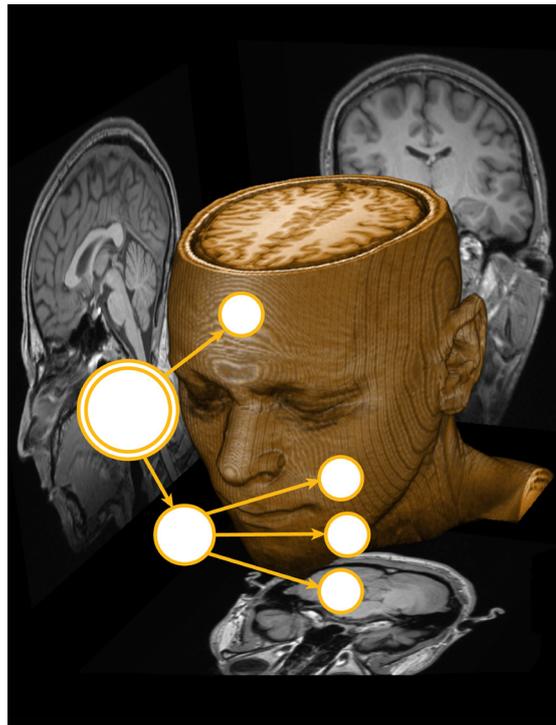
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# ACCESSING CONCEPTUAL REPRESENTATIONS FOR SPEAKING

Topic Editors:

**Peter Indefrey**, Heinrich Heine University Düsseldorf, Germany

**Ian FitzPatrick**, Heinrich Heine University Düsseldorf, Germany



Conceptual representations as recursive attribute-value structures or frames.  
Cover artwork created by Frauke Hellwig

For speaking, words in the lexicon are somehow activated from conceptual representations but we know surprisingly little about how this works precisely. Which of the attributes of the concept DOG (e.g. BARKS, IS WALKED WITH A LEASH, CARNIVORE, ANIMATE) have to be activated in a given situation to be able to select the word 'dog'? Are there things we know about dogs that are always activated for naming and others that are only activated in certain contexts or even never? To date, investigations on lexical access in speaking have largely focused on the effects of distractor nouns on the naming latency of a target noun. We have learned that distractors from the same semantic category (e.g. 'cat') hinder naming, but associatively related

distractors ('leash') may facilitate or hinder naming. However, associatively related words can have all kinds of semantic relationships to a target word, and, with few exceptions, the effects of specific semantic relationships other than membership in the same category as the target concept have not been systematically investigated.

This special issue aims at moving forward towards a more detailed account of how precisely conceptual information is used to access the lexicon in speaking and what corresponding format of conceptual representations needs to be assumed.

**Citation:** Indefrey, P., FitzPatrick, I., eds. (2016). *Accessing Conceptual Representations for Speaking*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-011-4

# Table of Contents

- 06 Editorial: Accessing Conceptual Representations for Speaking**  
Ian FitzPatrick and Peter Indefrey

## **Section 1: Conceptual Processing for Sentence Formulation**

- 08 What the eyes say about planning of focused referents during sentence formulation: a cross-linguistic investigation**  
Lesya Y. Ganushchak, Agnieszka E. Konopka and Yiya Chen
- 19 Agent-patient similarity affects sentence structure in language production: evidence from subject omissions in Mandarin**  
Yaling Hsiao, Yannan Gao and Maryellen C. MacDonald

## **Section 2: Relationships between Conceptual and Lexical Activation in Monolingual and Bilingual Speakers**

- 31 Different Loci of Semantic Interference in Picture Naming vs. Word-Picture Matching Tasks**  
Denise Y. Harvey and Tatiana T. Schnur
- 50 Investigating the flow of information during speaking: the impact of morpho-phonological, associative, and categorical picture distractors on picture naming**  
Jens Bölte, Andrea Böhl, Christian Dobel and Pienie Zwitserlood
- 66 The roles of shared vs. distinctive conceptual features in lexical access**  
Harrison E. Vieth, Katie L. McMahon and Greig I. de Zubicaray
- 78 Processing different kinds of semantic relations in picture-word interference with non-masked and masked distractors**  
Markus F. Damian and Katharina Spalek
- 91 Semantic gradients in picture-word interference tasks: is the size of interference effects affected by the degree of semantic overlap?**  
James Hutson and Markus F. Damian
- 105 Bilinguals implicitly name objects in both their languages: an ERP study**  
Katie Von Holzen and Nivedita Mani
- 117 EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations**  
João M. Correia, Bernadette Jansma, Lars Hausfeld, Sanne Kikkert and Milene Bonte

### **Section 3: Activation of Conceptual Attributes**

**127** *The role of the sound of objects in object identification: evidence from picture naming*

Claudio Mulatti, Barbara Treccani and Remo Job

**132** *Long-term repetition priming and semantic interference in a lexical-semantic matching task: tapping the links between object names and colors*

Toby J. Lloyd-Jones and Kazuyo Nakabayashi



# Editorial: Accessing Conceptual Representations for Speaking

Ian FitzPatrick<sup>1,2\*</sup> and Peter Indefrey<sup>1,2</sup>

<sup>1</sup> Institut für Sprache und Information, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, <sup>2</sup> Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands

**Keywords:** language production, conceptual representation, semantics, lexical access, conceptual attributes

## The Editorial on the Research Topic

### Accessing Conceptual Representations for Speaking

Systematic investigations into the role of semantics in the speech production process have remained elusive. This special issue aims at moving forward toward a more detailed account of how precisely conceptual information is used to access the lexicon in speaking and what corresponding format of conceptual representations needs to be assumed. The studies presented in this volume investigated effects of conceptual processing on different processing stages of language production, including sentence formulation, lemma selection, and word form access.

## CONCEPTUAL PROCESSING FOR SENTENCE FORMULATION

Using an eye-tracking paradigm in which participants are prompted to describe pictures of two-character transitive events, Ganushchak et al. show that contextually new referents are fixated with priority over contextually old (i.e., given) referents. The time course of the contextual effects on gaze patterns suggests that contextual information might well be taken into account during sentence formulation. Hsiao et al. present data from a sentence production task and a corpus study that show that speakers of Mandarin Chinese are more prone to omitting subject pronouns in their utterances when the subject and object of the sentence are conceptually similar (e.g., both animate or both inanimate) than when they are conceptually dissimilar.

## RELATIONSHIPS BETWEEN CONCEPTUAL AND LEXICAL ACTIVATION IN MONOLINGUAL AND BILINGUAL SPEAKERS

The majority of studies aimed at gaining further insights into classic distractor effects. Harvey and Schnur investigated semantic interference in picture naming and word–picture matching. Using a blocked-cyclical paradigm they show that semantic interference in naming generalizes to novel objects, but semantic interference in word–picture matching does not. This is taken as evidence that semantic interference effects in naming and word–picture matching arise at different processing stages. Naming novel items that corresponded to semantic categories that had been previously encountered in word–picture matching induced semantic interference. The latter result suggests a common origin of semantic interference across tasks.

Bölte et al. investigated the origin of semantic interference effects in the picture–picture paradigm. Participants named pictures of German compound words which were accompanied by categorically or associatively related distractor objects. Categorically related distractors facilitated naming at SOAs at which semantic processing is expected (in this case +200). The authors argue that the absence of semantic interference means that such distractors activate their conceptual–semantic information but do not activate the corresponding lemma.

## OPEN ACCESS

### Edited and reviewed by:

Manuel Carreiras,  
Basque Center on Cognition, Brain  
and Language, Spain

### \*Correspondence:

Ian FitzPatrick  
ian@ianfitzpatrick.eu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 21 July 2016

**Accepted:** 02 August 2016

**Published:** 18 August 2016

### Citation:

FitzPatrick I and Indefrey P (2016)  
Editorial: Accessing Conceptual  
Representations for Speaking.  
Front. Psychol. 7:1216.  
doi: 10.3389/fpsyg.2016.01216

Vieth et al. investigated semantic interference from distinctive features. Their first experiment showed no evidence that distractors that differed from target items on a distinctive feature (e.g., for HORSE-/zebra/the feature stripes) were processed differently from semantically matched distractors with no distinctive feature differences (e.g., HORSE-/donkey/). Further experiments showed that distractors denoting visible parts of target objects that are also found in other objects (e.g., GOAT—tail) slowed down naming of target items. The authors argue that this reflects competition from semantically related items (e.g., other animals with tails).

Damian and Spalek used a picture–word-interference paradigm with distractors that were either unrelated, categorically related, associatively related, or both categorically and associatively related. In addition the authors manipulated the visibility of distractors by presenting them in between forward and backward masks. Results replicate earlier (Finkbeiner and Caramazza, 2006; Dhooge and Hartsuiker, 2010) reports of semantic facilitation (rather than inhibition) for masked distractors. Importantly, however, the picture–word-interference effect did not seem to depend on individual subject differences in the ability to recognize the masked distractors. The authors take these results as more in line with competition threshold accounts (e.g., Piai et al., 2012) for picture–word interference rather than response exclusion accounts (Finkbeiner and Caramazza, 2006; Dhooge and Hartsuiker, 2010).

Hutson and Damian tested a prediction of the response exclusion account of the picture–word-interference effect, namely that for semantically closely related items, priming counteracts buffer-based interference. They found no evidence of degree of semantic relatedness in picture–word-interference. This result, they argue, is difficult to reconcile with either response exclusion accounts (which would need to abandon the notion of conceptual priming from semantically related distractors) or competitive accounts (which would need to postulate opposing effects of conceptual priming and semantic interference canceling each other out).

Two studies investigated relationships between conceptual and word form activation in bilingual speakers. Von Holzen and Mani show that bilinguals implicitly generate labels for pictures simultaneously in their first and second languages. Targets

preceded by phonologically related pictures showed lower N400 effects irrespective of whether the phonological relationship was within or between languages. This implies that the non-selected (non-target language) lemma can send activation cascading forward to the phonological level. Correia et al. studied the reverse flow of activation. Using multivariate pattern analysis of EEG data, they show that in bilingual listeners language invariant semantic representations can be decoded around 550 ms following the onset of a spoken word.

## ACTIVATION OF CONCEPTUAL ATTRIBUTES

Finally, two studies investigated the role of attribute retrieval in naming. Mulatti et al. show that white noise interferes with naming pictures of objects with typical sounds but not with objects without typical sounds. This suggests that an object's sound attribute is used during lemma retrieval. Lloyd-Jones and Nakabayashi examined the retrieval of object color information using a picture naming and semantic matching task. Their results suggest differential retrieval of color information for object names and object shapes.

## CONCLUSION

It becomes clear in this volume that effects of conceptual processing extend beyond the conceptual level and can affect many levels of processing. The range of conceptual relationships that are explored is just beginning to be expanded beyond categorical and associative relationships.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

This research was funded by the Deutsche Forschungsgemeinschaft (DFG) Collaborative Research Centre (CRC) 991.

## REFERENCES

- Dhooge, E., and Hartsuiker, R. J. (2010). The distractor frequency effect in pictureword interference: evidence for response exclusion. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 878–891. doi: 10.1037/a0019128
- Finkbeiner, M., and Caramazza, A. (2006). Now you see it, now you don't: on turning semantic interference into facilitation in a Stroop-like task. *Cortex* 6, 790–796. doi: 10.1016/S0010-9452(08)70419-2
- Piai, V., Roelofs, A., and Schriefers, H. (2012). Distractor strength and selective attention in picture-naming performance. *Mem. Cogn.* 40, 614–627. doi: 10.3758/s13421-011-0171-3

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 FitzPatrick and Indefrey. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# What the eyes say about planning of focused referents during sentence formulation: a cross-linguistic investigation

Lesya Y. Ganushchak<sup>1,2,3\*</sup>, Agnieszka E. Konopka<sup>4</sup> and Yiya Chen<sup>1,3</sup>

<sup>1</sup> Leiden University Centre for Linguistics, Leiden, Netherlands

<sup>2</sup> Education and Child Studies, Faculty of Social and Behavioral Sciences, Leiden University, Leiden, Netherlands

<sup>3</sup> Leiden Institute for Brain and Cognition, Leiden, Netherlands

<sup>4</sup> Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

## Edited by:

Ian FitzPatrick, Heinrich Heine  
Universität Düsseldorf, Germany

## Reviewed by:

Christoph Scheepers, University of  
Glasgow, UK

Susanne Brouwer, Utrecht  
University, Netherlands

## \*Correspondence:

Lesya Y. Ganushchak, Education and  
Child Studies, Faculty of Social and  
Behavioral Sciences, Leiden  
University, Pieter de la Court,  
Gebouw Postbus 9555, 2300 RB,  
Leiden, Netherlands  
e-mail: lganushchak@gmail.com

This study investigated how sentence formulation is influenced by a preceding discourse context. In two eye-tracking experiments, participants described pictures of two-character transitive events in Dutch (Experiment 1) and Chinese (Experiment 2). Focus was manipulated by presenting questions before each picture. In the Neutral condition, participants first heard “*What is happening here?*” In the Object or Subject Focus conditions, the questions asked about the Object or Subject character (*What is the policeman stopping? Who is stopping the truck?*). The target response was the same in all conditions (*The policeman is stopping the truck*). In both experiments, sentence formulation in the Neutral condition showed the expected pattern of speakers fixating the subject character (*policeman*) before the object character (*truck*). In contrast, in the focus conditions speakers rapidly directed their gaze preferentially only to the character they needed to encode to answer the question (the *new*, or *focused*, character). The timing of gaze shifts to the new character varied by language group (Dutch vs. Chinese): shifts to the new character occurred earlier when information in the question can be repeated in the response with the same syntactic structure (in Chinese but not in Dutch). The results show that discourse affects the timecourse of linguistic formulation in simple sentences and that these effects can be modulated by language-specific linguistic structures such as parallels in the syntax of questions and declarative sentences.

**Keywords:** focus planning, discourse context, sentence formulation, incrementality, eye-tracking

## INTRODUCTION

To produce a sentence, speakers must prepare a preverbal message and then encode it linguistically. These processes are assumed to proceed incrementally (e.g., Kempen and Hoenkamp, 1987). However, the amount of linguistic information that speakers prepare in advance of speaking can be highly variable (e.g., Konopka, 2012; Konopka and Meyer, 2014). While much work has been done on formulation of individual sentences produced out of context, a largely neglected area of research is how sentences are planned as a function of the discourse context in which they are produced. The aim of the present project is to investigate the timecourse of online sentence formulation within one particular discourse context—i.e., as a function of changes in informational focus.

Specifically, we consider formulation of simple event descriptions like *The policeman is stopping the truck* (Figure 1) in response to informational wh-questions. For examples, questions like “*What is the policeman stopping?*” provide a discourse context that establishes one referent in the event as contextually *old* information and the referent that is being asked about as *new*, and therefore *focused*, information (Gussenhoven, 2007). Thus,

in answer to this question, the typical answer (*The policeman is stopping the truck*) includes *policemen* as given information and *truck* as new (focused) information. In contrast, if the question is *Who is stopping the truck?*, the typical answer (*The policeman is stopping the truck*) includes *policeman* as the focused referent, indicating that it is the policeman, rather than a person of another profession, who is stopping the *truck*.

The issue we address here is to what extent focus may affect the way utterances are planned online. Sentence formulation is normally investigated by asking speakers to describe pictures of events (Figure 1) while their gaze and speech are recorded (Griffin and Bock, 2000; Bock et al., 2004; Griffin, 2004; Meyer and Lethaus, 2004; Gleitman et al., 2007; Kuchinsky and Bock, 2010; Konopka, 2013, 2014; Ganushchak et al., 2014; Konopka and Meyer, 2014; Van de Velde et al., 2014). On Griffin and Bock’s (2000) account, formulation begins with an apprehension phase (0–400 ms after picture onset) during which speakers encode the “gist” of the event. During this phase, fixations to the subject and object characters in the event do not differ from each other reliably. Event apprehension is then followed by a longer phase of linguistic encoding that lasts until the end of

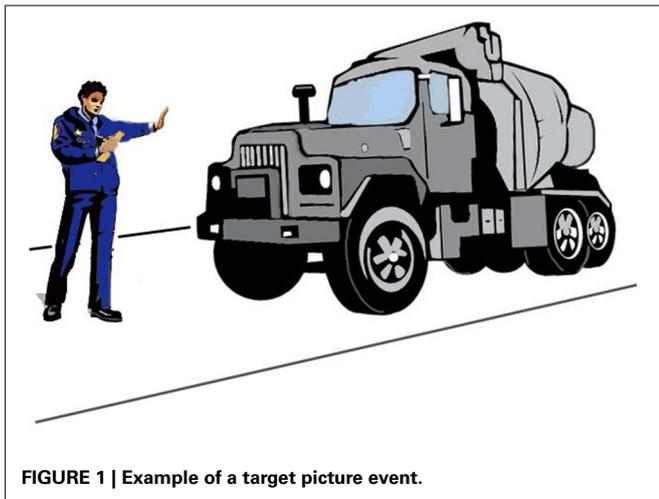


FIGURE 1 | Example of a target picture event.

articulation. In this time window (400 ms until the end of speech), participants normally look at characters in the display in the order of mention. Viewing times on a character and gaze shifts from one character to another after 400 ms are thus expected to vary with the ease of encoding each character (e.g., easy-to-name characters are fixated for less time than harder-to-name characters; see Konopka and Meyer, 2014; Van de Velde et al., 2014).

To compare formulation of sentences with and without focus, eye-tracked participants were asked to describe pictures shown on a computer screen in their native language: Dutch (Experiment 1) or Chinese (Experiment 2). Focus was manipulated by means of questions that preceded each picture. In the Neutral condition, participants were asked a question that was neutral with respect to discourse focus: “*What is happening here?*” In the remaining two conditions, the questions changed the discourse focus of the expected target event description. In the Subject Focus condition, participants were asked about the subject character (*Who is stopping the truck?*). In the Object Focus condition, participants were asked about the object character (*What is the policeman stopping?*). The expected target response had the same structure and content in all conditions (*The policeman is stopping the truck*).

How might discourse focus influence formulation? Differences in planning of the target responses were evaluated by comparing speakers’ eye movements to the two event characters prior to speech onset. On the one hand, it is possible that discourse focus does not immediately influence the timecourse of formulation. If so, viewing times for the subject and object characters should not differ across conditions: speakers should consistently fixate the subject character first and then direct their attention and gaze to the object character, reflecting order of mention. This outcome would be expected on the basis of research showing very tight gaze-speech coordination during formulation (e.g., Griffin and Bock, 2000), even when speakers talk about “old” or previously inspected referents (e.g., Meyer et al., 2004). On the other hand, if sentence formulation is sensitive to changes in information structure at the discourse level, then changes in the *old/new* (or *focused/unfocused*) status of event characters should influence the relative allocation of attention to these characters. In

this case, viewing patterns in the Subject and Object focus conditions should differ from the Neutral Focus condition: speakers should direct fewer fixations to the character that was mentioned in the question (the *old* character) but should preferentially fixate the character needed to answer the question (the *new*, or *focused*, character). Thus, in the Object Focus condition, speakers should rapidly direct their gaze to the object character, and in the Subject Focus condition, they should direct their gaze to the subject character.

We also test whether changes in gaze patterns are modulated exclusively by discourse context or if they also depend on the ease of encoding the target sentences linguistically. The questions in the Object and Subject Focus conditions mention one of the event characters, which establishes this character as *old* information in the discourse and provides speakers with a referential term they can use in their responses. Thus, by definition, the questions in the Focus conditions facilitate conceptual and linguistic planning of the *old* character. However, in addition to recognizing the *old* character in the event, speakers must also generate a suitable sentence structure to produce a full response to the preceding question. To test whether formulation additionally depends on the ease of linguistic encoding in the Focus conditions, Experiments 1 and 2 compare sentence formulation in the same task with speakers of two languages that differ in the word order of *wh*-questions: Dutch and Chinese. Dutch requires *wh*-fronting (*Who is stopping the truck? What is the policeman stopping?*), while Chinese is known for *in-situ wh*-questions (i.e., *wh*-words do not undergo movement but remain in the same surface syntactic position as the constituent being questioned; Cheng, 2009). This is illustrated in the following examples:

Subject focus: 誰在停止卡車 (*Who is stopping the truck?*)

Object focus: 警察在停止什麼 (The policeman is stopping what?)

So, the two languages have the same surface word order when the focus of the *wh*-question is on the subject character but very different orders when the focus of the *wh*-question is on the object character. Consequently, when prompted by an object-specific *wh*-question (i.e., Object Focus question), Chinese speakers are provided with linguistic material that they can repeat verbatim in their response without having to change the syntactic constituent order provided in the *wh*-question, while Dutch speakers need to generate a response with a word order different from that of the preceding question. If sentence formulation is sensitive to the amount of information provided in the preceding discourse context even at the syntactic structural level, we should observe a cross-linguistic difference in sentence formulation after Object Focus questions in Experiment 1 (Dutch) and Experiment 2 (Chinese): since Chinese speakers can “reuse” linguistic material from the question without syntactic restructuring when preparing their response, they may begin shifting their gaze to the *new* object character earlier than speakers of Dutch (who, besides encoding the object character, must also generate a suitable sentence structure).

Importantly, we test how early differences in fixation patterns to the subject and object characters emerge in the Object

and Subject focus conditions compared to the Neutral condition. Overall, differences occurring immediately after picture onset (0–400 ms, i.e., a window arguably corresponding to event apprehension) would indicate that focus information has an early effect on formulation of the target utterance—beginning during the encoding of the preverbal message. In contrast, differences across conditions emerging after 400 ms would indicate that focus information influences primarily the timing of linguistic encoding, after speakers have encoded the gist of the event they are about to describe.

## EXPERIMENT 1. FOCUS PLANNING: DUTCH

### METHODS

#### Participants

Thirty native speakers of Dutch, all students at Leiden University, participated in the experiment (24 women; age range 17–23 years). All participants were students at Leiden University. The study was conducted in accord with APA standards for ethical treatment of participants and was approved by the ethical committee board of Leiden University. Participants gave written informed consent prior to participating and received a small financial reward.

#### Materials

The stimulus lists consisted of 178 colored pictures displaying simple events (**Figure 1**). There were 58 target pictures of transitive events, 116 fillers, and 4 practice pictures. In the target pictures, the subject character was on the left in 77% of the cases<sup>1</sup>. Discourse focus was manipulated by means of questions presented before each picture.

- (A) Neutral question:  
*Wat gebeurt hier?* (What is happening here?)
- (B) Object Focus question:  
*Wat stopt de politiemans?* (What is the policeman stopping?)
- (C) Subject Focus question:  
*Wie stopt de vrachtauto?* (Who is stopping the truck?)

Modal target sentence: *De politiemans laat een vrachtauto stoppen* (The policeman is stopping the truck).

All questions were recorded by a native Dutch male speaker and were presented auditorily prior to picture onset.

#### Design and procedure

Lists of stimuli were created to counterbalance question type across target pictures. Each target picture occurred in Focus condition on different lists, so each participant saw each picture only once.

<sup>1</sup>We cannot say for sure whether the effects in the Neutral condition are due to “order of mention” or to a general left-to-right scanning preference. In the current study, we saw a stronger tendency for speakers to fixate the two characters in the order of mention when the agent appeared on the left hand-side of the screen. However, by comparison, we see very strong effects of the question manipulation on formulation. It is also important to note that all pictures appeared in all of the conditions, so the differences we see between conditions cannot be attributed to the agent placement.

Target pictures were interspersed among filler pictures, with at least two filler pictures separating any two target trials in each list. The fillers showed similar one-character and two-character events. However, the questions preceding filler pictures varied: e.g., the questions asked participants to name the color of an object, or to count how many of a given item appeared in the picture.

Participants were seated in a sound-proof room. Eye movements were recorded with an EyeLink 1000 eye-tracker (SR Research Ltd.; 500 Hz sampling rate). Eye calibration was done at the beginning of the experiment, using a 9-point calibration procedure. Participants first heard a question (presented through headphones). Experimenter then clicked with the mouse after completion of the question to proceed to the picture trials. Picture trials began with a fixation point presented at the top of the screen (drift correction): participants had to fixate the fixation point and press the space bar to display the picture. They were instructed to describe each picture with one sentence and were not under time pressure to produce the response. The experimenter clicked with the mouse when the participant finished speaking. On average, the pictures were displayed on the screen for 4191 ms ( $SD = 850$  ms). The task started with four practice trials.

#### Scoring and data analysis

Target sentences were scored as correct if participants used an active SVO structure. Trials where participants used a different structure (e.g., passive sentences) or made corrections during the description were excluded from analysis (7% of the data; Subject Focus: 1.1%; Object Focus: 1.4%; Neutral: 4.6%; error rates were lower than in other reported studies, largely because the experimental manipulations successfully constrained structure choice on target trials to SVO sentences).

Interest areas were drawn around each character in the target pictures (allowing a 2–3 cm margin around each character). Trials in which the first fixation was within the subject or object character interest area instead of the fixation point were also removed from the analyses (1% of the data). This left 883 trials for analysis.

Analyses were carried out a) on speech onsets to assess differences across conditions with respect to encoding difficulty in sentences with *new* and *old* subject and object characters, and b) on subject-directed fixations to assess differences in the timecourse of formulation across conditions.

Speech onsets were first log-transformed to remove the intrinsic positive skew and non-normality of the distribution, and then submitted to mixed-effects model analyses with participants and items as random effects (Baayen et al., 2008). Focus Location (Neutral, Object Focus, and Subject Focus) was entered as a fixed effect. By-subject and by-item random slopes for Focus Location and random intercepts were also included. Onsets in the three Focus Location conditions were compared with two contrasts using treatment coding. The first contrast compared the Neutral condition against the Object Focus condition; the second contrast compared the Neutral condition against the Subject Focus condition. Both contrasts thus assess how planning a sentence in response to a question that mentions one of the event

characters changes response latencies relative to the neutral condition. Next, a separate analysis was run with new contrasts to compare response latencies in the Subject and Object Focus conditions against one another.

For the timecourse analyses, the distribution of subject-directed fixations in sentences produced in the three conditions was compared with by-participant ( $\beta_1$ ) and by-item ( $\beta_2$ ) quasi-logistic regressions (Barr, 2008). Consistent with earlier work and based on visual inspection of the distributions, we selected three time windows (0–400, 400–800, and 800–1600 ms) for analysis. The first time window arguably corresponds to a period of event apprehension (Griffin and Bock, 2000; Konopka and Meyer, 2014), while the second and third time windows include the rise and fall of fixations to the subject character before speech onset in the Neutral condition (within each of these windows, changes of fixation proportions show a relatively linear pattern as a function of time). Fixations were aggregated into a series of 200 ms time bins for each participant in the by-participant analysis and each item in the by-item analysis in each condition. The dependent variable in each time bin was an empirical logit indexing the likelihood of speakers fixating the subject characters out of the total number of fixations observed in that time bin.

The models included Time Bin and Focus Location (Neutral, Subject Focus, and Object Focus) as fixed effects, and tested for interactions between these variables. All models included random by-participant and by-item random intercepts and slopes for the Time and Focus Location variables. For interactive models, the random effects structure included the interaction between Time and Focus Location; in additive models, the models included additive random slopes for Time and Focus Location. Main effects in these analyses indicate differences across conditions in the first bin of each window, while interactions with Time show how fixation patterns changed over the remaining bins in that time window. Thus, when we refer to an effect (a main effect) present at 0–200, 400–600, or at 800–1000 ms, we are describing a difference between conditions present at the first 200 ms of a time window. Interactions between the Focus Location factor and the Time factor then show how the pattern of fixations changed in the remaining time window (200–400, 600–800, and 1000–1600 ms, respectively). The log-likelihood ratio test ( $\chi^2$ ) was used to compare model fit in interactive and additive models, and thus test whether interactions with the Time variable significantly improved model fit (a reliable difference in this comparison indicates a better fit for the interactive model than the additive model). All interactions reported below were reliable by this criterion at  $p < 0.01$ .

As in the analyses of speech onsets, fixations in the three Focus Location conditions were compared with two contrasts, and the Object and Subject Focus conditions were compared against each other in a separate analysis.

## RESULTS

### Speech onsets

Participants started speaking significantly later in the Neutral condition than in the Object and Subject focus conditions ( $\beta = -0.24$ ,  $SE = 0.04$ ;  $t < -6$ ;  $\beta = -0.17$ ,  $SE = 0.04$ ;

**Table 1 | Mean response latencies in ms (and standard errors) per condition in Experiment 1 (Dutch) and in Experiment 2 (Chinese).**

	Object focus	Subject focus	Neutral
Experiment 1	1550 (412)	1555 (265)	2104 (623)
Experiment 2	1139 (452)	1610 (735)	1822 (588)

$t < -4$ ), for the two contrasts respectively; see **Table 1** for means). The difference in speech onset latencies between the Object Focus and Subject Focus conditions was not significant ( $t < 1$ ).

### Timecourse of sentence formulation

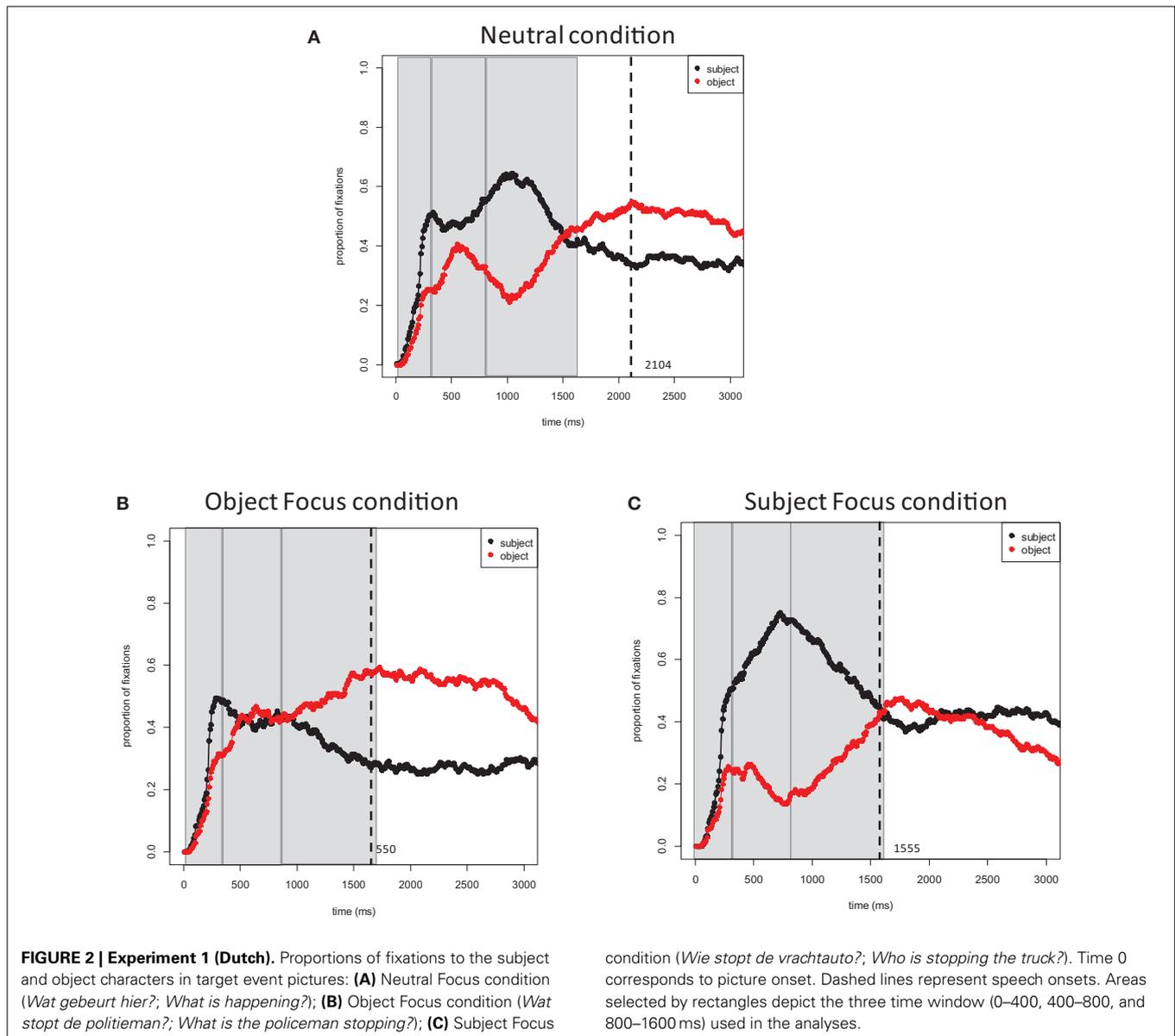
**Figure 2** plots the proportions of fixations to the subject and object characters in target pictures across conditions. **Figure 4A** then plots the proportions of fixations to the subject character in the target pictures across all three conditions. Results of all timecourse analyses are listed in **Table 2** (the by-participants and by-items analyses provided largely converging results and are thus not discussed separately).

**0–400 ms.** In all conditions, speakers rapidly directed their gaze to the subject character in the event within 400 ms of picture onset. All main effects and interactions in this time window did not reach significance (**Table 2A**).

**400–800 ms.** After 400 ms, speakers largely directed their gaze to the subject character in the Neutral condition. The first contrast in this analysis showed a weak difference in fixations to subject characters at the first time bin (i.e., 400–600 ms) in the Neutral condition and Object Focus condition (the effect was reliable in the by-item analysis). The interaction between Focus Location and Time was reliable: in the Neutral condition, speakers quickly directed their gaze to the subject character while in the Object focus condition, fixations to the subject character remained stable. The second contrast in the analysis showed that fixations to the subject character did not differ in the Neutral condition and Subject Focus condition at 400–600 ms. The interaction with Time for this contrast was again significant: speakers directed their gaze preferentially to subject characters in the Subject Focus condition while fixations to subject characters remained stable in the Neutral condition (**Table 2B**).

Comparing the Subject Focus and Object Focus conditions against one another in a separate analysis showed a significant interaction of Focus Location with Time. Thus, as time progressed, fixations to the subject character within this window increased in the Subject Focus condition but not in the Object Focus condition.

**800–1600 ms.** Speakers began shifting their gaze away from the subject character between 800 ms and speech onset. Carrying over from earlier windows, speakers were more likely to fixate subject characters in the Neutral condition than in the Object Focus condition during the first 200 ms of the time window (i.e., 800–1000 ms), but were more likely to fixate subject characters in the Subject Focus condition than in the Neutral condition. The



first contrast in the interaction between Time and Focus Location was significant, showing that fixations to the subject character decreased at a steeper rate in the Object Focus condition than in the Neutral condition. The second contrast in this interaction was also significant: fixations to subject characters decreased at a steeper rate in the Neutral condition than in the Subject Focus condition (Table 2C).

Finally, the comparison between Subject Focus and Object Focus conditions showed that there were more fixations to subject characters in the Subject Focus condition than in the Object Focus condition at the first 200 ms of the time window (i.e., 800–1000 ms). The interaction with Time was also significant: fixations to subject characters decreased at a steeper rate in the Subject Focus condition than in the Object Focus condition.

## DISCUSSION

Speakers' gaze patterns showed large differences in attention allocation to subject and object characters in target events across conditions. The pattern obtained in the Neutral condition replicated earlier findings, showing that participants largely fixate characters in the order of mention: first the subject character (*policeman*) and then the object character (*truck*; Griffin and Bock, 2000). Gaze shifts to the object character occurred well before speech onset.

In contrast, sentence formulation in the Subject Focus and Object Focus conditions was strongly influenced by the preceding discourse context. First, speech onsets were reliably shorter in these conditions than in the Neutral condition, suggesting that partial knowledge of the characters and of the relationship between characters in the upcoming event facilitated planning.

**Table 2 | Results of by-participant ( $\beta_1$ ) and by-item ( $\beta_2$ ) quasi-logistic regressions carried out over three time windows in Experiment 1 (Dutch) and Experiment 2 (Chinese).**

	Experiment 1		Experiment 2	
	Main effect of Focus location	Interaction of Focus Location with Time	Main effect of Focus location	Interaction of Focus Location with Time
<b>(A) 0–400 ms</b>				
1. Neutral vs. Object Focus	$\beta_1 = -0.12, SE = 0.24, t = -0.51$ $\beta_2 = -0.15, SE = 0.16, t = -0.92$	$\beta_1 = -0.32, SE = 0.91, t = -0.35$ $\beta_2 = -0.01, SE = 0.62, t = -0.01$	$\beta_1 = -0.13, SE = 0.21, t = -0.61$ $\beta_2 = -0.07, SE = 0.17, t = -0.42$	$\beta_1 = 0.07, SE = 0.74, t = 0.09$ $\beta_2 = -0.10, SE = 0.59, t = -0.17$
2. Neutral vs. Subject Focus	$\beta_1 = -0.16, SE = 0.24, t = -0.69$ $\beta_2 = -0.12, SE = 0.16, t = -0.72$	$\beta_1 = 1.26, SE = 0.88, t = 1.43$ $\beta_2 = 1.17, SE = 0.61, t = 1.91^\dagger$	$\beta_1 = -0.37, SE = 0.21, t = -1.79^\dagger$ $\beta_2 = -0.27, SE = 0.18, t = -1.63$	$\beta_1 = 0.85, SE = 0.77, t = 1.11$ $\beta_2 = 0.86, SE = 0.64, t = 1.33$
3. Object vs. Subject Focus	$\beta_1 = -0.04, SE = 0.21, t = -0.17$ $\beta_2 = 0.01, SE = 0.14, t = 0.09$	$\beta_1 = 0.86, SE = 0.70, t = 1.22$ $\beta_2 = 0.62, SE = 0.55, t = 1.13$	$\beta_1 = -0.12, SE = 0.17, t = -0.71$ $\beta_2 = -0.10, SE = 0.13, t = -0.76$	$\beta_1 = 0.38, SE = 0.61, t = 0.62$ $\beta_2 = 0.47, SE = 0.49, t = 0.95$
<b>(B) 400–800 ms</b>				
1. Neutral vs. Object Focus	$\beta_1 = 0.57, SE = 0.30; t = 1.91^\dagger$ $\beta_2 = 0.64, SE = 0.29; t = 2.18^*$	$\beta_1 = -1.91, SE = 0.47; t = -4.07^{***}$ $\beta_2 = -2.01, SE = 0.48; t = -4.17^{***}$	$\beta_1 = 1.16, SE = 0.42, t = 2.74^{**}$ $\beta_2 = 1.14, SE = 0.35, t = 3.27^{***}$	$\beta_1 = -3.23, SE = 0.70, t = -4.59^{***}$ $\beta_2 = -3.21, SE = 0.58, t = -5.50^{***}$
2. Neutral vs. Subject Focus	$\beta_1 = -0.03, SE = 0.33; t = -0.09$ $\beta_2 = -0.12, SE = 0.29; t = -0.41$	$\beta_1 = 1.53, SE = 0.48; t = 3.21^{***}$ $\beta_2 = 1.73, SE = 0.47; t = 3.68^{***}$	$\beta_1 = -1.12, SE = 0.39, t = -2.90^{**}$ $\beta_2 = -1.14, SE = 0.43, t = -2.64^{**}$	$\beta_1 = 3.34, SE = 0.63, t = 5.30^{***}$ $\beta_2 = 3.44, SE = 0.68, t = 5.05^{***}$
3. Object vs. Subject Focus	$\beta_1 = -0.31, SE = 0.26, t = -1.19$ $\beta_2 = -0.36, SE = 0.23, t = -1.59$	$\beta_1 = 1.74, SE = 0.38, t = 4.53^{***}$ $\beta_2 = 1.84, SE = 0.38, t = 4.84^{***}$	$\beta_1 = -1.15, SE = 0.34, t = -3.38^{***}$ $\beta_2 = -1.14, SE = 0.31, t = -3.66^{***}$	$\beta_1 = 3.29, SE = 0.57, t = 5.80^{***}$ $\beta_2 = 3.31, SE = 0.50, t = 6.60^{**}$
<b>(C) 800–1600 ms</b>				
1. Neutral vs. Object Focus	$\beta_1 = -1.60, SE = 0.35, t = -4.57^{***}$ $\beta_2 = -1.63, SE = 0.32, t = -5.13^{***}$	$\beta_1 = 0.63, SE = 0.29; t = 2.18^*$ $\beta_2 = 0.70, SE = 0.22; t = 2.76^{**}$	$\beta_1 = -1.93, SE = 0.46, t = -4.14^{***}$ $\beta_2 = -1.93, SE = 0.47, t = -4.14^{***}$	$\beta_1 = 0.54, SE = 0.39, t = 1.40$ $\beta_2 = 0.54, SE = 0.37, t = 1.40$
2. Neutral vs. Subject Focus	$\beta_1 = 1.60, SE = 0.32; t = 5.08^{***}$ $\beta_2 = 1.67, SE = 0.25; t = 5.91^{***}$	$\beta_1 = -0.84, SE = 0.27; t = -3.14^{**}$ $\beta_2 = -0.92, SE = 0.22; t = -4.17^{***}$	$\beta_1 = 2.19, SE = 0.39, t = 5.54^{***}$ $\beta_2 = 2.19, SE = 0.40, t = 5.54^{***}$	$\beta_1 = -1.2, SE = 0.29, t = -3.76^{***}$ $\beta_2 = -1.12, SE = 0.29, t = -3.76^{***}$
3. Object vs. Subject Focus	$\beta_1 = 1.60, SE = 0.27; t = 5.91^{***}$ $\beta_2 = 1.66, SE = 0.24; t = 6.92^{***}$	$\beta_1 = -0.74, SE = 0.23; t = -3.22^{***}$ $\beta_2 = -0.81, SE = 0.19; t = -4.23^{***}$	$\beta_1 = 2.04, SE = 0.34, t = 5.92^{***}$ $\beta_2 = 2.03, SE = 0.34, t = 5.92^{***}$	$\beta_1 = -0.81, SE = 0.27, t = -2.99^{**}$ $\beta_2 = -0.81, SE = 0.27, t = -2.99^{**}$

The table lists statistics for the main effects of the individual Focus comparisons and for interactions of the Focus condition with Time (all main effects for the Time variable were reliable, and are thus not listed). Two analyses were carried out on subject-directed fixations within each time window: the first two contrasts (1, 2) show results from the analysis performed with Focus Location as a three-factor variable (Neutral, Object, Subject), while the third contrast (3) shows results from the analysis performed with Focus Location as a two-factor variable (Object, Subject). Significance for individual effects was determined by treating t-values as if they were drawn of the normal distribution (see Barr, 2008).

$^\dagger p < 0.10$ .

\* $p < 0.05$ .

\*\* $p < 0.01$ .

\*\*\* $p < 0.001$ .

Second and more importantly, the distribution of fixations to the two characters across conditions was strongly influenced by the preceding discourse questions. Speakers had a strong preference for fixating the contextually *new* character with priority, both when this character was the sentence subject and when it was the sentence object. In the Object Focus condition, participants looked briefly at the subject character and shifted their gaze to the object character shortly after 400 ms of the picture onset, while in the Subject Focus condition, participants looked longer at the subject character and shifted their gaze to the object character only about 1600 ms after picture onset. Thus, even though the propositional content and the surface form of the target sentence were held constant across conditions, gaze-speech coordination during sentence formulation changed with discourse context.

## EXPERIMENT 2. FOCUS PLANNING: CHINESE

### METHODS

#### Participants

Thirty native speakers of Chinese (Northern regions) participated in the experiment (16 women; age range 23–29 years). All participants were students at Leiden University. Research reported in the current manuscript was conducted in accord with APA standards for ethical treatment of participants and was approved by the ethical committee board of Leiden University. Participants gave written informed consent prior to participating in the study and received a small financial reward after the experiment.

#### Materials

The pictures used in this experiment were a subset of the pictures described in Experiment 1. Fifteen target pictures were excluded as they were unlikely to elicit SVO descriptions in Chinese. Thus, in total, there were 129 colored pictures in Experiment 2 (43 target pictures, 82 fillers, and 4 practice pictures). In the target pictures, the subject character was on the left in 74% of the cases. As in Experiment 1, focus was manipulated by means of questions that preceded each picture. All questions were recorded by a native Chinese female speaker.

#### Design, procedure, and data analysis

The design, procedure and analyses were identical to Experiment 1. The target pictures remained on the screen for about 4541 ms ( $SD = 856$  ms). In total, 11% (Subject Focus: 2.6%; Object Focus: 3.3%; Neutral: 4.8%) of all target trials were removed due to erroneous responses and 1% of trials removed because the first fixation was within the subject or object character interest area instead of the fixation point. This left 527 trials for analysis.

## RESULTS

#### Speech onsets

Participants started speaking significantly later in the Neutral condition than in the Object Focus conditions ( $\beta = -0.56$ ,  $SE = 0.07$ ;  $t < -8$ ; see **Table 1** for means). The difference in speech onset latencies between the Neutral and Subject Focus conditions

was not significant ( $t < 1.5$ ). Participants also started speaking later in the Subject Focus conditions than in the Object Focus conditions ( $\beta = 0.34$ ,  $SE = 0.05$ ;  $t > 6$ ).

#### Timecourse of formulation

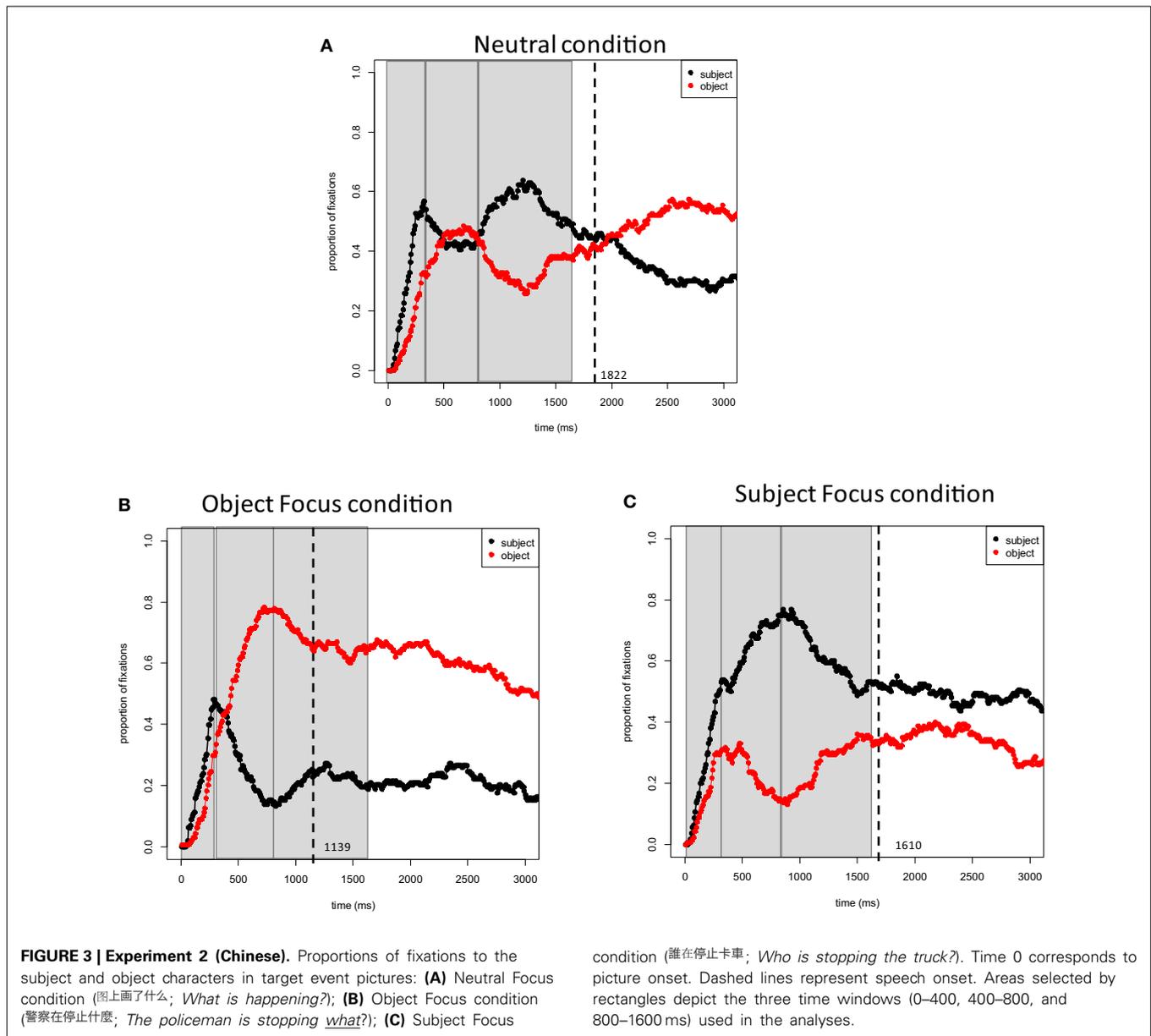
**Figure 3** plots the proportions of fixations to the subject and object characters in target pictures across conditions. **Figure 4B** again plots the proportions of fixations to the subject character in the target pictures across all three conditions. The overall distribution of fixations to the two characters was similar to Experiment 1, with the exception of the Object focus condition. Results of statistical tests are provided in **Table 2**.

**0–400 ms.** In all conditions, speakers rapidly directed their gaze to the subject character in the picture within 400 ms of picture onset. All main effects and interactions in this time window were not significant (**Table 2A**).

**400–800 ms.** Speakers were already more likely to fixate subject characters in the Object Focus condition than in the Neutral condition at the first 200 ms of the time window (i.e., 400–600 ms), which, in turn, had more fixation than in the Subject Focus condition. All interactions with Time were largely consistent with Experiment 1. The first contrast in the interaction between Focus Condition and Time was significant: fixations to subject characters decreased at a steeper rate in the Object Focus condition than in the Neutral condition. The second contrast in the interaction between Focus Location and Time was also significant: fixations to subject characters decreased in the Neutral condition but increased in the Subject Focus condition (**Table 2B**).

Comparing the Subject Focus and Object Focus conditions against one another in a separate analysis showed that initially (400–600 ms), speakers fixated subject characters more often in the Subject Focus condition than in the Object Focus condition. As time progressed, speakers also directed their gaze to subject characters in the Subject Focus condition and away from the subject characters in the Object Focus condition (resulting in an interaction of Focus Location with Time).

**800–1600 ms.** In the Neutral condition, speakers briefly directed their gaze to the subject character and then shifted their gaze away from this character between 800 and 1600 ms. In contrast, fixations in the Object and Subject Focus conditions were largely consistent with Experiment 1. Specifically, at the first 200 ms of the time window (i.e., 800–1000 ms), speakers were more likely to fixate subject characters in the Neutral condition than in the Object Focus condition, but were more likely to fixate subject characters in the Subject Focus condition than Neutral condition. The first contrast in the interaction between Focus Location and Time was not significant; the second contrast in this interaction was significant (**Table 2C**). Interactions with the Time variable are difficult to interpret because of non-linearities in the distribution of fixations in the Neutral condition. Thus for a rough comparison of fixations in this time window across conditions, a complementary analysis was carried out using average empirical logits calculated across the entire time window (i.e., the overall likelihood of speakers fixating the subject character) as the dependent



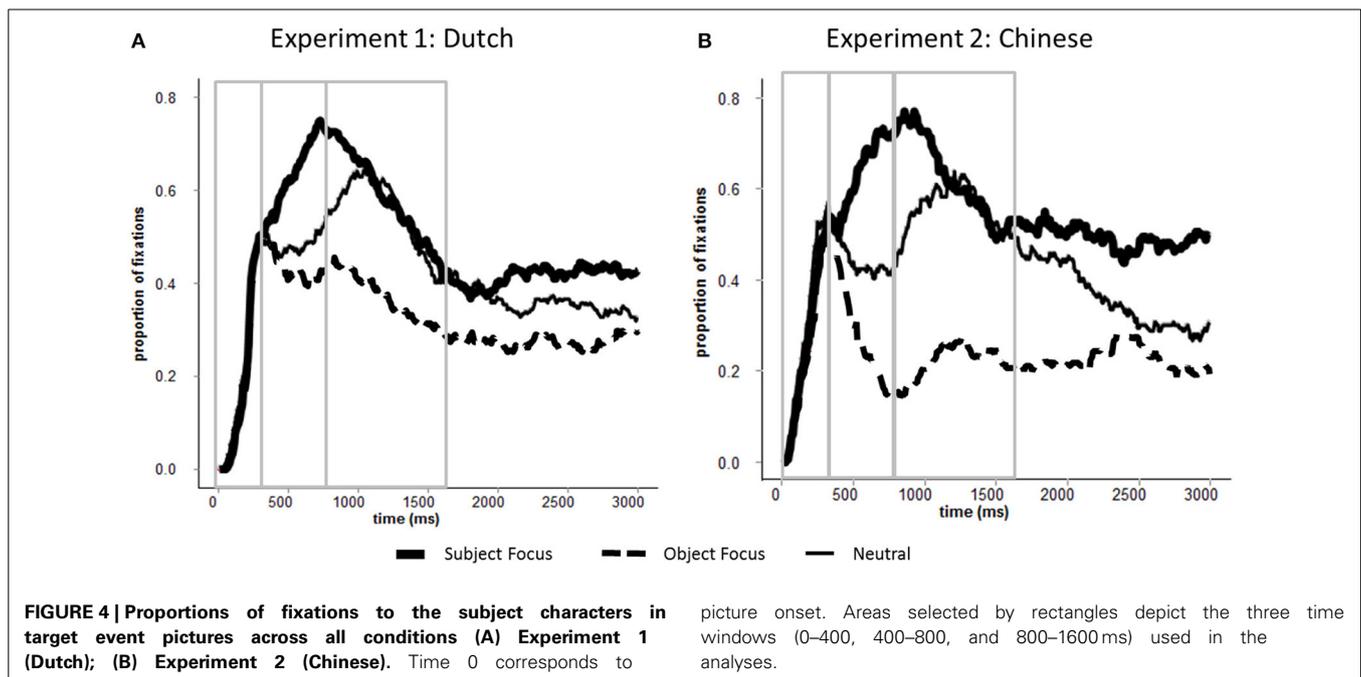
variable. This comparison showed the expected pattern: speakers were more likely to fixate subject characters in the Neutral condition than in the Object Focus condition ( $\beta_1 = -1.28$ ,  $SE = 0.15$ ,  $t = -8.46$ ;  $\beta_2 = -1.31$ ,  $SE = 0.12$ ,  $t = -11.08$ ) and were more likely to fixate subject characters in the Subject Focus condition than in the Neutral condition ( $\beta_1 = 0.75$ ,  $SE = 0.13$ ,  $t = 5.56$ ;  $\beta_2 = 0.75$ ,  $SE = 0.12$ ,  $t = 6.33$ ).

Finally, the Subject Focus and Object Focus conditions were compared against one another. As expected, the analysis showed that speakers were more likely to fixate the subject character in the Subject Focus condition than in the Object Focus condition at the first 200 ms of the time window (i.e., 800–1000 ms). The interaction with Time was also significant: fixations to subject characters decreased steeply in the Subject Focus condition but remained relatively stable in the Object Focus condition.

## DISCUSSION

Experiment 2 replicates the main findings of Experiment 1. First, speech onsets were longer in the Neutral condition than in the Object and Subject Focus conditions. The reduction in speech onset times was largest in the Object Focus condition<sup>2</sup>. Second,

<sup>2</sup>Note that speech onset latencies were somewhat different for Chinese and Dutch speakers. Specifically, Chinese speakers were overall faster than the Dutch participants. Chinese speakers were also faster in initiating speech in the Object Focus condition than the Subject Focus condition, while for Dutch speakers there was no reliable difference in speech onsets in these conditions. We compared speech onsets across the two groups in a complementary analysis with Focus Location (Neutral, Object Focus, Subject Focus) and Language (Chinese vs. Dutch) as fixed effects. The analysis showed a significant interaction between Focus Location and Language (Neutral vs. Object



and more importantly, Experiment 2 (Chinese) showed strong effects of the preceding discourse context on formulation. The pattern obtained in the Neutral condition again showed that participants looked at event characters in the order of mention, but in the Subject and Object Focus conditions, fixations to the two characters were strongly influenced by the preceding questions: after 400 ms, speakers preferentially and rapidly fixated the contextually *new* character.

Experiment 2 also shows the predicted cross-linguistic difference between Dutch and Chinese. Namely, shifts of gaze to the object character in the Object Focus condition began earlier than in Experiment 1: fixations to the object character increased immediately after 400 ms in Experiment 2 but only after 800 ms in Experiment 1 (see **Table 2B** for a comparison between experiments). To verify this finding, we ran additional analyses combining data from both experiments. The models included Time Bin, Focus Location (Neutral, Subject Focus, and Object Focus) and Language (Chinese and Dutch) as fixed effects. The analyses showed significant three-way interactions between these factors in the 400–800 ms time window (Neutral vs. Object Focus:  $\beta_1 = 3.08$ ,  $SE = 1.20$ ,  $t = 2.56$ ;  $\beta_2 = 2.15$ ,  $SE = 0.99$ ,  $t = 2.18$ ; Neutral vs. Subject Focus:  $\beta_1 = -2.39$ ,  $SE = 1.11$ ,  $t = -2.14$ ;  $\beta_2 = -1.78$ ,  $SE = 0.95$ ,  $t = -1.87$ ). As outlined earlier, this difference may be due to the fact that the surface word order in the Object Focus questions in Chinese provides speakers with a sentence preamble that they can repeat verbatim in their

response: availability of this material may have allowed Chinese speakers to direct their attention to the contextually new character earlier than Dutch speakers were able to do<sup>3</sup>. Consistent with this interpretation is also the large difference in speech onsets between the Object Focus and Subject Focus conditions in Experiment 2 (approximately 470 ms; this difference was only 5 ms in Experiment 1): Object Focus responses to questions in Chinese may have been easiest to prepare because speakers could repeat linguistic material from the question.

## GENERAL DISCUSSION

Two experiments compared the timecourse of formulation for sentences produced in response to three types of questions in Dutch and Chinese. The questions either provided no discourse context for the target event (Neutral condition) or specifically asked about one of the event characters (Object and Subject Focus conditions). The results showed that questions did not influence the distribution of attention to the two event characters immediately after picture onset (0–400 ms), i.e., during a period of message-level encoding. However, the highly linear pattern of formulation observed in the Neutral condition after 400 ms (e.g., Griffin and Bock, 2000; Konopka and Meyer, 2014) was different after Object Focus and Subject Focus questions: instead of fixating characters in the order of mention, speakers fixated primarily the *new* character, regardless of its position in the sentence.

Focus:  $\beta = 0.31$ ,  $SE = 0.07$ ,  $t = 4.48$ ; Neutral vs. Subject Focus:  $\beta = -0.25$ ,  $SE = 0.06$ ,  $t = -3.58$ ; Object Focus vs. Subject Focus:  $\beta = -0.29$ ,  $SE = 0.06$ ,  $t = -4.94$ ). This difference may be due to the fact that Dutch and Chinese participants initiated speaking at a different point relative to their progress with sentence preparation. However, we cannot conclude what this difference is due to in the current experiments, so it remains an interesting question for future cross-linguistic research.

<sup>3</sup>To verify whether this difference across experiments was due to differences in the syntax of *wh*-questions in Dutch and Chinese rather than to item differences, we also examined the timecourse of formulation in Experiment 1 (Dutch) for the subset of 43 pictures that were used in both experiments. The same pattern was observed for the smaller dataset as for the larger dataset reported in Experiment 1: Dutch speakers directed their gaze to the object character preferentially only approximately 800 ms after picture onset.

Differences in the likelihood of speakers fixating the subject and object characters in the Neutral condition and the two Focus conditions can be attributed to at least two factors. First, questions provided a discourse context that either did not draw attention to the subject and object characters (Neutral condition) or that did explicitly require preferential encoding of the contextually *new* character (Focus conditions). Second, explicit mention of one character in the question reduced the costs of retrieving its name when describing the target event and thus reduced the likelihood of speakers fixating this character (also see Konopka, 2014). Experiment 2 showed that reducing the costs of generating the target sentence itself in Chinese further reduced the likelihood of speakers fixating the old character.

The observed difference between Dutch and Chinese across experiments lends convincing evidence that sentence planning can be influenced by the linguistic context in which a target utterance is prepared and produced. Differences in the grammaticalized word orders in Chinese and Dutch facilitated production in Chinese as Chinese speakers could start by repeating verbatim the subject and verb of the preceding question without any further re-ordering of the syntactic constituents as is necessary for Dutch. The cross-linguistic difference therefore may be partly due to repetition priming and syntactic priming (e.g., Pickering and Branigan, 1998, 1999): given the compatible word order in the Object Focus question and the response in Chinese, priming is possible for Chinese speakers but not for Dutch speakers. To the extent that eye movements provide insight into the allocation of attention and resources to different encoding processes, large changes in the temporal coupling of gaze and speech suggest that context can strongly influence the incremental formulation of simple utterances. Specifically, the results of both experiments show strong effects of top-down guidance from the message level and contextual facilitation of linguistic encoding: on the basis of their encoding of event gist immediately after picture onset (0–400 ms) and their exposure to linguistic material in the question, speakers deployed their gaze only to the character they needed to encode to answer the question. Thus, eye movements in the Object and Subject Focus conditions show that shifts of gaze need not closely reflect the order of linguistic encoding operations. Rather, they are better indicators of *higher-level* communicative goals and recent linguistic experience: speakers direct their attention to whatever part of the display they need to process with priority to produce a contextually fitting response. Tight coordination of gaze and speech (e.g., Griffin and Bock, 2000) may therefore be more representative of formulation of sentences out of context, where all information in a to-be-described event is new and unfocused.

More generally, the results are compatible with theories of incrementality in sentence formulation that propose top-down guidance during the formulation process (Bock et al., 2004; Konopka and Meyer, 2014; see Gleitman et al., 2007, for an alternative, bottom-up account of sentence formulation). The key assumption of these theories is that sentence formulation begins with the formulation of a message-level representation that guides all subsequent encoding operations, as reflected in the ensuing pattern of eye movements to different parts of a to-be-described event. The results of the current experiments show that,

when message-level representations include information about discourse focus, the timecourse of sentence formulation changes immediately to reflect changes in speakers' communicative goals. The high degree of similarity in the timecourse of formulation across languages shows language-general adaptations in the incremental preparation of simple sentences.

## ACKNOWLEDGMENTS

We thank Margaret den Besten and Yifei Bi for help with data collection for the Dutch and Chinese experiments respectively. This research was supported by a VIDI Grant (NWO-061084338) and by ERC grant to Yiya Chen.

## REFERENCES

- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *J. Mem. Lang.* 59, 457–474. doi: 10.1016/j.jml.2007.09.002
- Bock, K., Irwin, D. E., and Davidson, D. (2004). “Putting first things first,” in *The Integration of Language, Vision, and Action: Eye Movements and the Visual World*, eds F. Ferreira and M. Henderson (New York, NY: Psychology Press), 249–278.
- Cheng, L. (2009). *Wh-in-situ* from the 1980s to now. *Lang. Linguist. Compass* 3, 767–791. doi: 10.1111/j.1749-818X.2009.00133.x
- Ganushchak, L. Y., Konopka, A. E., and Chen, Y. (2014). “Focus planning during sentence production: an eye-tracking study,” in *Poster Presented on International Seminar on Speech Production (ISSP)* (Cologne).
- Gleitman, L., January, D., Nappa, R., and Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *J. Mem. Lang.* 57, 544–569. doi: 10.1016/j.jml.2007.01.007
- Griffin, Z. M. (2004). The eyes are right when the mouth is wrong. *Psychol. Sci.* 15, 814–821. doi: 10.1111/j.0956-7976.2004.00761.x
- Griffin, Z. M., and Bock, J. K. (2000). What the eyes say about speaking. *Psychol. Sci.* 11, 274–279. doi: 10.1111/1467-9280.00255
- Gussenhoven, C. (2007). “Types of focus in English,” in *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*, eds C. M. Lee, M. Gordon, and D. Büring (Dordrecht: Springer), 83–100. doi: 10.1007/978-1-4020-4796-1\_5
- Kempen, G., and Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cogn. Sci.* 11, 201–258.
- Konopka, A. E. (2012). Planning ahead: how recent experience with structures and words changes the scope of linguistic planning. *J. Mem. Lang.* 66, 143–162. doi: 10.1016/j.jml.2011.08.003
- Konopka, A. E. (2013). “Discourse changes the timecourse of sentence formulation,” in *Poster presented at the 19th Architectures and Mechanisms for Language Processing Conference* (Marseille).
- Konopka, A. E. (2014). “Speaking in context: discourse influences formulation of simple sentences,” in *Poster Presented at the 27th CUNY Human Sentence Processing Conference* (Columbus, OH).
- Konopka, A. E., and Meyer, A. S. (2014). Priming sentence planning. *Cogn. Psychol.* 73, 1–40. doi: 10.1016/j.cogpsych.2014.04.001
- Kuchinsky, S. E., and Bock, K. (2010). “From seeing to saying: perceiving, planning, producing,” in *Paper Presented at the 23rd Meeting of the CUNY Human Sentence Processing Conference* (New York, NY).
- Meyer, A. S., and Lethaus, F. (2004). “The use of eye tracking in studies of sentence generation,” in *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*, eds J. M. Henderson and F. Ferreira (New York, NY: Psychology Press), 191–212.
- Meyer, A. S., van der Meulen, F. F., and Brooks, A. (2004). Eye movements during speech planning: talking about present and remembered objects. *Vis. Cogn.* 11, 553–576. doi: 10.1080/13506280344000248
- Pickering, M., and Branigan, H. (1998). The representation of verbs: evidence from syntactic priming in language production. *J. Mem. Lang.* 39, 633–651.
- Pickering, M., and Branigan, H. (1999). Syntactic priming in language production. *Trends Cogn. Sci.* 3, 136–141.

Van de Velde, M., Meyer, A. S., and Konopka, A. (2014). Message formulation and structural assembly: describing “easy” and “hard” events with preferred and dispreferred syntactic structures. *J. Mem. Lang.* 71, 124–144. doi: 10.1016/j.jml.2013.11.001

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 June 2014; accepted: 16 September 2014; published online: 02 October 2014.

*Citation:* Ganushchak LY, Konopka AE and Chen Y (2014) What the eyes say about planning of focused referents during sentence formulation: a cross-linguistic investigation. *Front. Psychol.* 5:1124. doi: 10.3389/fpsyg.2014.01124  
This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Ganushchak, Konopka and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Agent-patient similarity affects sentence structure in language production: evidence from subject omissions in Mandarin

Yaling Hsiao, Yannan Gao and Maryellen C. MacDonald\*

Department of Psychology, University of Wisconsin–Madison, Madison, WI, USA

## Edited by:

Peter Indefrey, University of Dusseldorf, Germany

## Reviewed by:

Antje Meyer, Max Planck Institute for Psycholinguistics, Netherlands  
Leah Roberts, University of York, UK

## \*Correspondence:

Maryellen C. MacDonald, Department of Psychology, University of Wisconsin–Madison, 1202 West Johnson Street, Madison, WI 53706, USA  
e-mail: mcmacdonald@wisc.edu

Interference effects from semantically similar items are well-known in studies of single word production, where the presence of semantically similar distractor words slows picture naming. This article examines the consequences of this interference in sentence production and tests the hypothesis that in situations of high similarity-based interference, producers are more likely to omit one of the interfering elements than when there is low semantic similarity and thus low interference. This work investigated language production in Mandarin, which allows subject noun phrases to be omitted in discourse contexts in which the subject entity has been previously mentioned in the discourse. We hypothesize that Mandarin speakers omit the subject more often when the subject and the object entities are conceptually similar. A corpus analysis of simple transitive sentences found higher rates of subject omission when both the subject and object were animate (potentially yielding similarity-based interference) than when the subject was animate and object was inanimate. A second study manipulated subject-object animacy in a picture description task and replicated this result: participants omitted the animate subject more often when the object was also animate than when it was inanimate. These results suggest that similarity-based interference affects sentence forms, particularly when the agent of the action is mentioned in the sentence. Alternatives and mechanisms for this effect are discussed.

**Keywords:** language production, sentence production, subject omission, grammatical encoding, interference

## INTRODUCTION

An important tool to understand the mapping from conceptual to lexical representations during language production is the picture-word interference paradigm, in which speakers name a picture and attempt to ignore a word printed on it. Picture naming in this situation is influenced by the relationship between the target picture and the distractor word. One classic result is that naming of a target picture (e.g., *cat*) is slower when the distractor word is of the same semantic category (e.g., *dog*) than when the word is semantically unrelated (e.g., *clock*; e.g., Rosinski et al., 1975; Glaser and Dünghelhoff, 1984, and many studies since). This result is interpreted to support the claim that lexical selection (settling on the word *cat* to name the picture) is a competitive process and is subject to interference from other activated words, in this case the highly semantically similar distractor word *dog*, making it harder to settle on the correct item for the utterance plan. This effect is reminiscent of behavior in the Stroop task (Stroop, 1935), where people are slower to name a color patch in the presence of a distractor color word (again, an element from the same semantic category). The effect may also be related to phonological interference among items in an utterance plan, in which partial phonological overlap among words leads to longer initiation latencies and higher levels of production errors. Although the picture-word paradigm yields a mix of interference and facilitation effects from phonological similarity, depending on timing and other factors (e.g., Schriefers et al., 1990; Meyer

and Schriefers, 1991), phonological overlap in phrase and sentence production yields longer latencies and more errors (Wilshire, 1998; Acheson and MacDonald, 2009; Janssen and Caramazza, 2009; Jaeger et al., 2012).

The sum of these results suggests that while planning an utterance, certain properties of words may interfere with one another in a variety of ways. There is still much to be learned about the nature of this interference, but this article addresses instead a *consequence* of interference in language production. In many of the studies noted above, the participants are constrained in the order in which they utter the words, or, in the case of picture-word interference tasks, have only one word to utter. In more unconstrained sentence production, however, producers often have a choice of word orders and sentence structures to convey an intended message. We investigate whether same-category interference, such as between *cat* and *dog*, also affects the sentence structures and word order that is developed during grammatical encoding. That is, we ask whether speakers' and writers' implicit choices of sentence structure are different in situations conveying a message with two same-category competitors (e.g., *cat*, *dog*) vs. situations in which the message does not require two same-category lexical items. Specifically, we speculate that the message is realized by the producer in a way that minimizes competition between the two semantically similar items. As a first step, our investigations compare messages with two animate sentence participants, so that the producer needs to convey that a human is acting on another

human, to those messages in which a human is acting on an inanimate object. For now, we will set aside the issue of whether any interference between the animate entities exists at a conceptual or lexical level or both (see Damian and Bowers, 2003, for discussion) and will simply refer to any interference of this sort as similarity-based interference.

### PRODUCTION CHOICES AND PRODUCTION DIFFICULTY

Almost any concept or message can be conveyed in a number of different ways—different sentence structures, word orders, and lexical items. MacDonald (2013a) argued that during the stages of utterance planning that precede articulation in language production, producers tend to settle on utterance forms that minimize the difficulty of utterance planning. She argued that these effort minimization biases were emergent from non-linguistic action and motor planning, where easier (more practiced, simpler, recently used, etc.) motor plans capture internal attentional resources over more complex plans, and they are therefore more likely to be implemented than the more complex alternatives. In incremental language production planning, in which elements of the utterance plan are developed and held in memory before the plan is executed, the presence of several semantically similar elements in the plan reduces the distinctiveness of these elements in memory (Acheson and MacDonald, 2009), thereby increasing the difficulty of developing and maintaining the ordered elements in the plan. One difficulty-minimization utterance planning bias that MacDonald identified was Reduce Interference, that producers tend to develop utterance plans that minimize similarity-based interference. Here we consider how similarity-based interference could affect the accessibility (readiness for articulation, Bock and Warren, 1985) of the interfering elements, and the consequences of variation in accessibility for choices of utterance form during language production.

A number of studies have shown clear effects of lexical accessibility on sentence structure and word order in production, so that more conceptually salient (accessible) nouns tend to be placed earlier in the utterance or at a syntactically more prominent position (e.g., Bock and Warren, 1985), such as the grammatical subject. For example, animate nouns, which are thought to be more salient and recalled more rapidly from long-term memory, are more likely than inanimate nouns to be uttered early and assume the surface subject position even when they are not agents of the event, resulting in the production of passive sentences like “*The boy was hit by the ball*,” rather than the active form “*The ball hit the boy*” (Bock et al., 1992). Structural relations of nouns in the sentence, in this case in English, are affected by the conceptual roles associated with the animacy of the referents of the nouns. Similarly, data from languages that allow flexible word order, such as Japanese, show that sentences in the object-subject-verb (OSV) word order with animate subjects tend to be recalled as subject-verb-object (SVO), associating animate nouns to the more prominent subject grammatical role, and to an earlier sentence position (Tanaka et al., 2005). These and similar results concerning the effects of accessibility on sentence form are relevant to effects of similarity-based interference, because if similarity-based interference can affect the accessibility of words to be placed in the utterance plan, then it is

plausible that these variations in accessibility could affect sentence form.

A second piece of evidence that makes it plausible that similarity-based interference could affect word order is that similarity-based interference affects utterance planning difficulty. Semantic or phonological similarity increases the rate of serial ordering errors in production: Dell and Reich (1981) studied the rate of word exchange errors, such as when the intended message *I wrote a letter to my mother*, is realized as *I wrote a mother to my letter*. They found that the exchanged words (e.g., *letter*, *mother*) have more phonological similarity than would be expected by chance, suggesting that the phonological similarity increases the chances of ordering errors during language production planning. Semantically related items in the utterance plan also yield longer initiation latencies, longer utterance durations and overall higher error rates, compared to conditions without semantic similarity. Acheson and MacDonald (2009) relate these and other similarity effects to contextual distinctiveness in serial ordering, in which nearby items in a memory representation (including an utterance plan during language production planning) tend to have more similar contextual representations than those farther apart, regardless of whether the contextual representation is external (e.g., list position in a recall task) or internal (e.g., distributional properties of syllable position in individual words: vowels and consonants are less likely to substitute each other in speech errors) to the items. In the short-term memory literature, similar constraints also apply: with short interval between presentation and recall, items interfere with one another in memory when sharing similarity in sound, meaning, location, or other dimensions (Anderson, 1983). Utterance planning has similar short-term memory demands (Acheson and MacDonald, 2009; MacDonald, 2013a), and so these same distinctiveness effects would be expected to influence serial ordering in language production, such that two similar (less distinct) items would be more likely to be exchanged or subject to error than two more distinct items.

Together these findings suggest that (a) similarity between entities in an utterance plan affects the difficulty of the planning of that utterance and the likelihood of errors, (b) similarity affects the accessibility of entities in the utterance plan, and (c) the accessibility of items influences sentence form and word order. Gennari et al. (2012) investigated the effect of similarity on sentence form using picture description tasks in three languages: English, Spanish, and Serbian. They studied active vs. passive relative clause production, as in *The baby that the woman is holding* vs. *The baby that's being held by the woman*. Unlike simple active and passive sentences, which have different noun orders, relative clauses in these three languages fix the position of the modified noun (i.e., head noun, *baby* in this example) in the clause-initial position and therefore allow better comparison between the active and passive forms. Thus in the active form *the baby that the woman is holding*, the head noun *baby* and embedded subject *woman* are near each other and are both in prominent grammatical roles (*the baby*: main clause subject, *the woman*: relative clause subject). In the passive, however, the agent of the holding action, *woman*, is produced in the by-phrase, which is optional. Gennari et al. (2012) found that in picture descriptions in all three languages, the rate of passives was higher when both

entities in the picture were animate (e.g., a woman holding a baby) than when an animate entity was acting on an inanimate entity (e.g., a woman holding a vase). Moreover, within the set of passive utterances, the rate of agentless passives, omitting the by-phrase, was higher for the animate patients (*The baby that's being held*) than for the inanimate ones (*The vase that's being held*). These results, which were replicated in English by Montag and MacDonald (2014), suggest that interference between the conceptually similar items *woman* and *baby* reduces the accessibility of *woman* and thereby promotes the passive form, where the agent *woman* is either demoted to a more minor part of the sentence (the by-phrase) or omitted entirely in the agentless passive.

To test that this effect stemmed from agent-patient similarity and was not simply an effect of animacy of a noun, Gennari et al. (2012) collected similarity ratings for the entities in their pictures. They found that in both Spanish and English (the two languages in which there were enough passives to conduct the analyses), the more similar the two entities to be described were, the more agentless passives were produced. A follow-up experiment using new pictures with only animate event participants confirmed this pattern: in both Spanish and English, participants produced more agentless passives (e.g., *The builder who was slapped*) when the agent was semantically similar (a miner) than when the agent was dissimilar (an astronaut). These results show how structure choice can emerge not simply from properties of a single noun, such as the head noun *builder* but also via the interaction between two event participants. When these entities are highly similar and thus create similarity-based interference, producers are more likely to omit mention of one of them from the utterance.

These studies shed light on potential underlying causes stemming from production constraints for the utterance forms that constitute distributional regularities in a person's language experience (MacDonald, 2013a). Most of the evidence for the biases mentioned above, however, comes from complex constructions, such as relative clauses. One concern is that a multitude of complex interactions among production constraints and task demands can be at work to create the patterns that Gennari et al. (2012) observed in relative clause production. Therefore, instead of using complex structures like relative clauses, we chose to examine simple sentences in Mandarin Chinese, a language that allows noun omission in certain discourse contexts. Typically the omitted element is thought to be a pronoun, because the discourse environment in which omission is possible is also the environment (prior mention) in which it is felicitous to use a pronoun. The omission phenomenon is variously described as pro-drop (i.e., that a pronoun is dropped), pronoun elision (i.e., omission), and null subject and null object, referring to an omission of the grammatical subject or object, respectively. Some languages, such as Spanish, permit omission of only the subject, while others, such as Mandarin and Japanese, permit omission of subjects, objects, and some other grammatical positions. Although omission phenomena have received a number of linguistic treatments, syntacticians commonly view sentences with omitted elements to have a different syntactic structure than the sentences in which the pronoun is present, although analyses may differ by languages (e.g., Biberauer et al., 2010; Camacho, 2013)

As our focus here will be on omitted subjects, we will refer to *omitted* or *null subjects*, even though Mandarin also allows omission of other grammatical positions. For example, in a scenario where two Mandarin speakers have been talking about a movie, one person can ask the other the question "Did you watch the movie?" in four different formats: (a) "You watched the movie?" in which both you and movie are overtly mentioned, (b) "You watched \_\_\_?" in which movie is omitted from the utterance, (a null object construction), (c) "\_\_\_ watched the movie?" a null subject construction, or (d) "\_\_\_ watched \_\_\_?" in which both the subject and object are omitted from the utterance. All four of these alternatives are grammatical in Mandarin, and the clarity of the message is not compromised as long as the context provides clear clue to what the omitted elements are, much as the message is clear in the English, "Want to go to a movie?," in which the pronoun *you* is omitted. Unlike many other pro-drop languages, Mandarin lacks a rich morpho-syntactic system that redundantly encodes the pronominal information with verbal inflections and other agreement systems (no number, gender, and tense agreement, no case marking). Therefore, Mandarin pro-drop may lend us a clearer lens in uncovering the production mechanisms behind null subjects and other omissions, perhaps more purely based on the lexical retrieval difficulty among the competing nouns.

In two studies reported below, we investigated the role of similarity-based interference on producers' use of null subject constructions in Mandarin. If the Gennari et al. (2012) relative clause production phenomena (i.e., agent omission in relative clause production in English and Spanish) generalize to a very different language and sentence structure, then producers should produce more null subject structures when the subject and object are similar than when they are dissimilar. We investigated this prediction in an analysis of a written corpus in Study 1 and in a spoken picture description task in Study 2, using animacy of the subject and object nouns as a proxy.

## STUDY 1 – CORPUS ANALYSIS

The corpus analysis presented here is an extension of one originally conducted by Hsiao and MacDonald (2013). Their original analysis focused on main and relative clause usage in Mandarin, with the goal of creating a training set for a computational model that closely matched Mandarin speakers' experience relevant to Mandarin relative clause comprehension. Among other sentence types, Hsiao and MacDonald extracted all simple (one clause) sentences with overt or null subject noun phrases from the parsed Chinese Treebank 7.0 (Xue et al., 2010). There were 4035 simple transitive sentences with overt direct object phrases, of which 2445 (61%) contained overt subjects and 1590 (39%) contained null subjects. These 4035 sentences formed the basis for our analyses here.

Hsiao and MacDonald (2013) hand-coded the animacy of all overt noun phrases in these sentences, but they did not code the animacy of the referent of the (omitted) subject nouns in the null subject sentences, that is, the animacy of the entity being discussed in the broader discourse context. In order to investigate whether null subject sentences are more frequent when the subject and object are conceptually similar than when they are less similar, we

used the surrounding sentence context to code the animacy of the intended referents for the omitted subjects.

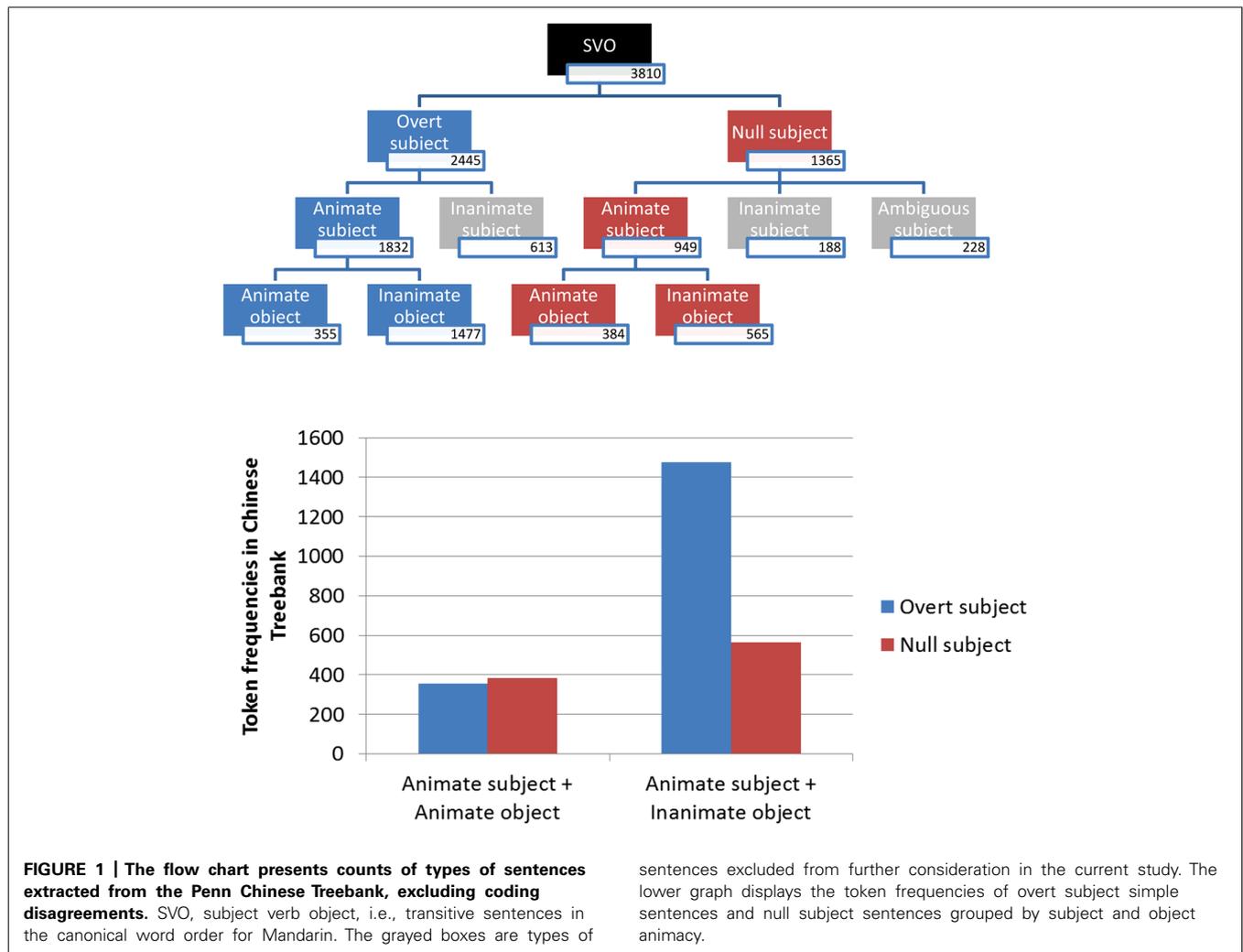
The 1590 null subject sentences were coded for the animacy of their omitted subjects. Two native Mandarin speakers who were blind to the hypotheses coded animacy of the null subject via the material in the verb phrase. For example, when the sentence read “\_\_\_ gave a thank-you speech,” the verb denotes an action that could only be completed by a human. Therefore, the omitted subject NP was coded as animate. For sentences like “\_\_\_ exceeds the percentage last year,” the omitted subject refers to some numerical value, which was coded as inanimate. Sentences for which the verb phrase did not clearly convey subject animacy, such as “\_\_\_ created uproar,” were coded as ambiguous. The overall inter-rater reliability was 85%. All items with a disagreement among coders were excluded from further analyses.

**RESULTS**

The coding results are summarized in the flow chart in **Figure 1**. Among a total of 1365 null subject sentences after excluding coder disagreements, 949 sentences were coded by both raters as having

animate referents for the null subjects, 188 were coded as having inanimate subjects, and 228 were agreed to be ambiguous, meaning that subject animacy could not be determined from the sentence context. Since animacy could not be established for the ambiguous items, they were excluded. We also excluded sentences with inanimate subjects, because there were too few observations in each cell when these items were partitioned into groups with animate vs. inanimate direct objects.

Among the 949 sentences with animate subject referents, 384 items had animate objects, and 565 were with inanimate objects. The bar graph in **Figure 1** compares these values to the patterns of overt subject usage that Hsiao and MacDonald (2013) found. Overt animate nouns, on the other hand, contained 355 sentences with animate objects, and 1477 with inanimate objects. These data show that there was a strong association between subject omission and the animacy of the direct object: when both the subject and the object were animate, the frequency of null subject sentences was higher than that of overt subject sentences; whereas when the subject was animate and the object was inanimate, the majority of them were overt subject sentences, [ $\chi^2(1, N = 2781) = 142, p < 0.05$ ].



**FIGURE 1 | The flow chart presents counts of types of sentences extracted from the Penn Chinese Treebank, excluding coding disagreements.** SVO, subject verb object, i.e., transitive sentences in the canonical word order for Mandarin. The grayed boxes are types of

sentences excluded from further consideration in the current study. The lower graph displays the token frequencies of overt subject simple sentences and null subject sentences grouped by subject and object animacy.

These results are consistent with the hypothesis that in conditions of similarity-based interference, speakers produce more null subject sentences. We also considered a second possibility, that similarity-based interference could affect the use of overt pronouns vs. full noun phrases, as some previous research has suggested that pronoun use varies as a function of whether the animacy of subjects and objects matches or not. Fukumura and van Gompel (2011) and Fukumura et al. (2011) found that in sentence completion tasks where the subject and object NPs were of the same animacy, participants referred to either one of them (depending on the manipulation: half of the time the subject NP and the other half the object NP) with pronouns less frequently than when both NPs were of different animacy. The finding suggests that similarity in meaning between the two nouns makes the referent's representation less accessible. However, in our study, the pronoun/full noun phrase contrast could not be investigated, because subject pronouns were too rare—the vast majority of sentences contained overt full noun phrases or null subjects, and subject pronouns comprised only about 3% of the extracted sentences. The low percentage of pronoun use may be attributed to the formal nature of written texts in Mandarin Chinese. Mandarin overt pronoun use varies with the social distance between the speaker and the interlocutor, and even the third party being referred to. The farther the social distance between the producer and the referent, the less likely a pronoun will be used (rather, role names are used for higher-ups, e.g., addressing your college professor as “Professor Wang” instead of “you”). This explains the rarity of pronoun use in the current corpus, which is composed of articles and transcripts from newspapers or news broadcasting normally written with formal language (Wang, 1987; Wang, 2007).

## DISCUSSION

The corpus results suggest that when the agent and patient are of similar and salient conceptual representations (animate entities), people producing a simple transitive sentence are more likely to omit the subject (agent). This pattern, as seen in unconstrained natural speech transcripts and texts outside of the laboratory, is a valuable piece of evidence for the relationship between similarity-based interference and subject omission in production. However, as with any unconstrained language sample, we cannot be sure whether other factors instead of or in addition to agent-patient similarity affected subject omission. For example, the sentences with animate direct objects may have tended to occur in different kinds of discourse contexts than those with inanimate objects. The use of null subjects is dependent on the referent being previously established (given) in the discourse, and it is possible that higher rates of null subjects in the animate direct object sentences may have been due to those sentences appearing in discourses in which the agent of the action had been more firmly established in the discourse compared to the sentences with inanimate direct objects. To address this concern, in the next experiment, we conducted a picture description task that controlled the discourse contexts to be equally plausible and appropriate for subject omission in all conditions and manipulated the animacy of patients/themes in the event while keeping the agents animate. If similarity-based interference affects the rate of subject omission in production, then we should find a

similar pattern to the one in the corpus analysis: more subject omission when both the agents and the patients of the action are animate.

## STUDY 2 – SENTENCE PRODUCTION TASK

### PARTICIPANTS

A total of 26 native Mandarin speakers were recruited from an Introductory Psychology class at the University of Wisconsin-Madison. All participants reported that they had been born or educated in China or Taiwan and spoke Mandarin Chinese as their dominant language. The majority of them were freshmen and sophomores who had spent less than 2 years in the United States. Participants received extra credit in the course for participation in the study.

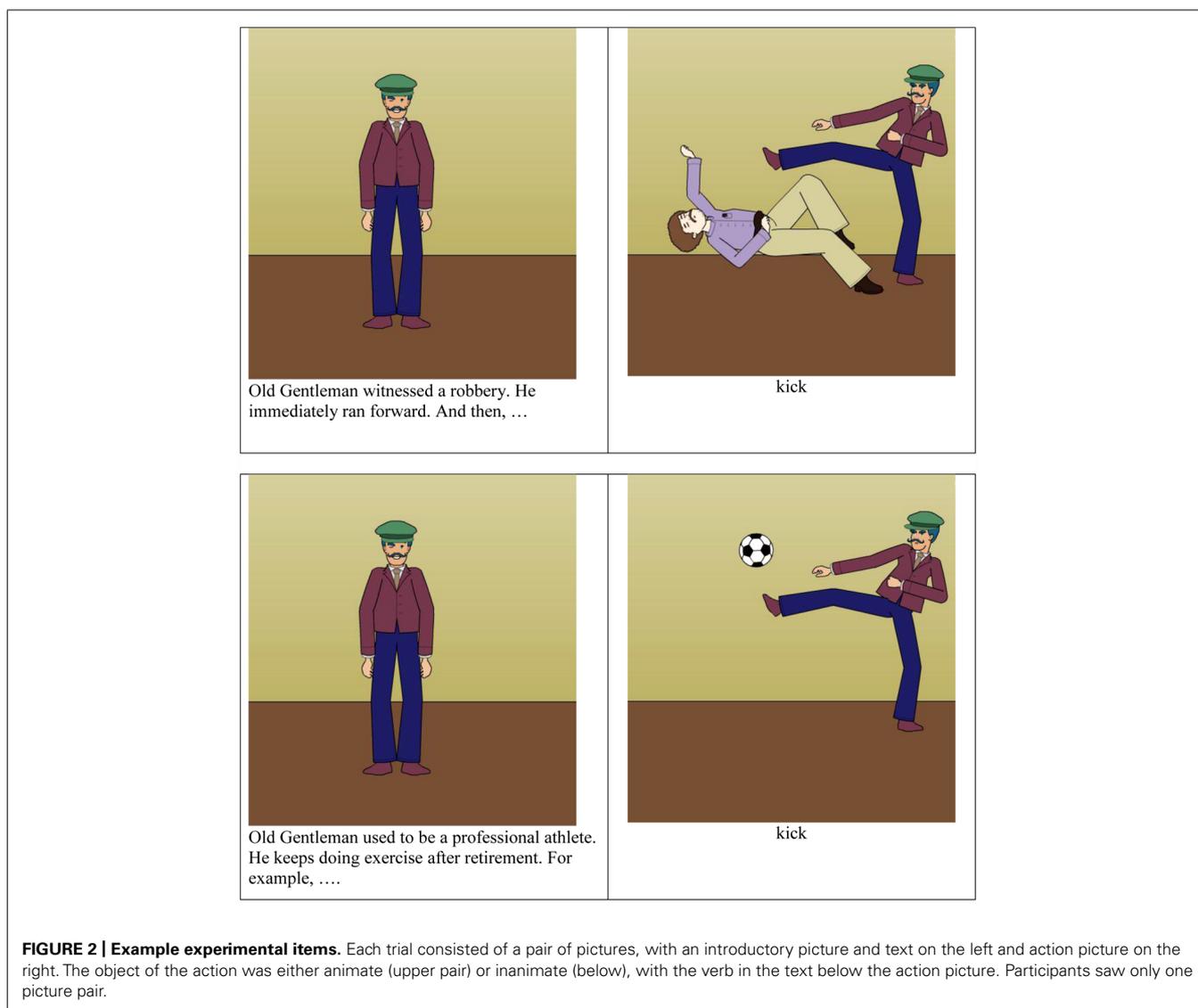
### MATERIALS

All pictures for the experiment were created using the online comic design website Pixton<sup>1</sup>. Twenty experimental picture triples were created. One member of the triple was an *introductory* picture, depicting a single standing human character with neutral facial expression. This picture introduced the agent of a subsequent action, creating a discourse context in which it would be felicitous to use either an overt pronoun or a null subject construction when referring to this character. The other two pictures were *action* pictures and showed the character acting on another entity. In one version, the entity being acted on was animate (another human), and in the other, it was inanimate.

The introductory picture was paired with one of the action pictures in each trial, with the introductory picture arranged to the left of the action picture. An example is shown in **Figure 2**. The two pictures were presented together in order to create a sense of continuous story flow and thus a better discourse environment for subject omission. Two or three sentences were written under the introductory picture, providing background information about the character (e.g., occupation, disposition, habits) and establishing the character as given in the discourse. The character's label was used as the grammatical subject of the first sentence (e.g., *Old Gentleman* for the examples in **Figure 2**) and a pronoun referring to the pictured character as the grammatical subject was used for subsequent sentences (e.g., *he*). In addition to introducing the character into the discourse, these introductory sentences also served to establish the plausibility of the event conveyed in the action picture. Because it was difficult to provide a single plausible discourse context for both an event involving an animate patient and one involving an inanimate object, the contexts differed for the two conditions where necessary to create a plausible sequence of events.

The action picture on the right appeared with a single word referring to an action, in order to encourage all participants to be consistent in their verb use when describing the action picture. For test trials, the action picture always depicted a transitive action performed by the human character introduced in the picture on the left. The human character exerted the action on an animate patient or an inanimate theme in the picture on the right. The two versions of the action pictures

<sup>1</sup>www.pixton.com



**FIGURE 2 | Example experimental items.** Each trial consisted of a pair of pictures, with an introductory picture and text on the left and action picture on the right. The object of the action was either animate (upper pair) or inanimate (below), with the verb in the text below the action picture. Participants saw only one picture pair.

were controlled to have the same background color and the same human character, which was made to have roughly the same action and position in the two action pictures. Thus the only difference between animate and inanimate action picture was the animacy of the object of the action. The verb used to describe the action was selected to be appropriate for both an inanimate and animate object. Two lists were created to counterbalance the assignment of animate or inanimate action pictures across participants, each of whom saw 10 animate and 10 inanimate objects in the experimental action pictures, and no participant saw both versions of the action pictures for a given item.

Thirty filler picture pairs were created. These were similar in form to the experimental items except that there was only one action picture matched with an introductory picture, and some of the action pictures depicted intransitive actions with no direct object. On some filler trials, the word under the action picture was a noun rather than a verb.

### PLAUSIBILITY NORMING

In order to ensure that the pairs of introductory and action pictures were equally plausible in the animate and inanimate object conditions, we conducted a rating study with a separate group of 48 native Mandarin-speaking participants, all of whom were from mainland China. The survey took 7–10 min to complete. Participants volunteered their time and were not compensated for participation.

The rating task had 20 test trials and 20 filler trials, each with introductory and action pictures with associated text, except that the single word underneath the action pictures that appeared in the main experiment was not presented in the rating study. The filler trials were 20 of the filler picture pairs from the main experiment, except that a portion of them had their text modified to be less plausible. This change was designed to create some variability in the range of events and to provide a manipulation check to determine whether raters were reading the text carefully and assessing plausibility in each trial.

Two lists were created, counterbalancing assignment of inanimate or animate action picture across subjects, so that each participant saw 10 animate and 10 inanimate objects in the experimental action pictures, and no participant saw both versions of the action pictures for a given item. The participants were asked to indicate the plausibility of the event in the action picture given the context sentences for the introductory picture, using an on-screen sliding scale of scores 1–7, with 1 referring to extremely implausible and 7 to very plausible. The study was hosted through the online surveying service Qualtrics<sup>2</sup>. On each trial, participants saw a picture, read the associated text, and used the mouse to adjust a slider on screen to correspond to their plausibility rating. Participants proceeded through the items at their own pace.

Statistical analyses of the plausibility data were conducted with mixed effects models with maximum random effects of participants and items, as suggested in Barr et al. (2013). The plausibility of the filler trials were rated significantly lower than experimental trials, with an average of 3.37 for fillers and 5.21 for experimental trials ( $\beta = 1.84$ ,  $SE = 0.26$ ,  $t = 7.20$ ,  $p < 0.001$ ). The fact that the average ratings of fillers and experimental items fell on the opposite sides of the neutral rating of 4 confirmed the success of our design to involve overall plausible events for the experimental items and for a portion of the fillers to be less plausible. These results also suggest that participants were reading carefully when rating the picture pairs. We further analyzed the ratings within experimental trials and found no difference between the ratings for the animate condition and those of the inanimate condition, with the former having an average of 5.03 and the latter 5.39 ( $\beta = 0.36$ ,  $SE = 0.22$ ,  $t = 1.6$ ,  $p = 0.11$ ). Thus even though inanimate direct objects are more common in the world (e.g., as in the corpus analysis in Study 1, in which inanimate objects were more common than animate ones at a rate of about 2:1), the null result here suggests that the discourse contexts we designed made the inanimate and animate conditions similarly plausible.

## PROCEDURE

E-prime 2.0.10 was used to create experimental scripts for the main production experiment. Participants were assigned to one of the two lists, each containing 20 test trials and 30 filler trials. These trials were interleaved so that no more than three test trials appeared in a row.

In each trial, participants were asked to read the context sentences under the introductory picture aloud and then continue with a description of the character's action depicted in the action picture, using any sentences regardless of structure and length, as long as the response contained the verb shown below the action picture. Participants were encouraged to describe the action picture soon after finishing reading the context sentences aloud, without pausing to consider elaborate continuations. Before the experiment started, participants practiced with two sample trials. Participants' responses were recorded digitally through a microphone.

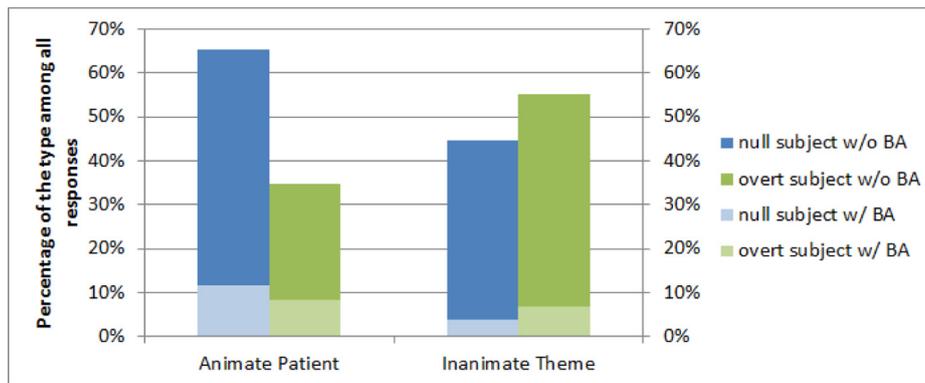
## RESULTS AND DISCUSSION

Participants' responses were transcribed and coded by a native Mandarin speaker. Utterances in which the agent was not the grammatical subject of the main clause (e.g., appearing in a conjunction clause or a passive sentence) were excluded from analyses. A total of 11% of responses were excluded. Responses were coded as null subject utterances when the grammatical subject position was empty and coded as overt subject utterances when a word occupied the subject position. All of the overt responses were pronouns; there were no responses that repeated the full NPs (i.e., character descriptions such as *Old Gentleman*) or used new descriptions such as *the old man*. The lack of repeated NP or full NP descriptions suggests that the introductory picture did establish the agent as given, allowing an overt pronoun or null subject continuation. The common use of overt pronouns in the spoken descriptions, in contrast to their rarity in the corpus, may have stemmed from the same social-discourse factor mentioned before: the speech modality here, and possibly the topics mentioned in the context sentences, are less formal than in the primarily written texts extracted from the corpus in Study 1.

Rates of subject omission in animate and inanimate object conditions are shown in **Figure 3**. Statistical analyses of participants' utterances employed mixed effects models with maximum random effects of participants and items (Barr et al., 2013). Comparing the rates of subject omission between the two conditions, we found that when the human character acted upon an animate patient, speakers omitted the subject NP 65% of the time, which was a reliably higher omission rate than when the human agent acted upon an inanimate object, with 44% omissions, ( $\beta = -0.21$ ,  $SE = 0.10$ ,  $t = -2.12$ ,  $p = 0.04$ ). The animacy effect remained significant after adding the plausibility ratings from the norming study as an additional factor to the model, and plausibility itself did not account for significant amount of variance in the responses ( $t = 0.01$ ). Given these animacy effects, a logical next step would be to identify whether finer-grained level of similarity beyond animacy should also show an effect on subject omission, as in the all-animate condition of Gennari et al. (2012). For example, pictures with two more similar human characters (same gender, similar age, occupation) could yield more null subjects than for pictures with more dissimilar human characters. We leave this to future research.

Some of the participants' responses employed the Mandarin disposal construction, also called BA construction (see **Figure 3**), which is a common form in describing Mandarin transitive events and typically expresses how an entity is handled, manipulated or dealt with (Li and Thompson, 1981). This construction was not included in the corpus analysis, which focused on simple sentences with SVO word order. In the production study, 85% of utterances were in this SVO word order and 15% were in the disposal construction, which has an SOV word order, with a light verb, such as *ba* or *jiang*, inserted between the subject and the object, as in *he ba robber kicked*, in the overt subject variant, or *ba robber kicked* in the null subject variant. The disposal construction is interesting from a production standpoint because it affords the producer an alternate word order, but the factors that promote use of this construction are beyond the scope of

<sup>2</sup>www.qualtrics.com



**FIGURE 3 | The rate of overt subject (green bars) vs. null subject (blue bars) responses in the picture description task.** The blue and green bars in each animacy condition add up to 100%. The lighter part of each bar indicates productions using the disposal constructions, and the darker portion shows productions with the standard SVO word order.

the current paper. Accordingly, our analyses focused simply on whether use of the disposal construction interacted with subject omission in some way. As **Figure 3** shows, there was a numerically higher percentage of disposal construction sentences in the animate condition (20%) than in the inanimate condition (11%), but this difference did not reach statistical significance ( $p = 0.1$ ) after including maximal random effects of subject and item. To test whether the use of the disposal construction might be related to subject omission, we added the percentage of null subjects as a factor in the model predicting disposal construction use, but the result was again not reliable ( $p = 0.9$ ). These results suggest that while the factors that promote production of the disposal vs. simple transitive construction are interesting and merit further study, the rate of subject omission does not appear to be tied to use of the disposal construction in this study.

## GENERAL DISCUSSION

The current study explored the effect of similarity-based interference in sentence production, using the presence of two human sentence participants as a condition of high similarity and a low similarity condition in which an animate entity acted on an inanimate one. The findings from both corpus analyses and a picture description experiment suggest that when Mandarin speakers are faced with developing an utterance plan containing two conceptually similar entities that may interfere with one another, they are more likely to omit one of the interfering elements than in the low similarity conditions.

The Mandarin null subject results here are similar to Gennari et al.'s (2012) results with relative clauses in English and Spanish, with higher rates of agentless passives (omission of the by-phrase) with a similar patient than a dissimilar one. Putting these results together with the ones in the current studies, there are consistent effects of agent-patient similarity on agent omissions across three quite different languages—English, Spanish, and Mandarin, across two sentence types—simple transitive sentences and relative clauses, and across paradigms—corpus analyses and picture descriptions. Together these results point to effects

of similarity-based interference on utterance form, specifically in choice of sentence structures that allow omission of the agent of the action—the null subject structures in the current studies, the agentless passives in English and Spanish in Gennari et al. (2012) and in Spanish a third agentless “impersonal” construction that Gennari et al. (2012) found is also more common under conditions of similarity-based interference. Thus over several different languages and structures, the unifying theme seems to be increased agent omission when the agent and patient of an action are similar compared to when they are less similar. In the next sections, we consider the evidence and opportunities for future research investigating the possible mechanisms underlying this agent omission effect, its relationship to other phenomena in production, and implications for theories of language production.

## INTERFERENCE, ACCESSIBILITY, AND INCREMENTALITY

There are several potential mechanisms that could link the similarity-based interferences effects in picture-word interference studies and the agent omissions that we've observed in sentence production. One possibility is that agent omission is an implicit strategy in language production: faced with interfering elements during utterance planning, speakers strategically choose an utterance form that reduces interference, i.e., choosing a form in which one of the interfering elements is placed some distance (in words) from the other, where the interfering elements are placed in very different syntactic positions (such as grammatical subject and adjunct, as in passives such as *The boy who was pushed by the girl*), or where one element is omitted altogether. On this view, structure choice is a direct (though unconscious) strategy to limit the interference and maintain fluency during production. An alternative view is that the utterance form is simply a consequence of the accessibility of the elements. On this more emergent view of omission, interference between similar elements leads to at least one of these elements being relatively inaccessible during utterance planning, with consequences for utterance form, as in other studies of accessibility in language production. Those studies often aim to increase an element's

accessibility, via priming, question-focusing, repetition, or other manipulations, with the consequence that speakers are able to retrieve highly accessible elements early and thus utter them early in an utterance (Bock and Warren, 1985; Bock, 1986). Interference has the opposite effect, decreasing accessibility, so that these low-accessible elements are delayed or omitted in the utterance. Thus both approaches link interference, accessibility of elements, and utterance form, but they differ in the extent to which they view this sentence-level planning phenomenon as strategic vs. emergent from the accessibility of elements of the utterance plan.

We do not believe that the experiments presented here or elsewhere distinguish these alternatives, and indeed it is not clear that the alternatives are completely incompatible. At issue is really the extent to which sentence planning is or can be under strategic control, which would accommodate strategic use of utterance forms to reduce interference between elements. Sentence form clearly can be under some deliberate strategic control on some occasions, and poets and other writers do consciously choose some sentence forms in some circumstances. It is less clear whether sentence form is always under a degree of strategic control, or whether it is more purely emergent from accessibility considerations at other times. The debate here seems similar to the question of the degree to which incrementality (planning ahead) during language production is under strategic control. Previous research does point to some amount of strategic control in the degree of advance planning (Ferreira and Swets, 2002; Wagner et al., 2010).

The analogy to incrementality here is interesting because the current data also bear on the question of the degree of advance planning during sentence production. By definition, similarity-based interference implies activation of both interfering entities, and therefore it suggests that there is sufficient advance planning to allow both entities to affect the development of the utterance plan. As such, the interference effects here argue against a “radical incrementality” perspective in which the first element (typically the subject) is planned and the sentence structure is adjusted thereafter to fit this encoding (e.g., Kempen and Hoenkamp, 1987; Levelt, 1989; de Smedt, 1996). Indeed, the Mandarin results are striking in this regard because material to be produced downstream (material in the verb phrase) affects whether the first position (the subject) will be uttered or not. Thus the current results are more consistent with a view in which an incremental production system is under some degree of strategic control of the speaker, and in which more advance planning may take place before production begins (Ferreira and Swets, 2002; Allum and Wheeldon, 2007; Wagner et al., 2010). This work is also consistent with results of Christianson and Ferreira (2005), who studied Odawa, a free word order language, using a picture description task that manipulated the agent and patient animacy and the focusing question. When the questions focused on an animate patient (e.g., “What is happening to the girl?” for a picture depicting a girl being pinched by a boy), participants’ answers tended to be passives even though the active object-first structures are appropriate answers, such as object-verb-subject or OSV. This result suggests that speakers would choose an overall less frequent sentence structure (i.e., passives) even though the language allows the dominant active voice

structure to appear with many word orders. Their results suggest that structure does not simply emerge from putting the most active element in sentence-initial position.

As researchers pursue these agent omission phenomena and the mechanisms that underlie them, it will be important to connect this work to another literature, the one addressing choice of referential form. That is, here we have been considering choice of sentence form, such as whether producers converge on an active or passive sentence, a full or agentless passive, or an overt or null subject, and most syntactic analyses consider these alternatives different syntactic constructions. However, the choice of a null vs. overt mention of an agent is also a choice of referential form—how producers choose to refer to some entity in the message. Typically studies in that literature investigate the conditions under which producers use (overt) pronouns vs. full noun phrases such as *the boy* or *Mary* (e.g., Arnold, 2010; Fukumura and van Gompel, 2011; Fukumura et al., 2011), but clearly speakers also choose omission to “refer” to entities for some languages, under certain discourse conditions and levels of interference. Indeed, some pronominal reference work describes cost functions for different referential forms (Almor and Nair, 2007). This point raises a related question: if the similar interfering elements (such as *cat-dog* or *old gentleman-robber*) are part of the producer’s message and thus a part of utterance planning, why is it that specifically overt mention is difficult? The answer, or perhaps a re-description, is that overt articulation appears to be especially sensitive to similarity-based interference. That is, perceiving or thinking about related elements (such as a cat and dog or an old man and robber) may not be more difficult than perceiving or thinking about less related ones (and may even be easier, given associative priming between related elements that is commonly found in perception, e.g., Neely, 1991 for a review.), but planning an utterance—retrieving, ordering, and/or phonologically encoding the lexical items, is especially sensitive to similarity, apparently even when the phonological realization of the referent is a pronoun. It may be conceptual representations that are phonologically realized in the utterance must be kept more active, guiding phonological encoding, than when there is no overt mention in the utterance, and that this longer or stronger activation is a source of higher difficulty. These speculations clearly merit additional research, and they suggest some continued interaction between levels of phonological encoding, where the phonological form is planned, and grammatical encoding, where the sentence form is developed (Janssen and Caramazza, 2009; Jaeger et al., 2012).

Another potentially related literature concerns the use/omission of other optional elements in an utterance, including the richness of inflections attached to a referential form. Kurumada and Jaeger (2013) investigated Japanese speakers’ production of the accusative case marker on direct object nouns such as *student* and *fire engine*; the accusative case marking is optional in spoken Japanese. Kurumada and Jaeger (2013) found higher rates of case marking for sentences that could be more ambiguous for the comprehender, a result that they attributed to producers’ aiming for communicative efficiency, i.e., using case marking when it is more necessary and omitting it when it is less essential. Thus across several different subfields, researchers are examining very

closely related phenomena concerning overt mention or omission and addressing questions of choice of form and the forces shaping those choices, so that studies of sentence form and studies of referential form should be able to inform each other.

#### **ALTERNATIVE ACCOUNTS: MESSAGE FACTORS, AUDIENCE DESIGN, COMMUNICATIVE EFFICIENCY**

We interpret speakers' use of null subjects as emergent from internal interference in speech planning, meaning that at least part of the motivation for omission is driven by producers' needs. Here we consider some potential alternative interpretations of these results and identify opportunities for future research to shed light on these alternatives.

Because the pictures in picture description experiments necessarily differ across conditions, it is always possible that producers' utterances are affected by some feature of the pictures other than the target of the experimental manipulation. Thus it is possible that the present null subject findings and Gennari et al. (2012) agentless passive results are due to some differences in the pictures in the semantically similar and dissimilar conditions. For example, the visual salience of to-be-described pictured elements is known to affect speakers' sentence structures in picture descriptions, perhaps because the task demands may implicitly encourage different amounts of description for visually salient vs. non-salient entities (Montag and MacDonald, 2014). Gennari et al.'s (2012) animate and inanimate entities do differ in salience (Montag and MacDonald, 2014), but in the present study, in which the action picture contains only two entities, both animate and inanimate conditions seem to have highly salient objects. Similarly, in Gennari et al.'s (2012) study with all animate entities, the pictures contained only three salient humans, without any apparent differences in salience across conditions. Thus it seems unlikely that visual salience or other picture properties affected rates of agent omission in the current picture description study or in Gennari et al. (2012). Moreover, there were no pictures in Study 1 here, which found more null subject in the speech/text corpus in all-animate sentences than in ones with inanimate objects.

A second possibility is that the message to be conveyed is different across the different picture conditions in a way that affects the felicity of overt mention of an agent. This possibility seems more relevant to some studies than others. For example, in Gennari et al.'s (2012) all-animate study, the high-similarity participants may have yielded more plausible scenarios than the low-similarity condition. Thus producers may have mentioned the agent of the action (i.e., used full passives like *The builder who was slapped by the astronaut* rather than agentless passives like *the builder who was slapped*) more often in the low similarity condition (astronaut slapping builder) than in the high similarity condition (miner slapping builder) because the low-similarity scene was more unusual, making the astronaut-agent more worthy of mention than the miner-agent. That explanation does not appear to hold for Gennari et al. (2012) animacy manipulations (e.g., holding a vase vs. a baby don't appear to have wide variations in plausibility) nor does it hold for the Mandarin production study here, where the two conditions were explicitly matched for plausibility. Thus while messages by necessity differ in

these animacy/semantic overlap manipulations, they appear not to be an obvious source of variation in pro-drop or other agent omissions.

Another potential alternative interpretation is that the speakers may vary the inclusion/omission of an agent to facilitate listeners' comprehension, in a form of audience design. On this view, speakers might omit agents that are similar to patients to help comprehenders avoid similarity-based interference. Similarity-based interference does exist in comprehension of at least complex sentences (Acheson and MacDonald, 2011; Van Dyke et al., 2014), but there are also priming effects (facilitation) from semantic overlap in comprehension (see Ledoux et al., 2006, for review). Thus there is not a straightforward argument for how agent omission would help the comprehender under some conditions and not others, and even if there were such an explanation, it is not clear how producers would calculate during online production when an omission would/wouldn't be helpful to the comprehender. Relatedly, the referential form literature (full noun phrases vs. pronouns) has considered the degree to which choice of form is made for the comprehender (e.g., Arnold et al., 2000; Fukumura et al., 2011). Comprehension studies that compare readers' processing of repeated full noun phrases vs. pronouns suggest that repeated noun phrases hinder comprehension compared to pronouns (Gordon et al., 1993; Gordon and Hendrick, 1997; Kennison and Gordon, 1997). One study of overt vs. null referential forms in comprehension found that for Mandarin comprehenders, overt pronouns and null forms were both easier than full repeated noun phrases (Yang et al., 1999). Yang et al. (1999) argued that overt and null pronoun forms contributed equally to discourse coherence. This finding does not support an audience design account of the null subject phenomena investigated here. It is likely that in some languages or some situations, the discourse status of null and overt pronouns are different to the point that one form is far more appropriate to convey a producer's message than another; indeed we saw almost no pronouns in the corpus analysis. In the picture description study, however, Mandarin speakers routinely produced both overt pronouns and null subjects, and their subject omissions are consistent with an explanation based on interference within utterance planning rather than being an audience design strategy to enhance comprehensibility for the perceiver.

In sum, in this as in all examples of variation in utterance form, producers' choices are likely to be multiply determined by message, production difficulty, and the need to be understood. It is unlikely that a single explanation for a choice of utterance form exists. Indeed, Jaeger et al. (2012) appear to advocate this multi-factor position when they argue that producers' choices can be traced to communicatively efficient production (Jaeger, 2013; Kurumada and Jaeger, 2013). On this view, choices of inclusion/omission of agents in the present studies and in Gennari et al. (2012) might be viewed as owing to communicative efficiency, that in some cases it is more efficient to omit the agent and in others to include it. Our argument here is not against communicative efficiency or other arguments for multiple forces shaping utterance form. Rather, our position is that "efficiency" needs to be engaged at a more mechanistic level with more specific hypotheses concerning (among other forces) the sources of production difficulty (MacDonald, 2013b). We see the current attempts to

link similarity-based interference and choice of utterance form as steps in that direction.

## REFERENCES

- Acheson, D. J., and MacDonald, M. C. (2009). Verbal working memory and language production: common approaches to the serial ordering of verbal information. *Psychol. Bull.* 135, 50–68. doi: 10.1037/a0014411
- Acheson, D. J., and MacDonald, M. C. (2011). The rhymes that the reader perused confused the meaning: phonological effects during on-line sentence comprehension. *J. Mem. Lang.* 65, 193–207. doi: 10.1016/j.jml.2011.04.006
- Allum, P. H., and Wheeldon, L. R. (2007). Planning scope in spoken sentence production: the role of grammatical units. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 791–810. doi: 10.1037/0278-7393.33.4.791
- Almor, A., and Nair, V. A. (2007). The form of referential expressions in discourse. *Lang. Linguist. Compass* 1, 84–99. doi: 10.1111/j.1749-818X.2007.00009.x
- Anderson, J. R. (1983). Retrieval of information from long-term memory. *Science* 220, 25–30. doi: 10.1126/science.6828877
- Arnold, J. E. (2010). How speakers refer: the role of accessibility. *Lang. Linguist. Compass* 4, 187–203. doi: 10.1111/j.1749-818X.2010.00193.x
- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., and Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course for pronoun resolution from eyetracking. *Cognition* 76, B13–B26. doi: 10.1016/S0010-0277(00)00073-1
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Biberauer, T., Holmberg, A., Roberts, I., and Sheehan, M. (2010). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.
- Bock, J. (1986). Meaning, sound, and syntax: lexical priming in sentence production. *J. Exp. Psychol. Learn. Mem. Cogn.* 12, 575–586. doi: 10.1037/0278-7393.12.4.575
- Bock, J., and Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21, 47–67. doi: 10.1016/0010-0277(85)90023-X
- Bock, K., Loebell, H., and Morey, R. (1992). From conceptual roles to structural relations: bridging the syntactic cleft. *Psychol. Rev.* 99, 150–171. doi: 10.1037/0033-295X.99.1.150
- Camacho, J. (2013). *Null subjects*. Cambridge: Cambridge University Press.
- Christianson, K., and Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition* 98, 105–135. doi: 10.1016/j.cognition.2004.10.006
- Damian, M. F., and Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychon. Bull. Rev.* 10, 111–117. doi: 10.3758/BF03196474
- Dell, G. S., and Reich, P. A. (1981). Stages in sentence production: an analysis of speech error data. *J. Verbal Learn. Verbal Behav.* 20, 611–629. doi: 10.1016/S0022-5371(81)90202-4
- de Smedt, K. (1996). “Computational models of incremental grammatical encoding,” in *Computational Psycholinguistics: AI and Connectionist Models of Human Language Processing*, eds A. Dijkstra and K. de Smedt (Philadelphia, PA: Taylor & Francis), 279–307.
- Ferreira, F., and Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *J. Mem. Lang.* 46, 57–84. doi: 10.1006/jmla.2001.2797
- Fukumura, K., and van Gompel, R. G. (2011). The effect of animacy on the choice of referring expression. *Lang. Cogn. Process.* 26, 1472–1504. doi: 10.1080/01690965.2010.506444
- Fukumura, K., van Gompel, R. G., Harley, T., and Pickering, M. J. (2011). How does similarity-based interference affect the choice of referring expression? *J. Mem. Lang.* 65, 331–344. doi: 10.1016/j.jml.2011.06.001
- Gennari, S. P., Mirković, J., and MacDonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cogn. Psychol.* 65, 141–176. doi: 10.1016/j.cogpsych.2012.03.002
- Glaser, W. R., and Döngelhoff, F. (1984). The time course of picture-word interference. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 640–654. doi: 10.1037/0096-1523.10.5.640
- Gordon, P. C., Grosz, B. J., and Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cogn. Sci.* 17, 311–347. doi: 10.1207/s15516709cog1703\_1
- Gordon, P. C., and Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition* 62, 325–370. doi: 10.1016/S0010-0277(96)00788-3
- Hsiao, Y., and MacDonald, M. (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Front. Psychol.* 4:767. doi: 10.3389/fpsyg.2013.00767
- Jaeger, T. F. (2013). Production preferences cannot be understood without reference to communication. *Front. Lang. Sci.* 4:230. doi: 10.3389/fpsyg.2013.00230
- Jaeger, T. F., Furth, K., and Hilliard, C. (2012). Incremental phonological encoding during unscripted sentence production. *Front. Psychol.* 3:481. doi: 10.3389/fpsyg.2012.00481
- Janssen, N., and Caramazza, A. (2009). Grammatical and phonological influences on word order. *Psychol. Sci.* 20, 1262–1268. doi: 10.1111/j.1467-9280.2009.02429.x
- Kempen, G., and Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cogn. Sci.* 11, 201–258. doi: 10.1207/s15516709cog1102\_5
- Kennison, S. M., and Gordon, P. C. (1997). Comprehending referential expressions during reading: evidence from eye tracking. *Discourse Process.* 24, 229–252. doi: 10.1080/01638539709545014
- Kurumada, C., and Jaeger, T. F. (2013). “Communicatively efficient language production and case-marker omission in Japanese,” in *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, eds P. Bello, M. Guarini, M. McShane, and B. Scassellati (Austin, TX: Cognitive Science Society), 858–863.
- Ledoux, K., Camblin, C. C., Swaab, T. Y., and Gordon, P. C. (2006). Reading words in discourse: the modulation of lexical priming effects by message-level context. *Behav. Cogn. Neurosci. Rev.* 5, 107–127. doi: 10.1177/1534582306289573
- Levelt, W. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT Press.
- Li, C. N., and Thompson, S. A. (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA: University of California Press.
- MacDonald, M. C. (2013a). How language production shapes language form and comprehension. *Front. Psychol.* 4:226. doi: 10.3389/fpsyg.2013.00226
- MacDonald, M. C. (2013b). Production is at the left edge of the PDC but still central: response to commentaries. *Front. Psychol.* 4:227. doi: 10.3389/fpsyg.2013.00227
- Meyer, A. S., and Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: effects of stimulus onset asynchrony and types of interfering stimuli. *J. Exp. Psychol. Learn. Mem. Cogn.* 17:1146–1160. doi: 10.1037/0278-7393.17.6.1146
- Montag, J. L., and MacDonald, M. C. (2014). Visual salience modulates sentence choice in relative clause production. *Lang. Speech* 57, 163–180. doi: 10.1177/0023830913495656
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. *Basic Process. Read. Vis. Word Recogn.* 11, 264–336.
- Rosinski, R. R., Golinkoff, R. M., and Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Dev.* 46, 247–253. doi: 10.2307/1128859
- Schriefers, H., Meyer, A. S., and Levelt, W. J. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *J. Mem. Lang.* 29, 86–102. doi: 10.1016/0749-596X(90)90011-N
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Tanaka, M., Branigan, H. P., and Pickering, M. J. (2005). The role of animacy in Japanese sentence production. *Paper Presented at the 18th Annual CUNY Conference on Human Sentence Processing*, Tucson, AZ.
- Van Dyke, J. A., Johns, C. L., and Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition* 131, 373–403. doi: 10.1016/j.cognition.2014.01.007
- Wagner, V., Jescheniak, J. D., and Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: effects of cognitive load on multiple lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 423–440. doi: 10.1037/a0018619
- Wang, L. (1987). *Xian Dai Han Yu Yu Fa [Modern Chinese Grammar]*. Taiwan: Lan Deng Wen Hua.
- Wang, Q. (2007). *An Investigation on the Semantics and Discourse and Sociopragmatic Functions of Personal Pronouns in Mandarin Chinese (Master's thesis)*. Available at: <http://www.airitilibrary.com/Publication/alDetailedMesh1?DocID=U0021-0712200716141159#References>
- Wilshire, C. E. (1998). Serial order in phonological encoding: an exploration of the word onset effect using laboratory-induced errors. *Cognition* 68, 143–166. doi: 10.1016/S0010-0277(98)00045-6

Xue, N., Jiang, Z., Zhong, X., Palmler, M., Xia, F., Chiou, F., et al. (2010). *Chinese Treebank 7.0*. Philadelphia: Linguistic Data Consortium.

Yang, C.-L., Gordon, P. C., Hendrick, R., and Wu, J. T. (1999). Comprehension of referring expressions in Chinese. *Lang. Cogn. Process.* 14, 715–743. doi: 10.1080/016909699386248

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 June 2014; accepted: 26 August 2014; published online: 16 September 2014.

*Citation:* Hsiao Y, Gao Y and MacDonald MC (2014) Agent-patient similarity affects sentence structure in language production: evidence from subject omissions in Mandarin. *Front. Psychol.* 5:1015. doi: 10.3389/fpsyg.2014.01015

This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Hsiao, Gao and MacDonald. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Different Loci of Semantic Interference in Picture Naming vs. Word-Picture Matching Tasks

Denise Y. Harvey<sup>1,2</sup> and Tatiana T. Schnur<sup>3\*</sup>

<sup>1</sup> Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA, <sup>2</sup> Moss Rehabilitation Research Institute, Elkins Park, PA, USA, <sup>3</sup> Baylor College of Medicine, Department of Neurosurgery, Houston, TX, USA

## OPEN ACCESS

### Edited by:

Ian FitzPatrick,  
Heinrich Heine Universität Düsseldorf,  
Germany

### Reviewed by:

Katharina Spalek,  
Humboldt-Universität zu Berlin,  
Germany

Markus F. Damian,  
University of Bristol, UK

### \*Correspondence:

Tatiana T. Schnur  
ttschnur@gmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 16 September 2014

**Accepted:** 27 April 2016

**Published:** 13 May 2016

### Citation:

Harvey DY and Schnur TT (2016)  
Different Loci of Semantic Interference  
in Picture Naming vs. Word-Picture  
Matching Tasks.  
*Front. Psychol.* 7:710.  
doi: 10.3389/fpsyg.2016.00710

Naming pictures and matching words to pictures belonging to the same semantic category impairs performance relative to when stimuli come from different semantic categories (i.e., semantic interference). Despite similar semantic interference phenomena in both picture naming and word-picture matching tasks, the locus of interference has been attributed to different levels of the language system – lexical in naming and semantic in word-picture matching. Although both tasks involve access to shared semantic representations, the extent to which interference originates and/or has its locus at a shared level remains unclear, as these effects are often investigated in isolation. We manipulated semantic context in cyclical picture naming and word-picture matching tasks, and tested whether factors tapping semantic-level (generalization of interference to novel category items) and lexical-level processes (interactions with lexical frequency) affected the magnitude of interference, while also assessing whether interference occurs at a shared processing level(s) (transfer of interference across tasks). We found that semantic interference in naming was sensitive to both semantic- and lexical-level processes (i.e., larger interference for novel vs. old and low- vs. high-frequency stimuli), consistent with a semantically mediated lexical locus. Interference in word-picture matching exhibited stable interference for old and novel stimuli and did not interact with lexical frequency. Further, interference transferred from word-picture matching to naming. Together, these experiments provide evidence to suggest that semantic interference in both tasks originates at a shared processing stage (presumably at the semantic level), but that it exerts its effect at different loci when naming pictures vs. matching words to pictures.

**Keywords:** semantic interference, lexical access, semantic access, generalization of interference, lexical frequency

## INTRODUCTION

Accessing words (lexical representations) and meanings (semantic representations) from the same vs. different categories can interfere with future access from the category. For example, patients with aphasia due to stroke tend to make semantic errors when naming pictures and/or matching words to pictures in the context of semantically related words (e.g., Schnur et al., 2006; Biegler et al., 2008; Harvey and Schnur, 2015). Moreover, naming pictures (e.g., Kroll and Stewart, 1994) or matching words to pictures (Campanella and Shallice, 2011) belonging to the

same semantic category has a detrimental effect on healthy participants' performance, known as semantic interference. That both picture naming and word-picture matching performance is sensitive to semantic contexts demonstrates that both tasks are semantically mediated (see Belke, 2013). However, because semantic interference in picture naming and word-picture matching tasks is usually investigated separately, this has led to different conclusions about the locus of interference in each task. In picture naming, evidence suggests that interference arises when mapping from semantic to lexical representations (hereafter, lexical locus; Levelt et al., 1999; Howard et al., 2006; Oppenheim et al., 2010), whereas in word-picture matching tasks evidence suggests that interference arises within the semantic system itself (Warrington and McCarthy, 1983, 1987; Warrington and Cipolotti, 1996; Forde and Humphreys, 1997, 2007; Gotts and Plaut, 2002; Campanella and Shallice, 2011). While semantic interference in picture naming tasks has been largely explored in healthy subjects, semantic interference in word-picture matching tasks is less often reported in the healthy population (cf. Biegler et al., 2008; Campanella and Shallice, 2011; Wei and Schnur, 2016). Here, we investigated the locus of semantic interference in naming and word-picture matching by testing in healthy participants whether interference was sensitive to semantic and lexical factors and transferred between the two tasks. Finding that interference is affected by the same factors and/or transfers across the two tasks can elucidate the extent to which processes governing access to semantic and lexical representations operate similarly across the two tasks. In turn, this work informs theories of lexical-semantic access, providing clues about the organization of the language system as a whole.

In both picture naming and word-picture matching tasks, repeatedly accessing semantically related stimuli has a negative effect on performance. For example, participants are slower to name pictures or match words to pictures when trials depict items belonging to the same categories (related context: e.g., CAT, DOG, BEAR, and COW) vs. different categories (unrelated context: e.g., CAT, TRAIN, SHIRT, and DESK)<sup>1</sup> (i.e., blocked naming and word-picture matching tasks; e.g., Damian et al., 2001; Damian and Als, 2005; Campanella and Shallice, 2011). Interference is thought to occur because activating the semantic system to produce a target word (i.e., "dog") or access a word's meaning (i.e., DOG) results in the co-activation of related words and meanings (e.g., "cat" and CAT) due to the high degree of semantic feature overlap amongst members of the same category (e.g., Collins and Loftus, 1975; see also Vigliocco et al., 2002; Forde and Humphreys, 2007). This is evidenced by the findings of graded semantic interference effects in both tasks (i.e., larger interference for semantically close vs. distant category members; naming: Vigliocco et al., 2002; Navarrete et al., 2012; word-picture matching: Crutch and Warrington, 2005, Experiment 1; Warrington and Cipolotti, 1996, Experiment 5). That naming and word-picture matching are sensitive to semantic contexts

demonstrates that interference in both tasks originates at the semantic level.

However, the locus of semantic interference in each task is thought to differ. By most accounts, semantic interference in naming exerts its effects at the lexical level (e.g., Howard et al., 2006; Oppenheim et al., 2010; see also Damian and Als, 2005; cf. Levelt et al., 1999; Damian et al., 2001), whereas semantic interference in word-picture matching exerts its effects at the semantic level (e.g., Forde and Humphreys, 1997, 2007; Campanella and Shallice, 2011). Computational models of semantic interference in naming (Howard et al., 2006; Oppenheim et al., 2010; see also Roelofs, 1992) assume that naming a picture (i.e., DOG) activates its lexical representation (i.e., "dog") and those sharing semantic features with the target (e.g., "cat") to a greater extent than those that do not share semantic features with the target (e.g., "shoe"). Producing the word "dog" increases its lexical representation's activation level, which negatively affects the subsequent selection of same-category lexical representations (e.g., "cat"). Accordingly, theories of semantic interference in naming (e.g., Howard et al., 2006; Oppenheim et al., 2010) assume that shared activation at the semantic level causes interference that exerts its effects at a lexical level. Theories of semantic interference in word-picture matching assume that activating the meaning of a given word (i.e., "dog") also activates related word meanings (e.g., CAT), which interfere with the ability to distinguish between same-category meanings on subsequent trials (Forde and Humphreys, 1997, 2007; see also Warrington and Cipolotti, 1996; Gotts and Plaut, 2002). Thus, semantic interference in naming originates at the semantic level, but has a lexical-level locus (see also Belke, 2013), whereas semantic interference in word-picture matching both originates and has its locus at the semantic level.

That semantic contexts are thought to interfere with word-picture matching performance at a semantic level seemingly contradicts the generally accepted view that semantic contexts facilitate performance on tasks requiring semantic but not lexical access for spoken output (Bajo, 1988; Belke, 2013). For example, semantic relationships facilitate the recognition of words preceded by a semantically related prime word (i.e., lexical decision task; e.g., McRae and Boisvert, 1998) and the categorization of pictured objects based on the direction (i.e., left or right) they face (i.e., orientation judgment task; Damian et al., 2001)<sup>2</sup> or based on their superordinate category (i.e., man-made or natural) membership (i.e., semantic classification task; Belke, 2013; see also Vitkovitch and Humphreys, 1991). However, tasks argued to tap semantic level facilitatory processes differ in a number of respects with those eliciting semantic interference. In the lexical decision task, "...co-activation of other words would not be costly because the task only requires participants to decide whether the presented string is a word or not..." (Vigliocco et al., 2004, p. 468) but not whether the word refers to a specific meaning. Further, judging the orientation of a pictured object (i.e., tip of a shoe) in terms of which

<sup>1</sup>We use quotations to denote lexical representations and picture name responses (e.g., "dog"), whereas capitalization denotes the semantic representation corresponding to a word (e.g., DOG).

<sup>2</sup>Damian et al. (2001) do not report if the facilitation observed in the orientation task was statistically significant. However, the RTs reported for 10 subjects are in the predicted direction: orientation judgments were faster for semantically related compared to unrelated objects (i.e., 388 ms vs. 396 ms, respectively).

direction it faces relies more on decoding the visual properties of the object (e.g., Humphreys et al., 1988, 1995) and not necessarily the semantic features corresponding to the object (see Belke, 2013). Lastly, the semantic classification task while likely requiring access to semantic information, does not necessitate accessing fine-grained semantic level distinguishing information, as all members of a semantic category are consistent with the classification of man-made or natural. By contrast, matching a word to its corresponding picture necessitates making fine-grained semantic decisions about the set of semantic features associated with that particular word, which in some ways is like naming a picture, as picture naming necessitates the selection of a word based on the set of semantic features that distinguish the target lexical representation from co-activated, semantically related, lexical representations (see Wei and Schnur, 2016 for a similar discussion). Thus, the assumption that semantic contexts facilitate processing at the semantic level may be an artifact of the types of tasks used to tap semantic-level processes (see Chen and Mirman, 2012 for a similar argument).

Does semantic interference occur in the healthy semantic system when discriminating a target from related meanings? Evidence of semantic interference in word-picture matching almost exclusively comes from neuropsychological studies of patients with aphasia secondary to stroke (cf. Biegler et al., 2008; Campanella and Shallice, 2011; Wei and Schnur, 2016). Consequently, the extent to which the healthy semantic system operates similarly when accessing words and meanings is not well understood. To our knowledge, only a few studies have investigated semantic interference in healthy younger adults' word-picture matching performance, demonstrating that semantic interference occurs in tasks tapping semantic-level processes (Campanella and Shallice, 2011; Mirman and Graziano, 2012; Wei and Schnur, 2016; see Biegler et al., 2008 for evidence of semantic interference in healthy older adults' word-picture matching performance). What remains unclear is whether the semantic context effects observed in word-picture matching occur due to the same processes that create interference in naming.

## The Current Research

The main goal of this research was to investigate whether semantic interference in naming and word-picture matching originate and/or exert their effects at a shared processing level(s). Accordingly, we explored how factors that tap semantic and lexical processing affect semantic interference in each task, and whether semantic interference at a shared processing level(s) allows for the effect to transfer across tasks.

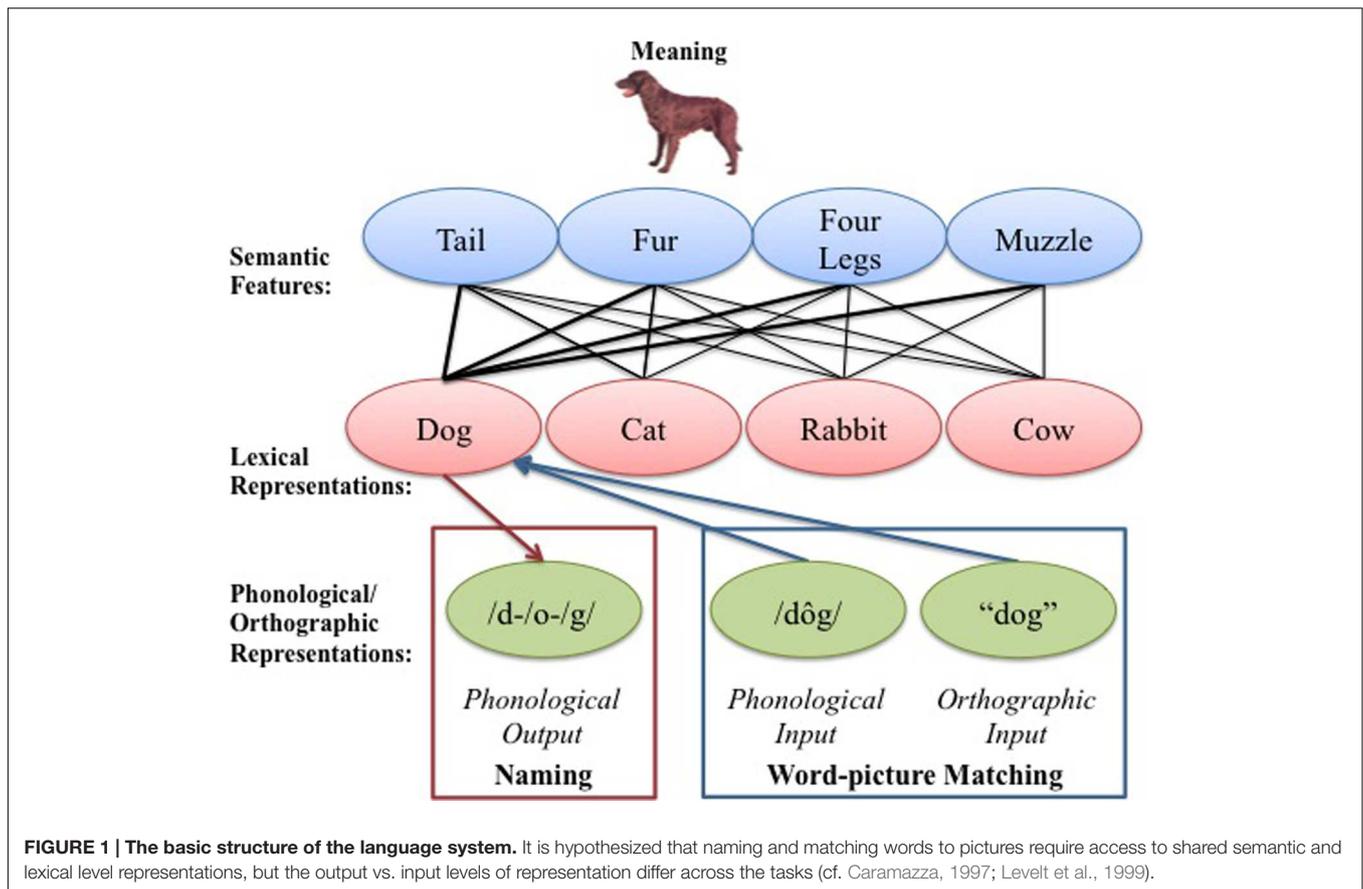
Here, we used cyclical variants of the blocked naming and word-picture matching tasks, where subjects name pictures or match words to pictures in related vs. unrelated contexts, and target items repeat multiple times (cycles) in different orders (e.g., Kroll and Stewart, 1994; Damian, et al., 2001; Campanella and Shallice, 2011; see also Wei and Schnur, 2016). Whether assuming a lexical- or semantic-level locus, interference is thought to emerge with repetition because competition increases with repeated access to same- vs. different-category items (e.g., Forde and Humphreys, 1995, 2007; Belke et al., 2005b; cf.

Navarrete et al., 2014 for an alternative account in naming). We hypothesize that because both picture naming and word-picture matching tasks require mapping between shared lexical and semantic representational levels (see **Figure 1**; reviewed in Howard, 1995; Levelt, 1999; cf. Caramazza, 1997), it suggests a shared origin and/or locus of interference in the two tasks. Specifically, in both tasks it is necessary to access the semantic features corresponding to the target representation (picture or word form) – a process that results in the co-activation of related representations. However, because the order with which lexical and semantic representations are activated occurs in reverse in the two tasks (semantic-to-lexical in naming and vice versa in word-picture matching), the level at which co-activated representations interfere with performance is thought to differ. Consequently, it remains an open question as to whether semantic interference in the two tasks is a reflection of the same underlying phenomena occurring at shared semantic and/or lexical representational levels.

## Origin of Interference

Because interference is assumed to originate at the semantic level in both picture naming and word-picture matching, it is predicted to generalize to novel category members (Forde and Humphreys, 1995; Belke et al., 2005b) and transfer across tasks (Belke, 2013). In naming, shared activation at the semantic level gives rise to the co-activation of related lexical representations (Howard et al., 2006; Oppenheim et al., 2010), resulting in the accumulation of semantic interference for both previously named and novel category members (e.g., Belke et al., 2005b). Similarly, in word-picture matching, accessing semantically related word meanings in succession renders disambiguating both previously accessed and novel word meanings belonging to the same category more difficult (e.g., Forde and Humphreys, 1995, 2007). Moreover, a shared semantic-level origin of interference in naming and word-picture matching predicts that interference will transfer across tasks. For example, if accessing the semantic system in word-picture matching results in the co-activation of both related semantic and lexical representations, then this should interfere with subsequent naming of novel same-category pictures. Thus, interference originating at the semantic level predicts both generalization of interference within each task, and transfer of interference across tasks.

To our knowledge, there are only two studies that have demonstrated interference generalization using the blocked-cyclic paradigms: one in healthy subjects' picture naming performance (Belke et al., 2005b, Experiment 3) and the other in an aphasic patient's comprehension performance (patient J.M.; Forde and Humphreys, 1995, Experiment 12). Belke et al. (2005b) examined whether interference generalized to cycles of naming novel items semantically related to those named previously. They found that semantic interference emerged after the first cycle and remained unchanged across subsequent cycles of both previously named and novel pictures. Forde and Humphreys (1995) found that in comprehension (i.e., auditory-to-written word matching), the interference effect was larger for the first cycle of novel words relative to the first cycle



of semantically related “old” words. Thus, generalization of semantic interference has been quantified in two different ways. Belke et al. (2005b) defined generalization as an unchanging effect (once it emerges after cycle 1) across old and novel items (cycles 2–8), whereas Forde and Humphreys defined generalization as larger interference for the first cycle of novel compared to the first cycle of old items. In the experiments reported here, we assessed generalization as an increase in interference for novel compared to old items collapsed across cycles because both characterizations (i.e., no interference at cycle 1 followed by unchanging interference for cycles 2–8, Belke et al., 2005b and larger interference for the first cycle of novel vs. the first cycle of old items, Forde and Humphreys, 1995) should when averaged across cycles yield larger interference for novel vs. old items. The first goal of this study was to replicate and extend the findings of interference generalization obtained in previous studies using blocked-cyclic naming and word-picture matching tasks to demonstrate that interference in both tasks originates at the semantic level.

### Locus of Interference

Because theories of semantic interference in naming and word-picture matching tasks assume different semantic interference loci (e.g., Forde and Humphreys, 2007; Oppenheim et al., 2010), this generates the prediction that lexical frequency, a factor thought to exert its effects primarily at the name retrieval

stage (e.g., Vitkovitch and Humphreys, 1991),<sup>3</sup> should affect semantic interference in naming but not word-picture matching as word-picture matching does not require access to lexical representations for spoken output (see Campanella and Shallice, 2011 for discussion). That subjects name pictures depicting high-frequency words faster than those depicting low-frequency words (e.g., Oldfield and Wingfield, 1965) and recognize high- vs. low-frequency words faster in word recognition tasks such as the lexical decision task (e.g., Scarborough et al., 1977; Balota and Chumbley, 1985) indicates that high- vs. low-frequency lexical representations have increased activation levels (Morton, 1969; McClelland and Rumelhart, 1981; Dell, 1986; Seidenberg and McClelland, 1989; Plaut et al., 1996; Caramazza, 1997; Barry et al., 2001; Kittredge et al., 2008), rendering them more available for selection in naming and identification in word recognition. Thus, a lexical, but not semantic, locus of interference predicts that interference will be affected by the lexical frequency of semantically related words when naming pictures, but not when matching words to pictures.

<sup>3</sup>Most agree that lexical frequency does *not* reflect semantic-level processes (e.g., Wingfield, 1968; Bartram, 1976; Meyer et al., 1998), but whether frequency has a lexical or phonological locus is debated (e.g., Jescheniak and Levelt, 1994; Finocchiaro and Caramazza, 2006; Navarrete et al., 2006). Because practice via frequently producing a word most likely strengthens connections between semantic-to-lexical *and* lexical-to-phonological representational levels, it is likely that lexical frequency effects are represented at both levels (if two levels are assumed: lexical vs. phonological; e.g., cf. Caramazza, 1997; Levelt et al., 1999).

Although lexical frequency is predicted to interact with semantic interference in naming but not word-picture matching, previous studies investigating these factors provide equivocal results. To our knowledge, there is only one study that examined the effect of lexical frequency on response times (RTs) in blocked-cyclical naming and although there was an overall effect of lexical frequency on naming, it did not interact with semantic interference (Santesteban et al., 2006). Campanella and Shallice (2011) manipulated semantic context (close vs. distant) and lexical frequency (high- vs. low-frequency) in a non-cyclical word-picture matching task, and found that healthy subjects were slower and less accurate in the semantically close, low-frequency condition compared to all other conditions (i.e., semantically close, high-frequency, semantically distant high-frequency, and semantically distant, low-frequency; Experiment 1). Experiment 2 used a cyclical variant of the task, testing only those items that gave rise to the largest interference effects in Experiment 1 (i.e., semantically close, low-frequency words), and found that interference increased across cycles of repeated word-picture matching. That semantic interference was numerically larger for low- compared to high-frequency words (Experiment 1) contradicts a semantic locus of interference, suggesting instead that interference in word-picture matching has a lexical locus. Thus, the second aim of this study was to test whether lexical frequency interacts with semantic interference in naming and/or word-picture matching to determine whether or not the locus of interference is shared across the two tasks.

Lastly, it remains an open question whether or not semantic interference observed in picture naming and word-picture matching arises due to the same or partially overlapping processing stages. To our knowledge, no one has tested whether semantic interference transfers across the two tasks. However, previous work has examined interference transfer to and from different levels of the language system, but the evidence here is mixed. While Navarrete et al. (2010) found that interference transferred from a task tapping semantically mediated lexical retrieval (i.e., picture + determiner naming) to one requiring lexical retrieval without semantic mediation (i.e., word + determiner naming) but not vice versa (Experiment 3), Belke (2013) did not replicate this finding (Experiment 4). Moreover, Belke (2013) demonstrated that picture naming affected subsequent semantic classification (i.e., man-made or natural) of categorically related objects but not vice versa (Experiment 5), which conflicts with previous evidence that semantic classification affects the subsequent naming of categorically related pictures (Vitkovitch and Humphreys, 1991, Experiment 2). Consequently, the extent to which interference transfers across tasks tapping shared representational levels remains unclear. Thus, the third goal of this study was to investigate whether a shared origin and/or locus of interference exists, as evidenced by increased semantic interference (and thus transfer) when performing the naming (or word-picture matching) task on novel items categorically related to those which appeared previously in the word-picture matching (or naming) task. If the origin *and* locus of interference is shared in naming and word-picture matching, then interference will be sensitive to both semantic and lexical factors within each task

(Experiments 1 and 2) and transfer across tasks (Experiment 3). However, if interference has a shared origin but different loci, then interference will generalize within tasks and transfer across tasks, but only inference in naming will interact with lexical frequency.

## MATERIALS AND METHODS

### Participants

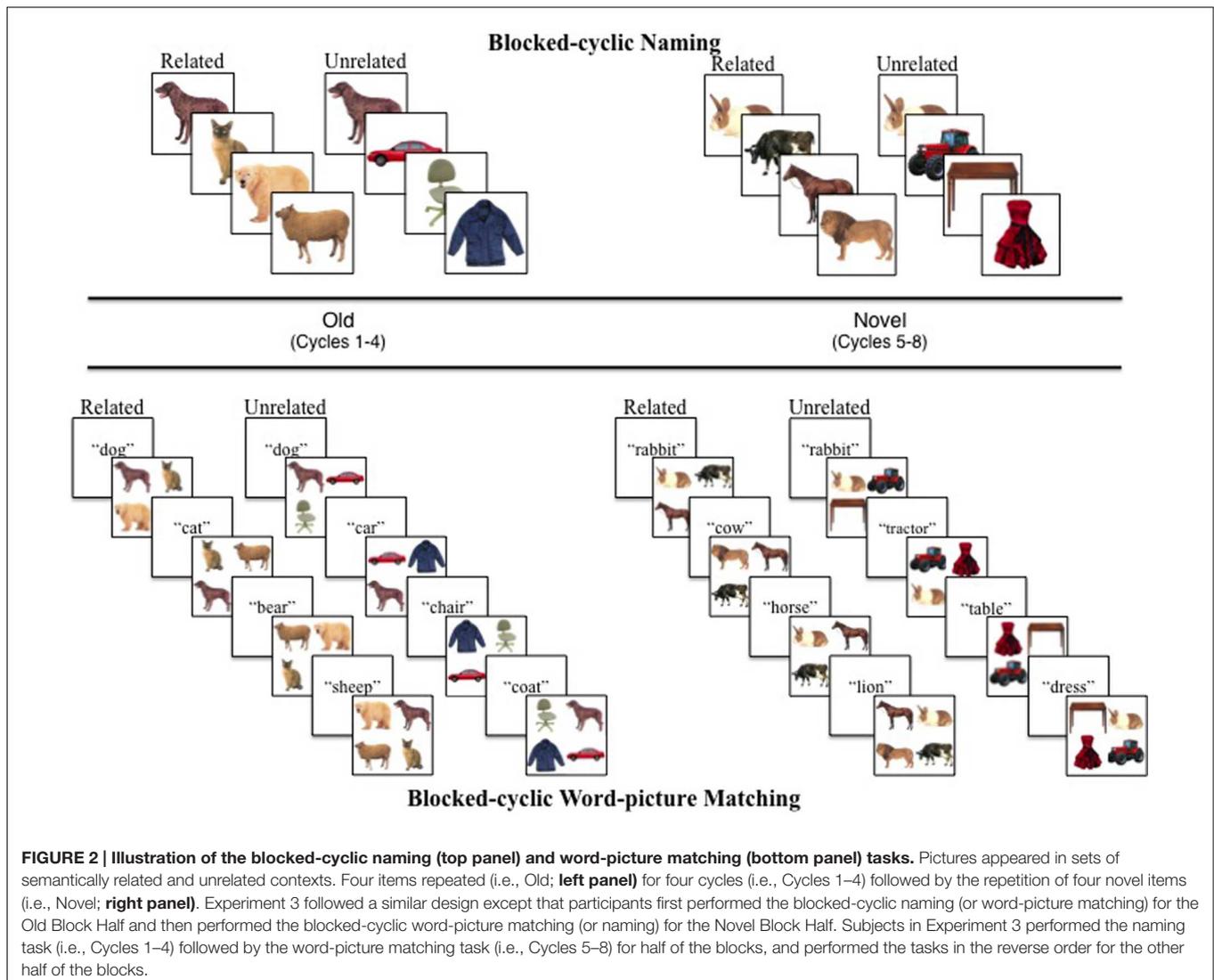
There were 94 participants total. Thirty-one participated in Experiment 1 [15 female, 16 male; mean (and range) age: 19 years (18–21)], 20 participated in Experiment 2 [12 female, 8 male; mean (and range) age: 19 years (18–22)], and 43 in Experiment 3 [25 female, 18 male; mean (and range) age: 19 years (18–22)]. Data from four participants who took part in Experiment 1 were excluded: two due to experimenter error and two due to equipment error. All were native English speakers with normal or corrected to normal vision attending Rice University, and received course credit for their participation. Informed consent in accordance with the IRB at Rice University was obtained from each participant.

### Materials and Design

Stimuli were 64 colored pictures of familiar objects belonging to eight semantic categories. Pictures were taken from the Bank of Standardized Stimuli (BOSS; Brodeur et al., 2010) and another image database (Viggiano et al., 2004), and scaled to 400 pixels × 400 pixels. Within each category, pictures consisted of all high- or low-frequency names, and were selected to minimize differences in other factors known to correlate with lexical frequency, such as familiarity and imageability (e.g., Morrison et al., 1997). Measures of lexical frequency, familiarity, and imageability of target stimuli were obtained from an online database<sup>4</sup> (see also Wilson, 1988). Half of the categories depicted objects with high-frequency names (mean 59.72; range 41–86), whereas the other half depicted objects with low-frequency names (mean 8.75; range 5–15; see Appendix A). Lexical frequency differed significantly for high- and low-frequency categories [ $t(62) = 5.02, p < 0.00001$ ], even after controlling for indices of imageability and familiarity [ $F(1, 60) = 20.44, p < 0.001$ ]. Picture names were either mono- or disyllabic, and the number of syllables did not differ between high- and low-frequency categories [ $t(62) = 1.59, p = 0.12$ ]. In Experiments 2 and 3, stimuli also included visually presented written word forms of the 64 target picture names.

Items in each of the eight semantic categories appeared together to form four high- and four low-frequency related blocks of trials consisting of eight items each. One item from each of the high- or low-frequency related categories appeared together in a set to form four high- and low-frequency unrelated blocks of trials, resulting in a total of 16 blocked sets (see Appendix B). Each block consisted of a set of four pictures that repeated for four cycles in different orders (i.e., Old) followed

<sup>4</sup>[http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa\\_mrc.htm#base/uwa\\_mrc.htm](http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm#base/uwa_mrc.htm)



by four cycles repeating the remaining four pictures in the set (i.e., Novel). For example, in the Related Condition, the Old set contained four same-category pictures (e.g., animal: BEAR, CAT, LION, and SHEEP), and the Novel set contained four novel pictures drawn from the same semantic category (e.g., animal: DOG, COW, RABBIT, and HORSE). The 8-item unrelated sets contained two exemplars from each of the four high- or low-frequency semantic categories, where one appeared in the Old Block Half (e.g., BEAR, CAR, SHOE, and CHAIR) and another appeared in the Novel Block Half (e.g., DOG, VAN, SHIRT, and RUG; see **Figure 2**). Stimuli appeared an equal number of times in each condition. Blocked sets appeared in pseudorandom order, such that no more than three blocks of the same Condition (Related and Unrelated) or Frequency (High and Low) appeared consecutively. Following these constraints, we created five stimulus presentation lists. Items appearing in the Old vs. Novel sets were counterbalanced across participants to ensure that any differences between semantic interference effects in the Old and Novel Block Halves were not due to the specific

items used. This resulted in a total of 10 lists of test materials. Together, there were 16 blocks with 32 trials each for a total of 512 trials per subject.

### Apparatus

Target stimuli were presented using DMDX software (Forster and Forster, 2003). To record naming performance (i.e., Experiments 1 and 3), a microphone headset triggered a voice key to collect RTs to the nearest millisecond (ms) and record verbal responses. An experimenter coded naming errors. To record word-picture matching performance (i.e., Experiments 2 and 3), participants made their response using a touch screen monitor, and DMDX software recorded RTs to the nearest ms and error data (i.e., tapping the wrong picture).

### Experiment 1: Semantic Interference in Picture Naming

To establish that semantic interference in blocked-cyclic naming originates at the semantic level but has its locus at the lexical level

(e.g., Belke, 2013), we tested whether (1) repeatedly naming a set of semantically related pictures renders subsequent naming of novel pictures drawn from the same semantic category more difficult (i.e., generalization of semantic interference), and (2) whether the lexical frequency of targets affects semantic interference magnitudes. To test the prediction that interference generalizes to novel category members (Howard et al., 2006; Oppenheim et al., 2010; see also Belke et al., 2005b), in Experiment 1 we examined whether semantic interference increased when naming novel items (collapsed across cycles) in comparison to having previously named different items from the same category. Subjects named sets of semantically related and unrelated pictures across four cycles (i.e., Old) immediately followed by naming novel semantically related or unrelated pictures for an additional four cycles (i.e., Novel; following Belke et al., 2005b). While Belke et al. (2005b) found that interference emerged on the second cycle and remained stable thereafter, a computational simulation of this experiment predicts that semantic interference increases incrementally (linearly) across cycles of previously named and novel category members (Oppenheim et al., 2010, Simulation 3). In either case, larger semantic interference for novel compared to old items (collapsed across cycles) is predicted to occur because either interference is absent on the first cycle of the block and present on later cycles (Belke et al., 2005b) or because interference continues to increase linearly across cycles of both old and novel items (Oppenheim et al., 2010). Thus, although we predict increased interference for novel items categorically related to those named previously (collapsed across cycles), it remains unclear *how* semantic interference in naming generalizes to novel category members in the blocked-cyclic naming task (i.e., stable vs. linear increase), as there is limited evidence to support either account of interference generalization in this task.

To investigate the contribution of lexical-level processing on semantic interference in naming and word-picture matching, we compared semantic interference for high- vs. low-frequency picture names. The semantically related and unrelated sets consisted of objects depicting words with similar frequency counts. This was done to replicate the word-picture matching findings of Campanella and Shallice (2011), and directly compare semantic interference in word-picture matching with that of naming. If semantic interference in naming has a lexical-level locus, we predict greater interference for low- vs. high-frequency picture names because their inherently lower activation levels render them more susceptible to interference from co-activated, same-category high-frequency words with inherently higher activation levels.

## Procedure

Prior to the experiment, subjects were familiarized with the picture stimuli and their corresponding names. In the learning phase, each picture appeared centrally on the computer screen with its written name displayed underneath the picture. The picture and name stayed on the screen until the subject pressed a key indicating that they understood the correct response for the stimulus. To keep the learning phase consistent across experiments, subjects were instructed to not name the pictures,

as naming the pictures could contaminate semantic interference effects observed in the word-picture matching variants of the task (Experiments 2 and 3).

Immediately after the learning phase, the experimental phase began. A single picture appeared in the center of the screen, and subjects were instructed to name the picture as quickly and accurately as possible into the microphone headset. If the microphone failed to trigger the voice key, then the subject would see the words “Speak up” before the next picture appeared which indicated that they should speak more loudly on the next trial. The picture remained on the screen for 1600 ms or until the subject made a response (similar to previous studies using the blocked-cyclic tasks; e.g., Damian and Als, 2005; Campanella and Shallice, 2011). Once a response was made, the next trial began immediately [i.e., 0 ms response stimulus interval (RSI), following Campanella and Shallice, 2011]. Subsequent trials either depicted same category (Related Condition) or different category items (Unrelated Condition). See **Figure 2**. The experiment lasted approximately 20 min.

## Statistical Analyses

We excluded from the analyses RTs for trials classified as an error (i.e., incorrect naming response or no response and voice-key malfunction) and responses faster than 250 ms or slower than 1550 ms (following Damian and Als, 2005). Valid RTs were analyzed using a repeated measures analysis of variance (ANOVA) with participants and items as random factors, yielding  $F_1$  and  $F_2$  statistics, respectively. Fixed factors included Condition (Related and Unrelated), Block Half (Old, Novel), Cycles (1–4), and Frequency (High-frequency and Low-frequency). All fixed factors were considered within-subject, within-item variables except for Frequency, which was a within-subject variable in the  $F_1$  analysis and a between-item variable in the  $F_2$  analysis.

## Results and Discussion

Response errors occurred on 4.4% of experimental trials. **Tables 1** and **2** summarize RT  $F$  statistics and mean RTs, respectively.

There were significant main effects of Condition, Block Half, Cycle, and Frequency. Participants responded more slowly in the Related (685 ms) compared to the Unrelated Condition [670 ms; Condition effect 15 ms, 95% confidence interval (CI) 7–23 ms]. RTs were faster in the Old (669 ms) vs. Novel Block Half (687 ms; Block Half effect 18 ms, 95% CI 12–24 ms). Participants also became faster across naming cycles (755, 667, 654, and 636 ms), replicating previous findings of repetition priming in studies using the blocked-cyclic naming task (e.g., Schnur et al., 2006; Navarrete et al., 2012, 2014). Lastly, naming latencies were faster for high- (666 ms) compared to low-frequency words (690 ms; Frequency effect 24 ms, 95% CI 15–33 ms), which replicates the lexical frequency effect found elsewhere (e.g., Oldfield and Wingfield, 1965; Jescheniak and Levelt, 1994; Griffin and Bock, 1998) and demonstrates the items were sensitive to this variable.<sup>5</sup>

<sup>5</sup>Although the effects of lexical frequency described here could be due to either frequency or age of acquisition (they are highly correlated due to the fact that frequently produced words tend to be those acquired early in life e.g., Brysbaert and

TABLE 1 | Experiment 1 ANOVA results.

	Subject			Item		
	Degrees of freedom		$F_1$	Degrees of freedom		$F_2$
	Numerator	Denominator		Numerator	Denominator	
Condition	1	30	13.51*	1	62	9.89*
Block Half	1	30	33.41*	1	62	40.97*
Cycle	3	90	213.66*	3	186	195.18*
Frequency	1	30	24.07*	1	62	7.60*
Condition × Block Half	1	30	4.83*	1	62	6.10*
Condition × Cycle	3	90	11.67*	3	186	10.96*
Condition × Frequency	1	30	4.32*	1	62	2.90
Block Half × Cycle	3	90	1.90	3	186	2.39
Block Half × Frequency	1	30	0.04	1	62	0.39
Cycle × Frequency	3	90	5.65*	3	186	2.31
Condition × Block Half × Cycle	3	90	2.47	3	186	2.45
Condition × Block Half × Frequency	1	30	4.39*	1	62	4.07*
Condition × Cycle × Frequency	3	90	0.69	3	186	1.12
Block Half × Cycle × Frequency	3	90	0.99	3	186	1.45
Condition × Block Half × Cycle × Frequency	3	90	0.34	3	186	0.36

Summary of  $F$  statistics for the RT ANOVA examining the effects of Condition (Related and Unrelated), Block Half (Old and Novel), Cycles (1–4), and Frequency (High-Frequency and Low-Frequency).

Significant main effects and interactions in both the  $F_1$  and  $F_2$  analyses appear shaded in gray, and an \* indicates a significant effect at  $p < 0.05$ .

Two-way interactions were significant between Condition and Block Half and Cycle. The Condition × Cycle interaction revealed that the semantic interference effect (Related – Unrelated) increased with repetition across cycles (collapsed across Block Half; –5, 9, 27, and 27 ms). The Condition × Block Half interaction revealed that semantic interference increased when naming novel (21 ms) vs. old pictures (8 ms; Condition × Block Half effect 13 ms, 95% CI 1–25 ms), indicating that semantic interference in the blocked-cyclic naming task generalizes to novel category items not previously named (Belke et al., 2005b; see Figure 3B). We assessed whether generalization of semantic interference manifested as a linear increase across cycles of old and novel items (Oppenheim et al., 2010, Simulation 3) vs. emerging on the second cycle and remaining stable thereafter (Belke et al., 2005b, Experiment 3). While interference increased linearly when all eight cycles are included in the analyses [ $F_1(1,30) = 21.18, p < 0.001$ ;  $F_2(1,63) = 23.64, p < 0.001$ ], the linear contrast is not significant when the first cycle is excluded from the analyses ( $p$ 's > 0.11), suggesting that interference emerges after Cycle 1 and remains stable thereafter (see Belke et al., 2005b). We also conducted analyses including Cycles 2–5 following the prediction put forth in Belke et al. (2005b, p. 683): “If the semantic blocking effect generalizes to new items, the difference between homogeneous and heterogeneous sets that we expected to observe in cycles 2–4 should prevail on the 5th cycle.” Consistent with Belke et al. (2005b) we find a significant main effect of Condition

[ $F_1(1,30) = 12.15, p = 0.002$ ;  $F_2(1,63) = 11.43, p = 0.001$ ] and Cycle [ $F_1(3,90) = 222.35, p < 0.001$ ;  $F_2(3,189) = 199.10, p < 0.001$ ], but no interaction between the two variables ( $F$ 's < 1.73,  $p$ 's > 0.16). Consistent with generalization of interference as defined in Forde and Humphreys (1995, Experiment 12), we also find larger semantic interference on the first cycle of novel items (i.e., Cycle 5) vs. the first cycle of old items (i.e., Cycle 1) [Related – Unrelated 12 ms vs. –22 ms, respectively;  $F_1(1,30) = 16.01, p = 0.002$ ;  $F_2(1,63) = 12.74, p = 0.001$ ].

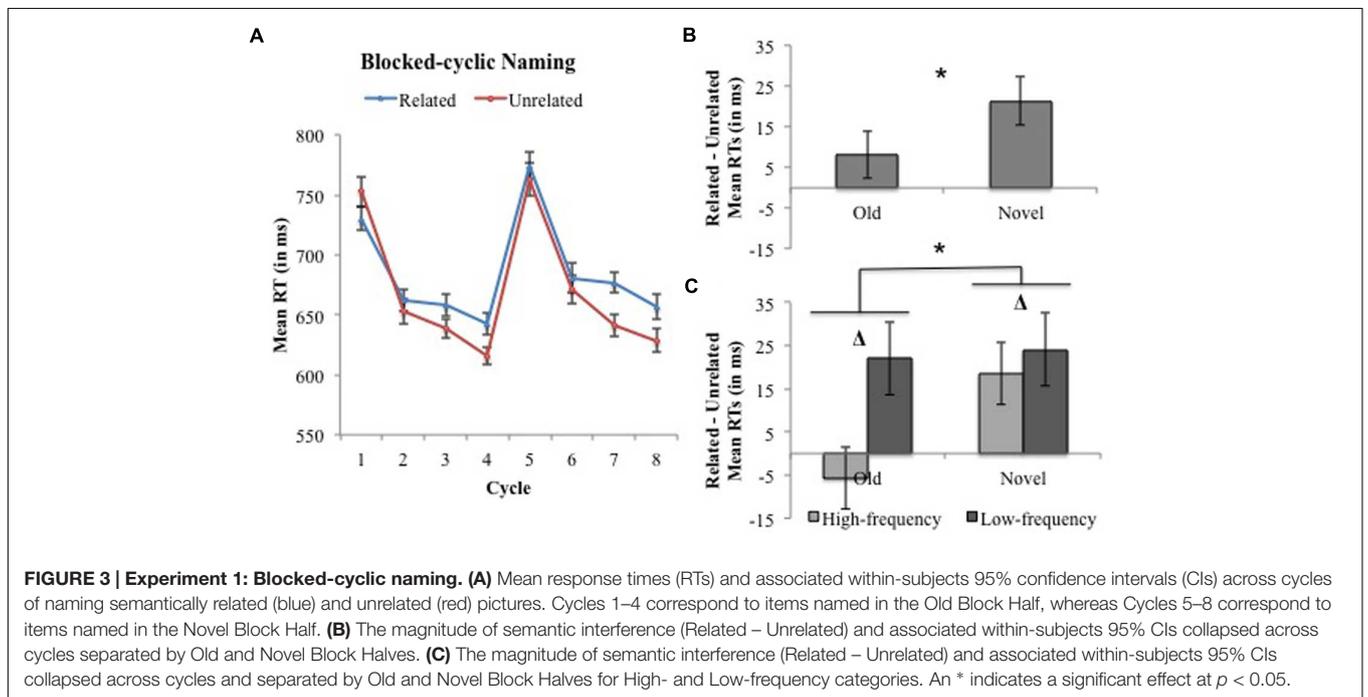
Analyses examining Frequency revealed marginally significant interactions between this variable and Cycle and Condition – but a significant three-way interaction between Frequency, Condition, and Block Half. The Frequency × Cycle marginal interaction was due to a reduction in the lexical frequency effect (High- < Low-frequency) with repetition (Low-frequency – High-frequency difference: 39, 12, 23, and 20 ms), which is consistent with previous studies (e.g., Scarborough et al., 1977; Jescheniak and Levelt, 1994; Griffin and Bock, 1998). The marginal interaction between Frequency and Condition indicated that low- compared to high-frequency items exhibited greater semantic interference (Related – Unrelated; 23 vs. 6 ms, respectively; Condition × Frequency effect 17 ms, 95% CI 1–32 ms). The Frequency × Cycle and Frequency × Condition interactions were significant by subject and marginally significant by item (Frequency × Cycle:  $p = 0.07$ ; Frequency × Condition:  $p = 0.09$ ), which may be because Frequency is a between-item variable in the  $F_2$  analyses, and the significant three-way interaction between Frequency, Condition, and Block Half. That is, semantic interference for low-frequency words exceeded that of high-frequency words in the Old Block Half [22 ms vs.

Ghyselinck, 2006, also see Barry et al., 2001), both are thought to reflect processes occurring when mapping meanings to words for naming (e.g., Belke et al., 2005a; Anderson, 2008).

TABLE 2 | Experiment 1 naming latencies.

		Cycles							
		Old				Novel			
		1	2	3	4	5	6	7	8
Related	High-frequency	706	652	637	625	755	666	668	646
	Low-frequency	754	672	679	660	794	696	686	668
	Mean	730	662	658	642	774	681	677	657
Unrelated	High-frequency	739	658	631	614	741	667	632	619
	Low-frequency	767	648	646	617	785	674	650	639
	Mean	753	653	639	616	763	671	641	629
Difference		-22	9	19	27	12	10	35	28

Mean RTs displayed in the shaded rows are collapsed across Frequency. Mean RTs (in ms) latencies separated by Condition, Block Half, Cycle, and Frequency.



–6 ms, respectively;  $F_1(1,30) = 7.92$ ,  $p < 0.001$ ;  $F_2(1,62) = 7.53$ ,  $p < 0.001$ ], but did not differ from high-frequency words in the Novel Block Half [24 ms vs. 19 ms, respectively;  $F_1(1,30) = 0.35$ ,  $p = 0.56$ ;  $F_2(1,62) = 0.23$ ,  $p = 0.63$ ; see Figure 3C).<sup>6</sup>

To summarize, the results from Experiment 1 are consistent with the assumption that semantic interference in naming originates at the semantic level but has its locus at the lexical level (e.g., Belke, 2013). That semantic interference increased when naming novel items categorically related to those

named previously (collapsed across cycles) demonstrates that interference originates at the semantic level as a result of shared activation. However, inconsistent with computational models of semantic interference in naming (e.g., Oppenheim et al., 2010; see also Howard et al., 2006) is the finding that the effect does *not* increase linearly across cycles of old and novel items. Instead, we replicate evidence elsewhere demonstrating that interference emerges on the second cycle and remains stable thereafter (Belke et al., 2005b), resulting in larger interference for the first cycle of novel vs. the first cycle of old items (Forde and Humphreys, 1995). This suggests that the larger interference effect observed for Novel relative to Old Block Halves (collapsed across cycles) reflects in part the presence of facilitation on the first cycle of old items vs. its absence on the first cycle of the novel items.

In addition, lexical frequency modulated semantic interference in naming, providing evidence of a lexical-level locus. However, the three-way interaction between lexical

<sup>6</sup>Interference for low- but not high-frequency words decreased numerically from the last cycle of old to the first cycle of novel items (low frequency Cycle 4 vs. Cycle 5 interference effect = 43 ms vs. 9 ms; high-frequency Cycle 4 vs. Cycle 5 interference effect = 11 ms vs. 14 ms; see Table 2), suggesting that lexical frequency differentially affects interference magnitudes across cycles, and thus interference generalization. However, *post hoc* analyses examining whether the interference effect across Cycles 1–8 and 2–5 differed for high- vs. low-frequency words were not significant ( $F$ 's  $< 1.30$ ,  $p$ 's  $> 0.25$  and  $F$ 's  $< 1.68$ ,  $p$ 's  $> 0.17$ , respectively).

frequency, semantic interference, and Block Half (Old vs. Novel) was surprising, as it might be expected that lexical frequency would have a consistent impact on interference across both old and novel items. We discuss this unexpected finding in the General Discussion, as it has implications for different mechanistic accounts of *how* semantic interference in naming occurs, i.e., due to competitive lexical selection (e.g., Roelofs, 1992; Levelt et al., 1999; Howard et al., 2006) vs. competitive learning (Oppenheim et al., 2010). Nonetheless, the results from Experiment 1 support a semantic origin and lexical locus of the interference effect in naming.

## Experiment 2: Semantic Interference in Word-Picture Matching

Although the locus of interference is thought to differ in the two tasks – lexical in naming (e.g., Howard et al., 2006; Oppenheim et al., 2010) and semantic in word-picture matching (e.g., Forde and Humphreys, 1997; Campanella and Shallice, 2011) – interference is assumed to originate at a shared semantic level. Thus, interference in word-picture matching is predicted to generalize to novel category items. Moreover, qualitative RT patterns suggest that semantic interference increases for low- relative to high-frequency words in word-picture matching (Campanella and Shallice, 2011), raising the possibility that interference in naming and word-picture matching have a shared origin *and* locus. We test this hypothesis in Experiment 2 using a word-picture matching variant of the blocked-cyclic naming task used in Experiment 1. In this task, subjects matched a visually presented word to its corresponding picture which appeared embedded in an array of three distractor pictures either semantically related or unrelated to the target picture (e.g., Biegler et al., 2008, Experiment 2B; Harvey and Schnur, 2015). If the locus of interference is shared across the two tasks, then semantic interference in word-picture matching will exhibit the same characteristics as those observed in picture naming (Experiment 1): increased semantic interference for novel compared to old categorically related words (i.e., generalization of semantic interference) and greater interference for low- compared to high-frequency words for old vs. novel pictures (i.e., an interaction between Condition, Frequency, and Block Half). Alternatively, if semantic interference in word-picture matching has its origin *and* locus within the semantic system (e.g., Forde and Humphreys, 1997, 2007; Campanella and Shallice, 2011), then it is predicted to generalize to novel category members, but not interact with lexical frequency.

### Procedure

Subjects first completed the learning phase identical to that of Experiment 1. Immediately thereafter, subjects completed a practice phase, which followed the same parameters as the actual experiment but used items not depicted in the experiment (i.e., Old: BEE, ORANGE, SCISSORS, and DOLFIN; Novel: PLANT, GRAPES, PEN, and SHARK). This was done to demonstrate the speed with which the target words appear in the blocked-cyclic word-picture matching task. The procedure was as follows. A visually presented word appeared in the center of the screen for 300 ms followed by an array of four pictures: one corresponding

to the previous target word and three distractor pictures. Subjects were instructed to select the picture that matches the previously presented word by tapping the picture on the touch screen monitor. Distractor pictures depicted words appearing as other targets in the cycle, and therefore either belonged to the same semantic category as the target (Related Condition) or belonged to different semantic categories as the target (Unrelated Condition). See **Figure 2**. All other experiment parameters were identical to Experiment 1.

### Statistical Analyses

Response times analyses did not include trials classified as an error (i.e., selecting an incorrect picture) or responses faster than 250 ms or slower than 1550 ms. Valid RTs were analyzed using the same repeated measures ANOVAs as those used in Experiment 1: Random factors included participants and items, yielding  $F_1$  and  $F_2$  statistics, respectively. Fixed factors included Condition (Related and Unrelated), Block Half (Old and Novel), Cycles (1–4), and Frequency (High-frequency and Low-frequency). All fixed factors were considered within-subject, within-item variables except for Frequency, which was a within-subject variable in the  $F_1$  analysis and a between-item variable in the  $F_2$  analysis.

## Results and Discussion

Response errors occurred on 1.1% of experimental trials. **Tables 3** and **4** summarize RT  $F$  statistics and mean RTs, respectively. See **Figure 3A** for the full pattern of results.

There was a main effect of Condition, due to slower RTs in the Related (676 ms) compared to the Unrelated Condition (598 ms; Condition effect 79 ms, 95% CI 63–94 ms). However, in contrast to the naming results obtained in Experiment 1, main effects of Block Half, Frequency, and Cycle were not significant. Lastly, two- and three-way interactions between Condition, Block Half, and Frequency were not significant (see **Figure 4**). As in Experiment 1, we examined interference generalization in word-picture matching following Belke et al. (2005b; i.e., stable interference across Cycles 2–5) and Forde and Humphreys (1995; i.e., larger Cycle 5 vs. Cycle 1 interference). We found that although interference remained stable across Cycles 2–5 (i.e., main effect of Condition [ $F_1(1,19) = 74.62$ ,  $p < 0.001$ ;  $F_2(1,63) = 33.17$ ,  $p < 0.001$ ], but no interaction with Cycle ( $F$ 's  $< 1.19$ ,  $p$ 's  $> 0.31$ )), interference did *not* differ on the first cycle of old vs. the first cycle of novel items (Related – Unrelated 86 ms vs. 83 ms, respectively;  $F$ 's  $< 0.07$ ,  $p$ 's  $> 0.80$ ). Together, these findings demonstrate that although semantic interference occurs in healthy participants' word-picture matching performance, it does *not* manifest in the same manner as that which occurs in picture naming (i.e., Experiment 1).

The findings from Experiment 2 are only partially consistent with the assumption that the origin and locus of semantic interference in word-picture matching is at the semantic level (e.g., Warrington and Cipolotti, 1996; Forde and Humphreys, 2007) for two reasons. First, interference did not increase for novel relative to old category items (i.e., no generalization of interference) – a result at odds with a previous

**TABLE 3 | Experiment 2 ANOVA results.**

	Subject			Item		
	Degrees of freedom			Degrees of freedom		
	Numerator	Denominator	$F_1$	Numerator	Denominator	$F_2$
Condition	1	19	97.42*	1	62	29.71*
Block Half	1	19	1.35	1	62	0.10
Cycle	3	57	7.81*	3	186	1.41
Frequency	1	19	0.05	1	62	0.12
Condition × Block Half	1	19	1.99	1	62	0.03
Condition × Cycle	3	57	1.24	3	186	0.64
Condition × Frequency	1	19	1.91	1	62	4.42*
Block Half × Cycle	3	57	1.12	3	186	1.14
Block Half × Frequency	1	19	0.43	1	62	0.38
Cycle × Frequency	3	57	2.21	3	186	0.95
Condition × Block Half × Cycle	3	57	0.38	3	186	1.57
Condition × Block Half × Frequency	1	19	0.05	1	62	0.06
Condition × Cycle × Frequency	3	57	0.75	3	186	1.28
Block Half × Cycle × Frequency	3	57	2.65	3	186	1.33
Condition × Block Half × Cycle × Frequency	3	57	1.20	3	186	0.72

Significant main effects and interactions in both the  $F_1$  and  $F_2$  analyses appear shaded in gray, and an \* indicates a significant effect at  $p < 0.05$ . Summary of  $F$  statistics for the RT ANOVA examining the effects of Condition (Related and Unrelated), Block Half (Old and Novel), Cycles (1–4), and Frequency (High-frequency and Low-frequency).

**TABLE 4 | Experiment 2 word-picture matching response latencies.**

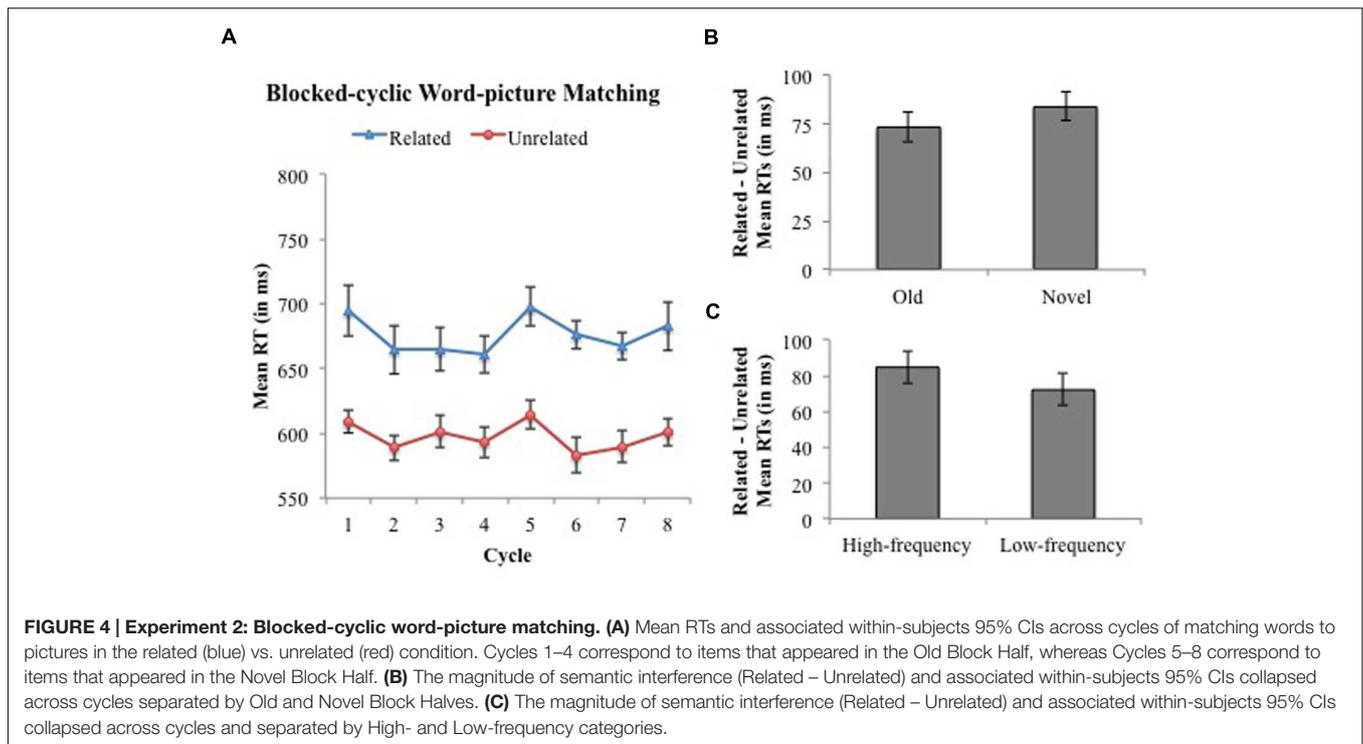
		Cycles							
		Old				Novel			
		1	2	3	4	5	6	7	8
Related	High-frequency	702	671	661	663	701	658	669	703
	Low-frequency	687	658	668	659	694	694	665	663
	Mean	695	664	665	661	698	676	667	683
Unrelated	High-frequency	618	587	585	592	608	577	584	597
	Low-frequency	599	590	617	593	621	589	595	605
	Mean	609	589	601	593	614	583	590	601
Difference	86	75	63	68	83	93	77	82	

Mean RTs displayed in the shaded rows are collapsed across Frequency. Mean RTs (in ms) separated by Condition, Block Half, Cycle, and Frequency.

neuropsychological finding suggesting a semantic-level locus (Forde and Humphreys, 1995, Experiment 12). It is possible that the visual similarity between target and distractor pictures in related vs. unrelated picture arrays contaminates the interference effect, masking potential changes in interference across cycles. Belke et al. (2005b) suggested that in naming, the absence of interference on the first cycle argues against a visual (similarity) locus of the effect, as interference due to visual similarity should be largest at the first presentation of items (i.e., Cycle 1) and reduced on subsequent presentations. We had participants from Experiment 2 rate the visual similarity of target pictures appearing together in an array after completing the word-picture matching task to assess the contribution of visual similarity on word-picture matching RTs. Two pictures appeared side-by-side, and participants rated the visual similarity of the two pictures on a 5-point scale (1 = not at all visually similar, 5 = very highly

visually similar). Items appearing in the related vs. unrelated picture arrays were rated as more visually similar [ $t_1(19) = 9.46$ ,  $p < 0.001$ ;  $t_2(63) = 14.41$ ,  $p < 0.001$ ]. An analysis of covariance revealed that the condition effect remained significant after controlling for visual similarity in the analysis by subject [ $F_1(1,18) = 15.35$ ,  $p = 0.001$ ], but not by item [ $F_2(1,62) = 2.20$ ,  $p = 0.14$ ], suggesting that for some of the picture arrays visual similarity contributed to the interference effect.

Second, consistent with a non-lexical locus of interference (e.g., Forde and Humphreys, 1997, 2007), interference was *not* sensitive to lexical frequency, a hypothesized lexical, not semantic effect. This is in contrast to Campanella and Shallice (2011) who found that lexical frequency had some effect on semantic interference in word-picture matching. While it is not entirely clear why we did not replicate Campanella and Shallice, we hypothesize it may be due to a difference between experiment



designs. Campanella and Shallice (Experiment 1) included a baseline (unrelated) condition but the items in this baseline condition differed from those tested in the experimental (related) condition (and cycle was not manipulated as a factor). In contrast, in the present study, items served as their own controls by appearing in both the related and unrelated conditions (allowing us to draw stronger conclusions concerning the effects of relatedness across items; see also Biegler et al., 2008; Harvey and Schnur, 2015; Wei and Schnur, 2016). Thus, it may be the case that their findings were driven by item-specific differences between related and unrelated conditions.

### Experiment 3: Transfer of Semantic Interference

Experiments 1 and 2 demonstrated that the locus of interference in naming and word-picture matching tasks differs, but leave open the possibility that semantic interference originates at a shared processing level. Experiment 3 tested this hypothesis by investigating whether accessing the semantic system when performing the word-picture matching task with semantically related vs. unrelated words subsequently impacts naming novel pictures drawn from the same vs. different semantic categories.<sup>7</sup> If the origin of semantic interference in the two tasks is *not* shared, then semantic interference when naming before (i.e., Old) and after word-picture matching (i.e., Novel) will be of comparable

<sup>7</sup>Because we found that semantic interference remained stable across old and novel items in word-picture matching, it would not be expected to increase for novel items semantically related to those tested in a previous block or task. Thus, the failure to demonstrate interference generalization in word-picture matching precludes inferences concerning interference transfer from naming to word-picture matching.

magnitudes, i.e., unaffected by interference in word-picture matching. However, if interference in word-picture matching and naming originate at a shared level, then semantic interference should transfer across tasks, whereby naming novel pictures semantically related to those accessed previously in word-picture matching should result in greater semantic interference than when naming precedes word-picture matching (i.e., novel > old same-category naming).

### Procedure

The procedure followed that of the previous experiments except in Experiment 3 subjects either named pictures or performed word-picture matching in the first Block Half (i.e., Old) and switched tasks for the second Block Half (i.e., Novel) within both related and unrelated blocks of trials (see Figure 2). For a given subject, half of the blocks began with word-picture matching (i.e., Cycles 1–4) followed by naming (i.e., Cycles 5–8), whereas the other half began with naming (i.e., Cycles 1–4) followed by word-picture matching (i.e., Cycles 5–8). While we did not expect to find changes when switching from naming to word-picture matching (see footnote 6), we included this manipulation to determine if the semantic interference magnitude for naming increased in the Novel Block Half (i.e., after word-picture matching) relative to naming in the Old Block Half (i.e., before word-picture matching), where a larger magnitude when switching to naming would indicate interference transferred from word-picture matching to picture naming. Subjects always switched tasks halfway through the block (i.e., when presented with cycles of novel stimuli) an equal number of times throughout the experiment. The order of the task switch (from naming to word-picture matching and vice

versa) was constrained so that no more than three consecutive blocks occurred with the same task switching direction. Further, there were an equal number of Related and Unrelated as well as High- and Low-frequency blocks for each task switching direction. We created an additional 10 lists of test materials by counterbalancing across subjects items appearing in each Block Half for a given task (i.e., blocked-cyclic naming or word-picture matching), resulting in a total of 20 lists.

As in Experiment 2, participants completed a practice phase with the same stimuli used in Experiment 2 to familiarize them with not only the fast presentation rate, but also the task-switching procedure. Because the task-switching was somewhat unpredictable, subjects were given a cue (i.e., #####) when the task switched from word-picture matching to naming, and the visually presented target word served as a cue when switching from naming to word-picture matching. Thus, participants always practiced the word-picture matching task before the picture naming task in order to familiarize them with the switch cue. When a single picture appeared on the screen, subjects were instructed to name the picture into the microphone headset as quickly and accurately as possible. When a word appeared followed by an array of pictures, subjects were instructed to select the picture corresponding to the previously presented word by tapping the picture on the touch screen monitor. All other experiment parameters were identical to Experiments 1 and 2.

### Statistical Analyses

Response times for trials were excluded in the same manner as Experiments 1 and 2. Valid RTs were analyzed using repeated measures ANOVAs with participants and items as random factors, yielding  $F_1$  and  $F_2$  statistics, respectively. Fixed factors included: Task (Blocked-cyclic naming and Blocked-cyclic word-picture matching), Condition (Related and Unrelated), Block Half (Old vs. Novel), and Cycles (1–4). All fixed factors were considered within-subject, within-item variables.

### Results and Discussion

Response errors occurred on 3.9% of experimental trials. **Tables 5 and 6** summarize RT  $F$  statistics and mean RTs, respectively.

There were main effects of Task, Condition, and Cycle. The effect of Task revealed that RTs were faster for blocked-cyclic naming (666 ms) compared to word-picture matching (707 ms; Task effect 41 ms, 95% CI 15–67 ms). As expected, the main effect of Condition was due to slower RTs in the Related (713 ms) vs. Unrelated Condition (660 ms; Condition effect 53 ms, 95% CI 47–60 ms). The Cycle effect was due to decreasing response latencies from the first (751 ms) to the remaining three cycles (669, 662, and 665 ms; see **Figures 5A,B**).

The interaction between Task and Cycle revealed that RTs decreased with repetition to a greater extent in blocked-cyclic naming (749, 645, 638, and 629 ms) than word-picture matching (753, 691, 684, and 700 ms), and Block Half also interacted with Cycle, but not in a meaningful way. The interaction between Task and Condition was due to smaller semantic interference (Related – Unrelated) in naming (9 ms) than word-picture matching (96 ms; Task  $\times$  Condition difference 87 ms, 95% CI 73–101 ms). The significant three-way interaction between

Task, Condition, and Block Half was due to larger semantic interference when naming in the Novel Block Half (i.e., after word-picture matching) compared to naming in the Old Block Half (20 ms vs. –2 ms, respectively). The increase in naming-induced semantic interference for Novel vs. Old Block Halves was confirmed with a simple effects comparison [Condition  $\times$  Block Half effect 22 ms, 95% CI 7–35 ms;  $t_1(42) = -3.01$ ,  $p < 0.01$ ;  $t_2(63) = -2.69$ ,  $p < 0.01$ ]. That interference in naming was greater following word-picture matching suggests that semantic interference transferred from word-picture matching to picture naming (see **Figure 5C**).

We were surprised, however, to find that interference in the Old Block Half of naming was numerically smaller in Experiment 3 (–2 ms) compared to Experiment 1 (8 ms). We assessed *post hoc* whether interference in naming across the two experiments was similar in magnitude (thus similar in terms of “generalization”) with repeated measures ANOVAs that included Experiment (Experiments 1 and 3) as a between-subjects, within-item variable. The results mirrored the main findings of interference in naming reported above (i.e., main effects of Condition, Related > Unrelated), Cycle (RTs decreased across cycles), and interactions between Condition and Block Half (Novel > Old) and Condition and Cycle (semantic interference increased across cycles;  $F$ 's > 10.35,  $p$ 's < 0.002). Critically, the factor Experiment did not modulate any interactions with Condition ( $F$ 's < 1.45,  $p$ 's > 0.23). The increase in semantic interference from old to novel same-category naming was the same regardless of the task performed in the Old Block Half (i.e., 21 ms after naming in Experiment 1 and 20 ms after word-picture matching in Experiment 3). Likewise, analyses of interference across Cycles 2–5 (i.e., Belke et al., 2005b) and comparisons between interference on Cycle 1 vs. 5 (Forde and Humphreys, 1995) were consistent with the findings of Experiment 1 [i.e., main effects of Condition and Cycle ( $F$ 's > 9.75,  $p$ 's < 0.005), but no Condition  $\times$  Cycle interaction ( $F$ 's < 2.17,  $p$ 's > 0.09), and greater Cycle 5 vs. Cycle 1 interference (Related – Unrelated = 12 ms vs. –21 ms, respectively;  $F$ 's > 12.95,  $p$ 's < 0.001)], where here too Experiment did not modulate interactions with Condition ( $F$ 's < 1.90,  $p$ 's > 0.17). This suggests that interference in naming and word-picture matching originate at a shared (semantic) level of the language system (see Belke, 2013 for a similar rationale).

### GENERAL DISCUSSION

To bridge the gap between theories of lexical-semantic access in naming vs. word-picture matching tasks (e.g., Forde and Humphreys, 1997, 2007; Levelt et al., 1999), we examined whether the origin and/or locus of semantic interference is shared across the two tasks. Accordingly, we tested the extent to which interference generalized to novel category items and interacted with lexical frequency in picture naming and word-picture matching variants of the blocked-cyclic paradigm, while also assessing whether interference transferred across the two tasks. In line with a semantically mediated lexical locus of interference in naming (cf. Belke, 2013), Experiment 1 demonstrated that semantic interference increased when naming

**TABLE 5 | Experiment 3 ANOVA results.**

	Subject			Item		
	Degrees of freedom			Degrees of freedom		
	Numerator	Denominator	$F_1$	Numerator	Denominator	$F_2$
Task	1	42	9.83*	1	63	37.77*
Condition	1	42	249.21*	1	63	73.30*
Block Half	1	42	0.01	1	63	0.09
Cycle	3	126	155.33*	3	189	197.44*
Task × Condition	1	42	152.87*	1	63	80.41*
Task × Block Half	1	42	1.31	1	63	2.28
Task × Cycle	3	126	26.88*	3	189	23.31*
Condition × Block Half	1	42	0.00	1	63	0.09
Condition × Cycle	3	126	1.49	3	189	1.47
Block Half × Cycle	3	126	21.25*	3	189	23.64*
Task × Condition × Block Half	1	42	9.01*	1	63	10.76*
Task × Condition × Cycle	3	126	2.24	3	189	1.77
Task × Block Half × Cycle	3	126	2.95*	3	189	1.88
Condition × Block Half × Cycle	3	126	1.89	3	189	2.85*
Task × Condition × Block Half × Cycle	3	126	0.87	3	189	1.23

Significant main effects and interactions in both the  $F_1$  and  $F_2$  analyses appear shaded in gray, and an \* indicates a significant effect at  $p < 0.05$ . Summary of  $F$  statistics for the RT ANOVA examining the effects of Task (Blocked-cyclic naming and Blocked-cyclic word-picture matching), Condition (Related and Unrelated), Block Half (Old and Novel), and Cycles (1–4).

**TABLE 6 | Experiment 3 response latencies.**

		Cycles							
		Old				Novel			
		1	2	3	4	5	6	7	8
Blocked-cyclic naming	Related	731	647	652	641	763	648	638	640
	Unrelated	751	655	644	627	752	631	618	608
	Difference	–20	–8	8	14	11	17	20	32
	Mean	741	651	648	634	758	640	628	624
Blocked-cyclic Word-picture matching	Related	784	750	743	751	822	728	712	750
	Unrelated	673	651	632	647	733	636	652	651
	Difference	111	99	111	104	90	93	60	100
	Mean	729	701	687	699	778	682	682	700

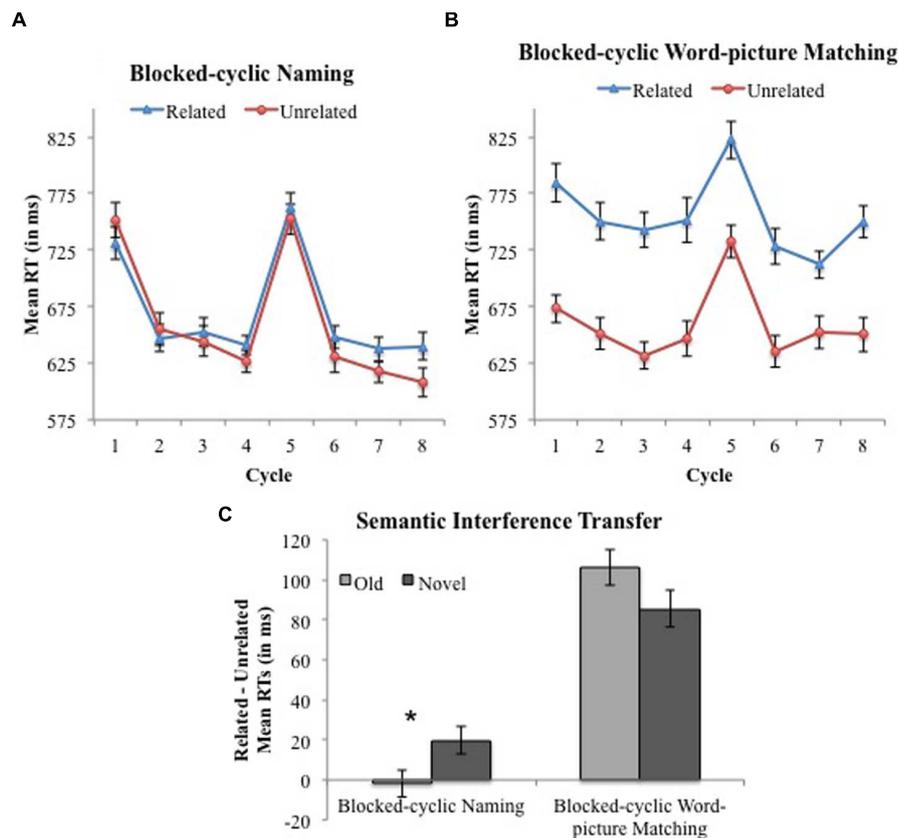
Mean RTs displayed in the shaded rows are collapsed across Condition. Mean RTs reflect data from all participants, as each participant performed the naming and word-picture matching tasks in both the Old and Novel Block Halves (see Procedure in Experiment 3). Mean RTs (in ms) separated by Task, Condition, Block Half, and Cycle.

novel pictures drawn from the same category as those named previously where the effect differed based on the lexical frequency of target items. Experiment 2 demonstrated that although interference occurs in word-picture matching, it did not change in magnitude for old vs. novel items or for high- vs. low-frequency words. Lastly, Experiment 3 revealed that semantic interference increased when naming novel items categorically related to those accessed previously in word-picture matching (as compared to when naming preceded word-picture matching). Together, these experiments suggest that the locus of semantic interference in picture naming and word-picture matching differs (lexical vs. semantic), but that both interference effects originate

at a shared (semantic) level. In the following, we discuss how these findings inform existing theories of lexical-semantic access and semantic interference phenomena in the blocked-cyclic paradigms.

### Semantic Interference in Naming

That semantic interference in naming generalized to novel category pictures and differed based on lexical frequency is consistent with a semantically mediated lexical locus of the effect. However, in order to account for the full pattern of results, additional assumptions must be adopted. For example, Oppenheim et al. (2010) predicts linearly increasing semantic



**FIGURE 5 | Experiment 3: Blocked-cyclic naming and word-picture matching.** Mean RTs and associated within-subjects 95% CIs across cycles of naming pictures and matching words to pictures in the semantically related (blue) and unrelated (red) conditions. **(A)** Cycles 1–4 correspond to items named in the Old Block Half prior to performing the word-picture matching task, whereas Cycles 5–8 correspond to items named in the Novel Block Half after having performed the word-picture matching task with semantically related words in the Old Block Half. **(B)** Cycles 1–4 correspond to items that appeared in the word-picture matching task in the Old Block Half prior to naming pictures, whereas Cycles 5–8 correspond to items that appeared in the word-picture matching task in the Novel Block Half after having named semantically related pictures in the Old Block Half. **(C)** The magnitude of semantic interference (Related – Unrelated) and associated within-subjects 95% CIs collapsed across cycles and separated by Old and Novel Block Halves for blocked-cyclic naming and word-picture matching. Here, “Old” bars illustrate the semantic interference effect observed before switching tasks, whereas “Novel” refers to the semantic interference effect observed after switching tasks. An \* indicates a significant effect at  $p < 0.05$ .

interference across cycles of old and novel items (Simulation 3). However, this prediction was not confirmed, as semantic interference remains stable from when it emerged on the second cycle to the introduction of novel category items (seen here in Experiment 1 and in Belke et al., 2005b, Experiment 3). Although there may be increasing semantic interference across cycles in the blocked-cyclic naming task, according to Belke (2008) and Belke and Stielow (2013), this task promotes the use of top-down cognitive control processes which masks the accumulation of semantic interference across cycles by biasing activation toward within-set category representations and away from set-external category members (i.e., biased selection account; see also Thompson-Schill and Botvinick, 2006; Belke, 2013; Crowther and Martin, 2014). Thus, top-down control may reduce, if not eliminate, semantic interference on the first cycle of novel items (i.e., Cycle 5), providing an explanation for why the Oppenheim et al. (2010) Simulation 3 does not fully capture the lack of interference change across cycles.

Although we find that interference in naming interacted with lexical frequency – consistent with a lexical-level locus (Roelofs, 1992; Levelt et al., 1999; Howard et al., 2006; Oppenheim et al., 2010), the finding that lexical frequency differentially impacted interference for old vs. novel items was not expected. Models that assume interference occurs due to competitive lexical selection (e.g., Roelofs, 1992; Levelt et al., 1999; Howard et al., 2006) predict greater interference for both old and novel low-frequency words due to their inherently lower activation levels, which makes them more susceptible to competition from related high-frequency words. However, a recent account proposes that interference arises due to a learning mechanism that strengthens target lexical-semantic connections while weakening those of related representations (i.e., competitive learning; Oppenheim et al., 2010). Lexical-semantic connections change in magnitude (strengthen or weaken) in proportion to their activation levels, or error in becoming active on a given trial (i.e., delta rule learning; e.g.,

Chang et al., 2000; Gupta and Cohen, 2002). Naming the same low- vs. high-frequency words (i.e., old items) results in greater semantic interference because low-frequency words have a greater learning potential due to their inherently lower activation levels (e.g., Oldfield and Wingfield, 1965; Morton, 1969). However, novel high-frequency words will be more active than novel low-frequency words due to their inherently higher activation levels, and thus greater “unlearning” potential. In turn, novel high- compared to low-frequency related words should undergo greater lexical-semantic connection weight weakening, rendering them functionally similar to low-frequency words. Thus, the competitive learning account provides a potential explanation as to why the interaction between frequency and semantic interference differed for old vs. novel items. Although this account can only be verified by computational modeling, on the face of it, the competitive learning account explains both the finding of increased interference for low-frequency words and the interaction between this characteristic and naming old vs. novel items. However, this extension of Oppenheim et al.’s (2010) account would also predict repetition (across cycles) modulates the observed interactions between lexical frequency and semantic interference – a prediction that was not borne out by the results. That both generalization of semantic interference and lexical frequency were not sensitive to repetition (i.e., cycle) suggests that there may be other factors at play in the blocked-cyclic naming task. Future work is needed to clarify the different mechanisms underlying semantic interference when repeatedly naming the same and novel categorically related high- vs. low-frequency words in the blocked-cyclic naming task.

## Semantic Interference in Word-Picture Matching

Semantic interference in the word-picture matching task differed from that observed in picture naming, suggesting that the locus of interference differs from naming. Although we observed semantic interference in word-picture matching (replicating previous results, e.g., Biegler et al., 2008; Campanella and Shallice, 2011), interference did not increase from old to novel categorically related words, which is inconsistent with a semantic-level origin and/or locus of the effect. However, there are several other explanations to consider. First, as discussed in Experiment 2, because interference did not change across cycles but was of the same magnitude from the first cycle onward, this makes it impossible to detect any generalization (or change) from old to novel items. Other variants of the word-picture matching task may be more sensitive to change of interference and thus may provide better tools to investigate generalization (cf. Wei and Schnur, 2016). Our findings suggest that the high degree of visual similarity among related vs. unrelated picture arrays had some effect on the magnitude of interference in word-picture matching (see Results and Discussion in Experiment 2). However, consistent with a semantic-, and not lexical-level locus of interference (e.g., Warrington and Cipolotti, 1996; Gotts and Plaut, 2002; Forde and Humphreys, 2007), lexical frequency did not affect

semantic interference magnitudes in word-picture matching. On the assumption that lexical frequency reflects word-level processing, this factor should *only* affect the time it takes to recognize the target word (e.g., Scarborough et al., 1977), and should *not* affect the time it takes to select the word’s depicted referent. This is because the word is presented first for a fixed amount of time, and differences in RTs for each condition reflect differences in the time it takes to select the picture rather than recognizing the word itself. Nonetheless, the failure to detect lexical frequency semantic interference effects in the word-picture matching task is consistent with a semantic-level locus of interference, but future work is needed to determine if a semantic locus of interference allows for generalization to novel word meanings using a task that minimizes potential confounds such as visual similarity among distractors pictures in the related vs. unrelated arrays.

## Transfer of Semantic Interference across Tasks

Lastly, the transfer of semantic interference from word-picture matching to naming suggests overlap in where interference originates in the two tasks. Because the naming results suggest a semantic origin and lexical locus of interference, whereas the word-picture matching results suggest a non-lexical locus of interference, together this suggests that the common origin of interference across the tasks is a semantic one. However, what is the evidence to rule out a lexical locus? First, lexical frequency did not interact with semantic interference in word-picture matching suggesting a non-lexical locus (Experiment 2). Second, although participants may have tacitly named the targets and pictures in the array (cf. Biegler et al., 2008), allowing semantic interference to seemingly transfer across tasks, there is evidence which argues against this possibility. For example, it is unlikely that subjects covertly named the four pictures in the array before responding because in Experiment 3, subjects named one picture on average within 667 ms, whereas average RTs for the word-picture matching task were 708 ms. Additionally, if subjects named the targets during the word-picture matching task, then semantic interference should have manifested as it did for naming (i.e., interacted with block half and lexical frequency). Previous work also demonstrates that semantic interference occurs in word-picture matching tasks that do *not* promote a silent naming strategy (by requiring subjects to select the picture most associated with the word), regardless of whether the distractor pictures in the array are related (Biegler et al., 2008, Experiment 3) or unrelated to the target picture (Wei and Schnur, 2016). That semantic interference occurs when matching associatively related words and pictures but does *not* occur when naming associatively related pictures (unless participants are primed with the scene/event name characterizing the associative relationship; Abdel Rahman and Melinger, 2011; cf. de Zubicaray et al., 2014), suggests that interference in word-picture matching is not due to silent naming. Third, models of semantic interference in naming (Howard et al., 2006; Oppenheim et al., 2010) assume that interference occurs *only* after having previously

named from the category, and therefore do *not* predict that accessing the semantic system in word-picture matching leads to interference when subsequently naming novel category pictures. Thus, the transfer of semantic interference from word-picture matching to naming indicates a shared semantic-level origin of the effect. However, because interference did not change in magnitude when word-picture matching was tested alone (Experiment 2), we were unable to examine interference transfer from naming to word-picture matching. Consequently, future work is needed to better understand the mechanisms by which semantic interference arises in word-picture matching and those that allow for interference to transfer across tasks.

## CONCLUSION

We examined whether the origin and/or locus of semantic interference in picture naming and word-picture matching is shared. We found that interference in naming generalized to novel category members and interacted with lexical frequency, a pattern that supports a semantic origin and lexical locus of interference in naming. In word-picture matching, however, the evidence for a semantic-level origin and locus of interference was mixed. We observed semantic interference which did not interact with lexical frequency, suggesting a non-lexical locus of the effect. Yet interference did not generalize to novel category members, arguing against a semantic locus of the effect. Because semantic interference in word-picture matching transferred to picture naming, this suggests a common origin of interference in the two tasks. We propose the shared origin is at the semantic

level, whereby accessing semantic representations contributes to interference effects in both tasks. However, future research is needed to provide more conclusive evidence regarding the shared origin of semantic interference and the mechanism by which interference transfers across word-picture matching and naming tasks.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## ACKNOWLEDGMENTS

These experiments were part of the doctoral dissertation thesis of DH. A subset of these results was presented at the International Workshop on Language Production (2014; Geneva, Switzerland). We thank Eman Bahrani and Kelsey Tomlinson for their help with data collection. This research was supported by a fellowship through the William Orr Dingwall Foundation and the Rice University School of Social Sciences Dissertation Research Improvement Grant awarded to DH.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.00710>

## REFERENCES

- Abdel Rahman, R. A., and Melinger, A. (2011). The dynamic microstructure of speech production: semantic interference built on the fly. *J. Exp. Psychol.* 37, 149–161. doi: 10.1037/a0021208
- Anderson, J. D. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *J. Fluency Disord.* 33, 135–155. doi: 10.1016/j.jfludis.2008.04.001
- Bajo, M. T. (1988). Semantic facilitation with pictures and words. *J. Exp. Psychol.* 14, 579–589.
- Balota, D. A., and Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: lexical access and/or production? *J. Mem. Lang.* 24, 89–106. doi: 10.1016/0749-596X(85)90017-8
- Barry, C., Hirsh, K. W., Johnston, R. A., and Williams, C. L. (2001). Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *J. Mem. Lang.* 44, 350–375. doi: 10.1006/jmla.2000.2743
- Bartram, D. J. (1976). Levels of coding in picture-picture comparison tasks. *Mem. Cogn.* 4, 593–602. doi: 10.3758/BF03213223
- Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychon. Bull. Rev.* 15, 357–363. doi: 10.3758/PBR.15.2.357
- Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: implications for models of lexical-semantic encoding. *J. Mem. Lang.* 69, 228–256. doi: 10.1016/j.jml.2013.05.008
- Belke, E., Brysbaert, M., Meyer, A. S., and Ghyselinck, M. (2005a). Age of acquisition effects in picture naming: evidence for a lexical-semantic competition hypothesis. *Cognition* 96, B45–B54. doi: 10.1016/j.cognition.2004.11.006
- Belke, E., Meyer, A. S., and Damian, M. F. (2005b). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Q. J. Exp. Psychol.* 58, 667–692. doi: 10.1080/02724980443000142
- Belke, E., and Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: evidence from blocked and continuous manipulations of semantic context. *Q. J. Exp. Psychol.* 66, 2135–2160. doi: 10.1080/17470218.2013.775318
- Biegler, K. A., Crowther, J. E., and Martin, R. C. (2008). Consequences of an inhibition deficit for word production and comprehension: evidence from the semantic blocking paradigm. *Cogn. Neuropsychol.* 25, 493–527. doi: 10.1080/02643290701862316
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., and Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE* 5:e10773. doi: 10.1371/journal.pone.0010773
- Brysbaert, M., and Ghyselinck, M. (2006). The effect of age of acquisition: partly frequency related, partly frequency independent. *Vis. Cogn.* 13, 992–1011. doi: 10.1080/13506280544000165
- Campanella, F., and Shallice, T. (2011). Refractoriness and the healthy brain: a behavioural study on semantic access. *Cognition* 118, 417–431. doi: 10.1016/j.cognition.2010.08.005
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208. doi: 10.1080/026432997381664
- Chang, F., Dell, G. S., Bock, K., and Griffin, Z. M. (2000). Structural priming as implicit learning: a comparison of models of sentence production. *J. Psycholinguist. Res.* 29, 217–230. doi: 10.1023/A:1005101313330
- Chen, Q., and Mirman, D. (2012). Competition and cooperation among similar representations: toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychol. Rev.* 119, 417–430. doi: 10.1037/a0027175

- Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407
- Crowther, J. E., and Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: the role of cognitive control and learning. *Front. Hum. Neurosci.* 8:9. doi: 10.3389/fnhum.2014.00009
- Crutch, S. J., and Warrington, E. K. (2005). Gradients of semantic relatedness and their contrasting explanations in refractory access and storage semantic impairments. *Cogn. Neuropsychol.* 22, 851–876. doi: 10.1080/02643290442000374
- Damian, M. F., and Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *J. Exp. Psychol.* 31, 1372–1384.
- Damian, M. F., Vigliocco, G., and Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition* 81, B77–B86. doi: 10.1016/S0010-0277(01)00135-4
- de Zubicaray, G., Johnson, K., Howard, D., and McMahon, K. (2014). A perfusion fMRI investigation of thematic and categorical context effects in the spoken production of object names. *Cortex* 54, 135–149. doi: 10.1016/j.cortex.2014.01.018
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283
- Finocchiaro, C., and Caramazza, A. (2006). The production of pronominal clitics: implications for theories of lexical access. *Lang. Cogn. Process.* 21, 141–180. doi: 10.1080/01690960400001887
- Forde, E., and Humphreys, G. W. (1995). Refractory semantics in global aphasia: on semantic organisation and the access–storage distinction in neuropsychology. *Memory* 3, 265–307. doi: 10.1080/09658219508253154
- Forde, E. M., and Humphreys, G. W. (1997). A semantic locus for refractory behaviour: implications for access storage distinctions and the nature of semantic memory. *Cogn. Neuropsychol.* 14, 367–402. doi: 10.1080/026432997381529
- Forde, E. M., and Humphreys, G. W. (2007). Contrasting effects of repetition across tasks: implications for understanding the nature of refractory behavior and models of semantic memory. *Cogn. Affect. Behav. Neurosci.* 7, 198–211. doi: 10.3758/CABN.7.3.198
- Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods* 35, 116–124. doi: 10.3758/BF03195503
- Gotts, S. J., and Plaut, D. C. (2002). The impact of synaptic depression following brain damage: a connectionist account of “access/refractory” and “degraded-store” semantic impairments. *Cogn. Affect. Behav. Neurosci.* 2, 187–213. doi: 10.3758/CABN.2.3.187
- Griffin, Z. M., and Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *J. Mem. Lang.* 38, 313–338. doi: 10.1006/jmla.1997.2547
- Gupta, P., and Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychol. Rev.* 109, 401–448. doi: 10.1037/0033-295X.109.2.401
- Harvey, D. Y., and Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex* 67, 37–58. doi: 10.1016/j.cortex.2015.03.004
- Howard, D. (1995). Lexical anomia: or the case of the missing lexical entries. *Q. J. Exp. Psychol.* 48, 999–1023. doi: 10.1080/14640749508401426
- Howard, D., Nickels, L., Coltheart, M., and Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition* 100, 464–482. doi: 10.1016/j.cognition.2005.02.006
- Humphreys, G. W., Lamote, C., and 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, 535–586. doi: 10.1080/09658219508253164
- Humphreys, G. W., Riddoch, M. J., and Quinlan, P. T. (1988). Cascade processes in picture identification. *Cogn. Neuropsychol.* 5, 67–104. doi: 10.1080/02643298808252927
- Jescheniak, J. D., and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *J. Exp. Psychol.* 20, 824–843.
- Kittredge, A. K., Dell, G. S., Verkuilen, J., and Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cogn. Neuropsychol.* 25, 463–492. doi: 10.1080/02643290701674851
- Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1016/j.jecp.2008.10.004
- Levelt, W. J. M. (1999). Models of word production. *Trends Cogn. Sci.* 3, 223–232. doi: 10.1016/S1364-6613(99)01319-4
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75. doi: 10.1017/S0140525X99001776
- McClelland, J. L., and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. an account of basic findings. *Psychol. Rev.* 88:375. doi: 10.1037/0033-295X.88.5.375
- McRae, K., and Boisvert, S. (1998). Automatic semantic similarity priming. *J. Exp. Psychol.* 24, 558–572.
- Meyer, A. S., Sleiderink, A. M., and Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition* 66, B25–B33. doi: 10.1016/S0010-0277(98)00009-2
- Mirman, D., and Graziano, K. M. (2012). Individual differences in the strength of taxonomic versus thematic relations. *J. Exp. Psychol.* 141, 601–609. doi: 10.1037/a0026451
- Morrison, C. M., Chappell, T. D., and Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Q. J. Exp. Psychol.* 50, 528–559. doi: 10.1080/027249897392017
- Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* 76, 165–178. doi: 10.1037/h0027366
- Navarrete, E., Basagni, B., Alario, F. X., and Costa, A. (2006). Does word frequency affect lexical selection in speech production? *Q. J. Exp. Psychol.* 59, 1681–1690. doi: 10.1080/17470210600750558
- Navarrete, E., Del Prato, P., and Mahon, B. Z. (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Front. Psychol.* 3:38. doi: 10.3389/fpsyg.2012.00038
- Navarrete, E., Del Prato, P., Peressotti, F., and Mahon, B. Z. (2014). Lexical selection is not by competition: evidence from the blocked naming paradigm. *J. Mem. Lang.* 76, 253–272. doi: 10.1016/j.jml.2014.05.003
- Navarrete, E., Mahon, B. Z., and Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychol.* 134, 279–289. doi: 10.1016/j.actpsy.2010.02.009
- Oldfield, R. C., and Wingfield, A. (1965). Response latencies in naming objects. *Q. J. Exp. Psychol.* 17, 273–281. doi: 10.1080/17470216508416445
- Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2010). The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. *Cognition* 114, 227–252. doi: 10.1016/j.cognition.2009.09.007
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., and Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* 103, 56–115. doi: 10.1037/0033-295X.103.1.56
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F
- Santesteban, M., Costa, A., Pontin, S., and Navarrete, E. (2006). The effect of word-frequency on lexical selection in speech production: evidence from semantic homogeneous naming contexts. *Cognitiva* 18, 75–84. doi: 10.1174/021435506775462436
- Scarborough, D. L., Cortese, C., and Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *J. Exp. Psychol.* 3, 1–17.
- Schnur, T. T., Schwartz, M. F., Brecher, A., and Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: evidence from aphasia. *J. Mem. Lang.* 54, 199–227. doi: 10.1016/j.bandl.2008.11.003
- Seidenberg, M. S., and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96:523. doi: 10.1037/0033-295X.96.4.523
- Thompson-Schill, S. L., and Botvinick, M. M. (2006). Resolving conflict: a response to Martin and Cheng (2006). *Psychon. Bull. Rev.* 13, 402–408. doi: 10.3758/BF03193860
- Viggiano, M. P., Vannucci, M., and Righi, S. (2004). A new standardized set of ecological pictures for experimental and clinical research on visual object processing. *Cortex* 40, 491–509. doi: 10.1016/S0010-9452(08)70142-4

- Vigliocco, G., Vinson, D. P., Damian, M. F., and Levelt, W. J. M. (2002). Semantic distance effects on object and action naming. *Cognition* 85, B61–B69. doi: 10.1016/S0010-0277(02)00107-5
- Vigliocco, G., Vinson, D. P., Lewis, W., and Garrett, M. F. (2004). Representing the meanings of object and action words: the featural and unitary semantic space hypothesis. *Cogn. Psychol.* 48, 422–488. doi: 10.1016/j.cogpsych.2003.09.001
- Vitkovitch, M., and Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: it's in the links. *J. Exp. Psychol.* 17, 664–680. doi: 10.1037/0278-7393.17.4.664
- Warrington, E. K., and Cipolotti, L. (1996). Word comprehension. the distinction between refractory and storage impairments. *Brain* 119, 611–625. doi: 10.1093/brain/119.2.611
- Warrington, E. K., and McCarthy, R. (1983). Category specific access dysphasia. *Brain* 106, 859–878. doi: 10.1093/brain/106.4.859
- Warrington, E. K., and McCarthy, R. A. (1987). Categories of knowledge further fractionations and an attempted integration. *Brain* 110, 1273–1296. doi: 10.1093/brain/110.5.1273
- Wei, T., and Schnur, T. T. (2016). Long-term interference at the semantic level: evidence from blocked-cyclic picture matching. *J. Exp. Psychol.* 42, 149–157. doi: 10.1037/xlm0000164
- Wilson, M. (1988). MRC psycholinguistic database: machine-usable dictionary, version 2.00. *Behav. Res. Methods* 20, 6–10. doi: 10.3758/BF03202594
- Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *Am. J. Psychol.* 81, 226–234. doi: 10.2307/1421267

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Harvey and Schnur. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Investigating the flow of information during speaking: the impact of morpho-phonological, associative, and categorical picture distractors on picture naming

Jens Bölte<sup>1\*</sup>, Andrea Böhl<sup>2</sup>, Christian Dobel<sup>3</sup> and Pienie Zwitserlood<sup>1</sup>

<sup>1</sup> Institut für Psychologie, Westfälische Wilhelms-Universität Münster, Münster, Germany, <sup>2</sup> Institut für Lernsysteme GmbH, Hamburg, Germany, <sup>3</sup> HNO Klinik, Universitätsklinikum Jena, Jena, Germany

## OPEN ACCESS

### Edited by:

Ian FitzPatrick,  
Heinrich Heine Universität Düsseldorf,  
Germany

### Reviewed by:

Dirk Koester,  
Bielefeld University, Germany  
Markus F. Damian,  
University of Bristol, UK  
Iva Ivanova,  
University of California, San Diego,  
USA

### \*Correspondence:

Jens Bölte,  
Institut für Psychologie, Westfälische  
Wilhelms-Universität Münster,  
Fliedner Straße 21, Münster, Germany  
boelte@uni-muenster.de

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

Received: 30 April 2014

Accepted: 23 September 2015

Published: 12 October 2015

### Citation:

Bölte J, Böhl A, Dobel C  
and Zwitserlood P (2015)  
Investigating the flow of information  
during speaking: the impact  
of morpho-phonological, associative,  
and categorical picture distractors on  
picture naming.  
Front. Psychol. 6:1540.  
doi: 10.3389/fpsyg.2015.01540

In three experiments, participants named target pictures by means of German compound words (e.g., *Gartenstuhl*–garden chair), each accompanied by two different distractor pictures (e.g., lawn mower and swimming pool). Targets and distractor pictures were semantically related either associatively (garden chair and lawn mower) or by a shared semantic category (garden chair and wardrobe). Within each type of semantic relation, target and distractor pictures either shared morpho-phonological (word-form) information (*Gartenstuhl* with *Gartenzweig*, garden gnome, and *Gartenschlauch*, garden hose) or not. A condition with two completely unrelated pictures served as baseline. Target naming was facilitated when distractor and target pictures were morpho-phonologically related. This is clear evidence for the activation of word-form information of distractor pictures. Effects were larger for associatively than for categorically related distractors and targets, which constitute evidence for lexical competition. Mere categorical relatedness, in the absence of morpho-phonological overlap, resulted in null effects (Experiments 1 and 2), and only speeded target naming when effects reflect only conceptual, but not lexical processing (Experiment 3). Given that distractor pictures activate their word forms, the data cannot be easily reconciled with discrete serial models. The results fit well with models that allow information to cascade forward from conceptual to word-form levels.

**Keywords:** picture–picture paradigm, morphology, spoken word production, cascading activation, discrete activation, semantic relatedness, associative relatedness, categorical relatedness

## Spoken Word Production

The production of a simple greeting such as “Hi” is the result of series of cognitive processes that precede articulation. Processes such as conceptualization, message generation, lexical selection, morpho-phonological processing, phonetic encoding, and monitoring all take place prior to articulation (Dell, 1986; Butterworth, 1989; Levelt, 1989; Levelt et al., 1999). How information flows between how many different processing levels is a much-debated topic, distinguishing between serial-discrete (“two-step”) models, fully cascading models and fully interactive models

(see Levelt, 1989). Interactive models allow for bidirectional information flow (from conceptual to phonological information, and vice versa). The major difference between discrete and fully cascading models concerns the information that is activated at certain processing stages, which are detailed below.

In the current study, we tested predictions derived from discrete and fully cascading models. We assessed the flow of information during speaking by investigating how distractor pictures that are not targets for speech production influence the speed with which a target picture is named. We varied the relationship between the distractor and target pictures to assess how “deeply” distractor pictures are processed. Target and distractor pictures could be semantically related (target “sunbed”, distractors “beach ball”, and “flippers”), and in addition, their names could share a morpheme (target “sheepdog”, distractors “sheep pen”, and “sheep wool”). An impact of these types of relatedness on picture naming is informative about the flow of information in speech production. To elucidate different predictions by the models that are put to test here, we briefly sketch these models.

Models of speech production agree that speaking makes demands on the following types of information. The first, conceptual/semantic information of the to-be-expressed concepts is often considered not to be lexical but part of semantic memory. Lexical information consists of grammatical aspects (e.g., word class, gender) and information about the form of words, including their morphological make up (cf. “collie” and “sheepdog”) and phonological specification (e.g., /d/ /o/ /g/). But models disagree with respect to the processing flow from conceptual to phonological information. In the serial models (Garrett, 1980; Levelt et al., 1999), speaking proceeds serially, in ordered steps, from conceptual processing to articulation. Critically, there are two distinct steps; the first step allows cascading of information, such that many representations can be active at adjacent levels of processing. The second step is only initiated when a selection process has delivered a single, complete output (cf. Levelt, 1989; Roelofs, 1997; Levelt et al., 1999; Bloem and La Heij, 2003). In discrete, two-step models, concepts activate multiple lexical entries at an initial level, labeled “lemma level”. Lemmas code the grammatical features (word class, gender, and so on), but not the morpho-phonological make up of lexical entries. Many related concepts (*dog*, *cat*, *collie*) can be active during speech production, and the activation cascades to their corresponding lemmas. Which lexical entry will be uttered is decided at the lemma level, by means of a competitive selection process (Roelofs, 1992). Selection is more difficult/takes more time when co-activated lemmas come from the same semantic category as the target (e.g., lemon–orange), because they compete more for selection than unrelated entries, or than related entries that have less semantic overlap (e.g., lemon–sour). Selection of one lemma as the target for production implies that only one lexical entry will activate its morpho-phonological word-form, and this is where cascading comes to a halt<sup>1</sup>.

<sup>1</sup>Roelofs (2008a) formulated an exception, allowing for the incidental cascading to word-form information. One and the same concept may activate multiple word forms, as is the case for near synonyms (e.g., *sofa* and *couch*).

In contrast, processing stages in fully cascading models, although temporally ordered, deliver multiple, even partial, outputs to consecutive stages, allowing for the simultaneous activation of many word forms (Dell, 1986). Some of these models do not adopt a separate lemma level (Stemberger, 1985; Humphreys et al., 1988; Caramazza, 1997; Peterson and Savoy, 1998). The selection as to which word will be uttered is non-competitive; to cite Mahon et al. (2007, p. 203) “the level of activation of a non-target does not affect the selection of the target”. Thus, there are two crucial differences between these models; (1) discrete, two-step models predict interference, reflecting competition during selection due to the presence of same-category stimuli, but fully cascading models do not and (2) cascading models allow and predict that word-form (morphological and phonological) information is simultaneously available for more than one lexical entry, but discrete two-step models do not. Jescheniak and Schriefers (1998), Peterson and Savoy (1998), Rapp and Goldrick (2000), as well as Goldrick (2006) offer overviews of the discrete/cascading controversy.

## Cascaded or Discrete Processing, Paradigms, and Evidence

In the following, we summarize the evidence in favor of fully cascaded, and against discrete, processing in speech production, and introduce the paradigms used together with their basic findings. Next, we present the manipulations and predictions for the three experiments of our study.

So far, evidence for cascaded processing comes from (1) speech errors, (2) picture naming experiments with word distractors, and (3) picture-naming experiments with picture distractors – the paradigm that we also used here. Speech-error data from patients and simulations of speech-error data argue against discrete models (Rapp and Goldrick, 2000). The relevant error type concerns mixed errors. A mixed error is a word that is semantically and phonologically related to the intended word (e.g., saying *cat* instead of *calf*). Taking error distributions into account, such errors are more likely to occur than pure semantic errors (e.g., saying *dog* instead of *cat*; Dell and Reich, 1981; Martin et al., 1996). Rapp and Goldrick (2000) argue that mixed errors can only occur in fully cascading models and/or interactive models, but not in discrete serial models. Roelofs (2004), however, argues that mixed errors result from erroneously selecting two lemmas instead of one. In his view, erroneous selection of multiple lemmas is not restricted to mixed errors but is also the basis for blend errors (e.g., *close* + *near* → *clear*, cf. Roelofs, 1992) and for activating multiple word forms of near synonyms.

The next source of evidence comes from picture–word interference (PWI) studies. In paradigms with word distractors, a picture that has to be named is accompanied by (written or spoken) words that can be ignored. Such PWI studies consistently show that picture naming is faster when distractor words are related in form (picture of a calf, distractor “cart”) than when not (picture of a calf, distractor word “bowl”; Meyer and Schriefers, 1991; Levelt et al., 1999, for an overview). This also holds for

cases of large form overlap, when target and distractor word share a morpheme (picture of a sheepdog, distractor “sheep wool”), even when there is no obvious semantic relation between the concepts specified by picture and distractor word (e.g., picture of a hummingbird, distractor “jailbird”; see Lüttmann et al., 2011a). Semantically related distractors that do not share the target’s semantic category (picture of a cow, distractor “milk”) tend to speed target naming. This is often interpreted as stemming from the non-lexical, conceptual level (see La Heij et al., 1990; Alario et al., 2000). However, picture naming is slowed when the distractor comes from the same semantic category as the target (picture of a calf, distractor “sheep”). This is interpreted either as evidence for competitive lexical selection (Schriefers et al., 1990; Roelofs, 1992; Levelt et al., 1999), or as originating from post-lexical problems, occurring when a semantically related distractor word occupies a prominent place in the serial output buffer, thus hindering the timely output of the picture name (Mahon et al., 2007).

With respect to the issue of full or partial cascading, experiments with word distractors that are related in both meaning and form to the target picture (e.g., target picture *calf*, distractor word “cat”) revealed interactive effects: form relatedness counteracts the negative consequences of a shared semantic category between target and distractor (Starreveld and La Heij, 1995, 1996; Damian and Martin, 1999). Moreover, near synonyms or cognates (for bilinguals) activate multiple word forms (Jescheniak and Schriefers, 1998; Peterson and Savoy, 1998; Costa et al., 2000), also supporting the notion of full cascading.

Finally, some studies using multiple pictures instead of pictures and words also argue for a continuous cascade of information. In picture–picture paradigms, a target picture for naming is accompanied by one or more distractor pictures that should not be named (Glaser and Glaser, 1989; Morsella and Miozzo, 2002; Damian and Bowers, 2003; Navarrete and Costa, 2005; Meyer and Damian, 2007; Oppermann et al., 2008, 2014; Roelofs, 2008a). Morsella and Miozzo (2002) asked their participants to name one of two differently colored, superimposed line drawings, and to ignore the other. Faster picture-naming latencies were obtained for phonologically related (*bed-bell*) than for unrelated pictures (*hat-bell*); see also Damian and Bowers, 2003; Navarrete and Costa, 2005; Meyer and Damian, 2007; Roelofs, 2008a), suggesting that the distractor picture activates its phonological representation, which then (because of phonological overlap) speeds up target naming. Jescheniak et al. (2009), who failed to replicate this data pattern, suggest that differences in amount of phonological overlap, the inclusion of the distractor pictures in the response set, and/or subtle differences in name agreement might be responsible for the divergent results. Importantly, and despite the absence of semantic effects in Morsella and Miozzo (2002), the presence of phonological effects argues for the full cascading of activation.

The absence of semantic effects (e.g., *table-bed*) in Morsella and Miozzo (2002) is rather startling, given that language production proceeds from semantic to phonological representations. In general, studies using picture–picture paradigms showed diverging results for categorically related distractor pictures: facilitation (Bloem and La Heij, 2003;

Roelofs, 2008a), interference (Glaser and Glaser, 1989), or no effects (Humphreys et al., 1995; Morsella and Miozzo, 2002; Damian and Bowers, 2003; Navarrete and Costa, 2005). It is not yet fully understood what causes the different result patterns. With picture distractors, it does not seem mandatory that all available conceptual information is automatically encoded lexically, and the task, target set, attention to the distractor picture, and material manipulations might play an important role.

One important factor concerns the availability of distractor pictures as (potential) targets – sometimes manipulated by including all pictures in the target set. This fits with data from Aristei et al. (2012), who presented two pictures simultaneously that both had to be named to produce a novel compound (e.g., *lion dog*). Participants were slower in producing such novel noun–noun compounds when the two pictures were categorically related (*lion dog*) than when not (*chair dog*). Aristei et al. (2012) argue that this provides evidence for lexical competition.

Similar conclusions can be drawn from studies by Oppermann et al. (2008, 2014), who presented a target and a distractor picture simultaneously, while spoken words that were semantically related, phonologically related or unrelated to the distractor picture served as additional distractors. When target and distractor objects were similar in shape, semantically related distractor words slowed down target picture naming relative to unrelated distractor words. This suggests that the concepts of the target and distractor pictures enter the lexicalization process provided that distractor pictures capture sufficient activation, because they are similar in shape to the target and are “boosted” by related distractor words.

Thus, whether semantic effects can be registered in picture–picture paradigms seems to depend on the amount of attention to the distractor picture (Jescheniak et al., 2014), on how to signal the target picture and/or on the particular task implemented (Glaser and Glaser, 1989; Bloem and La Heij, 2003; Damian and Bowers, 2003).

Note that evidence for cascading semantic information *per se* does not distinguish between fully cascading and discrete, two-step models, but the direction of semantic effects (facilitation, interference) does. Interference, due to same-category membership of distractors and targets, is predicted by two-step models but not by fully cascading models. It plays an important role in the discussion about lexical-competition (discrete models), and fully cascading models provide an explanation of such interference effects in terms of a post-lexical response-buffer. We will discuss this further below.

## The Picture–Picture Paradigm, Conditions, and Predictions

To further test the predictions of discrete and fully cascading models, we opted for the picture–picture paradigm, because its suitability to test for activation of lexical form (morphology, phonology) of non-target pictures. We presented three different pictures, one of which was the target for naming. Which picture had to be named was either signaled by a cue that appeared with varying delays (Experiments 1 and 2), or was unequivocally signaled by presenting the target picture with some delay after

the non-target (distractor) pictures (Experiment 3). We used multiple distractors (1) because effects can be larger with two than with one distractor (Melinger and Abdel Rahman, 2004) and (2) to create more uncertainty as to which picture has to be named eventually.

A first manipulation concerned the nature of semantic overlap between distractor and target pictures, which was either associative or categorical. Note that both models allow for the activation of multiple concepts (of all three pictures). To our knowledge, associatively related distractors (e.g., *sailor* and *ship*) or distractors representing semantic features of the target object (e.g., *porthole* and *ship*) have not been investigated so far within the picture–picture paradigm. It is well established that associatively and categorically related distractors have different effects in the PWI paradigm (Bölte et al., 2003, 2005; Costa et al., 2005; Mahon et al., 2007). Why words that are semantically associated or that represent semantic features of the target picture facilitate, whereas words that specify a same category member inhibit picture naming, is still a matter of intense debate (see Costa et al., 2005; Mahon et al., 2007; Abdel Rahman and Melinger, 2009; Janssen, 2013; Roelofs et al., 2013; Mahon and Navarrete, 2014). Whereas both associative and categorical similarity should induce priming at the level of conceptual representations, they seem to differ at lexical or post-lexical levels. According to discrete models, the activated lemmas of same-category concepts cause havoc during the selection of the lexical entry that is the target for speaking (Roelofs, 1992; Levelt et al., 1999), because they are confusable with the target and seem such valid responses (saying “dog” to a picture of a cat is more likely than saying “purr”). If we obtain categorical competition effects in a picture–picture paradigm, this is clear evidence for the existence of a competitive lexical selection process, and argues against prominent cascading models (Caramazza, 1997). Note that categorical interference from pictures also speaks against the response-exclusion hypothesis (Finkbeiner and Caramazza, 2006; Mahon et al., 2007). According to this hypothesis, the interference by categorically related distractor words observed in PWI is due to the fact that these distractors, because they are words, enter the articulatory response buffer that channels verbal responses for output. Words that are semantically related to the correct response (the picture name) are harder to remove from this buffer than unrelated words, hence, the interference. Most importantly, this holds for verbal stimuli only, not for pictures (see Jescheniak et al., 2014).

As stated above, discrete and fully cascading models also make different predictions concerning the impact of morpho-phonologically related distractor pictures on the speed of target-picture naming. We used German compound words, as distractors (*garden hose*, *garden gnome*) and targets (*garden chair*), because such stimuli have the advantage of sharing both semantic and form information. We crossed the type of semantic relation (associative vs. categorical) with form overlap, in terms of shared morphemes (initial or final morphemes of compound names). To our knowledge, combining semantic and form overlap has not been done before with the picture–picture paradigm (not even with partial overlap, as in “cart” and “calf”). The critical evidence for full cascading is when distractor pictures

also activate their word-form information. This should not be the case according to discrete, two-step models.

As stated earlier, form-relatedness has been reliably demonstrated with the PWI paradigm, when a target picture (e.g., of a *football*) is accompanied by a distractor word that shares phonemes or morphemes with the target (e.g., “foodstuff” or “footstool”; cf. Meyer and Schriefers, 1991; Zwitserlood et al., 2002; Lüttmann et al., 2011a). In picture–word paradigms, distractor words automatically activate lexical information. Their processing proceeds from phonemes or graphemes via word-form and syntactic information to concepts. Word distractors can thus influence picture naming at all (lexical) levels. This is different for picture distractors that can only influence the lexical processing of the target if the distractors themselves activate their lexical information. Thus, if naming a “football” is easier when the distractor pictures show a “footprint” and a “footstool”, this provides clear evidence for the activation of morpho-phonological information belonging to the distractor pictures, and for full cascading of information during speech production. In contrast, the lack of activation of the distractor pictures’ word forms supports discrete, only partially cascaded models.

We thus included the following target-distractor conditions in our study. The relation between a target picture (e.g., a *garden chair*)<sup>2</sup> and its two different distractor pictures was either (1) associative with morpho-phonological<sup>3</sup> overlap (+A+M) in the first constituent (e.g., *garden hose*, *garden gnome*), (2) same-category combined with morpho-phonological overlap +C+M) in the second constituent (e.g., *rocking chair*, *office chair*), (3) merely associative (+A–M; e.g., a *swimming pool*, *lawn mower*) or (4) merely categorically related (+C–M; e.g., *office desk*, *shoe rack*), thus without morpho-phonological overlap, or (5) completely unrelated (e.g., *billiard ball*, *sock suspender*).

Our rationale to use both types of semantic relation is as follows: if effects in the picture–picture paradigm solely originate at a conceptual level, effects should be similar for categorically and associatively related distractors. If interference – or reduced facilitation, relative to associatively related pictures – is observed for categorically related distractors, this is evidence for their lexical coding. Such effects provide clear evidence for competitive lexical selection (cf. Levelt et al., 1999), and against fully cascading models as well as against the response-exclusion hypothesis that only applies to words, not to pictures (Mahon et al., 2007). Note that reliable interference due to categorically related context pictures has rarely been observed in picture–picture studies reported so far, which either suggests that distractor pictures are not lexically coded automatically (cf. Damian and Bowers, 2003; Jescheniak et al., 2014), or that conceptual facilitation and lexical competition cancel each other out.

We also implemented the distinction between same category and association with pictures whose names are

<sup>2</sup>German compounds are written without spaces.

<sup>3</sup>We use the term morpho-phonological overlap to signal that the target constituent overlaps phonologically with the distractor constituent. The phonological overlap constitutes at the same time a free morpheme. Morpho-phonological overlap is different from pure phonological overlap (Roelofs and Baayen, 2002; Zwitserlood et al., 2002).

morpho-phonologically related to the target picture's name. Morphological relatedness is not specified at the conceptual level (Caramazza et al., 1988; Levelt et al., 1999; Janssen et al., 2008). If all effects are conceptual, without any lexical involvement, these should behave in the same way as associatively or categorically related pictures whose name is morpho-phonologically unrelated to the target. If distractor pictures are lexically processed, but at the lemma level only (in discrete models), the same predictions hold as formulated above for morphologically unrelated distractors. But if distractor pictures are processed all the way down to their word-form level, where morphology is specified, we expect facilitation due to morpho-phonological relatedness. In PWI studies, where form effects are obvious because the distractors are words, facilitation was observed with distractors and targets overlapping at word onset and offset, both with monomorphemic words (e.g., *power* and *towel* with the picture of a *tower*) and with morphologically related (e.g., *tea rose* and *rosebush* with the picture of a *rose*) distractors (Meyer and Schriefers, 1991; Zwitserlood, 1994; Zwitserlood et al., 2002; Belke, 2005; Lüttmann et al., 2011b).

When distractor pictures are encoded at the level of word form, we expect additional facilitation due to shared morphemes, relative to an unrelated baseline, in both morpho-phonological conditions (+A+M and +C+M). The size of effects might differ because of lexical competition in the +C+M condition. The purely associatively related distractor condition (+A–M) that does not induce much lexical competition should also reveal facilitation, but the categorically related distractors (+C–M) should show no effect or even interference. This is because they are conceptually related to the target (resulting in facilitation) but also lead to interference due to lexical competition with the target. Keep in mind that the presence of interference, or reduced facilitation, in the +C conditions speaks for competitive lexical selection (Levelt et al., 1999), but is incompatible with full cascading models (Caramazza, 1997) and with response-exclusion (Mahon et al., 2007).

Finally, we manipulated the signaling of the target picture, either by a cue (an arrow, Experiments 1 and 2) or by a time delay (Experiment 3). We varied the onset of the target cue (Experiments 1 and 2) relative to the stimuli display (SOA). This had two functions. First, given that it is unclear whether multiple pictures automatically activate their lexical information, a longer uncertainty as to which picture has to be named (implemented by a larger SOA) might invite a lexical activation of all pictures. A large SOA might invite the lexical coding of more than one picture, but a small SOA should not.

The next issue concerns the time course of lexical activation. In the PWI paradigm, the impact of semantic and phonological distractors on picture naming depends on the temporal relation between word distractor and target. Categorical and associative effects are largest if the distractor precedes the target, while phonological effects arise when the distractor follows the target or is presented simultaneously with the target (Glaser and Döngelhoff, 1984; Schriefers et al., 1990; Meyer and Schriefers, 1991; Alario et al., 2000; Jescheniak et al., 2005). Similarly, providing more or less time before it becomes clear which

picture is to be named might lead to the involvement of different processing levels. An SOA of 200 ms between the onset of the pictures and the cue may well be too short for the activation of word-form information, but an SOA of 600 ms should suffice. So, the SOA manipulation was used to invite or discourage the (strategic) lexical coding of all (or some) pictures before the target was signaled. In Experiment 3, it was clear to the participants that the two objects that appeared first were never to be named, because the target was signaled by means of an onset delay. In this case, lexical activation of distractor pictures might be completely absent.

We also monitored eye-movements, in addition to voice-key latencies. The reason was to investigate whether targets had to be fixated for correct naming, and whether distractors had to be attended overtly to affect target naming. Previous research using eye-movements required their participants to name all displayed objects (cf. Meyer et al., 1998). In such tasks, participants look at the object until its phonological form is planned. On the other hand, Dobel et al. (2007) showed that fixations of scene elements are not necessary to identify (and name) agents, actions and patients of action scenes. Unlike in the study by Meyer et al. (1998), participants were not asked to give speeded responses, and sometimes were even prevented from making eye-movements into the scene, because of very short scene presentation durations. So, speakers can name visual stimuli without overt attention, but they may well look at objects to facilitate object recognition and name retrieval (Meyer et al., 2012). It is still unknown whether distractor pictures have to be fixated at all to affect target naming.

## Experiment 1: Cue Onset 600 ms

### Method

#### Participants

Forty participants from the Westfälische Wilhelms-University of Münster took part in the experiment. They were either paid 4 € or received course-credit for their participation. All had normal or corrected-to-normal vision and were native speakers of German.

#### Material

We used pictures that are named with noun–noun compounds to implement the morpho-phonological similarity, concurrent with semantic similarity, between target and distractor pictures. Material selection was a multi-phased procedure. First, we selected noun–noun compounds from the Celex lexical database, discarding all compounds that were not depictable (Baayen et al., 1993). Next, distractors were constructed for each target (*Gartenstuhl*, lawn chair) such that there were three to five distractors per Distractor Type: (1) +A+M, associatively and morpho-phonologically related (e.g., *Gartenzwerg*, garden gnome; *Gartenschlauch*, garden hose), (2) +C+M, categorically and morpho-phonologically related (e.g., *Schaukelstuhl*, rocking chair; *Bürostuhl*, office chair), (3) +A–M, associatively but not morpho-phonologically related (*Rasenmäher*, lawn mower; *Schwimmbecken*, swimming pool) and (4) +C–M, categorically but not morpho-phonologically related (e.g., *Schreibtisch*, desk;

*Schuhregal*, shoe rack), and finally, control distractors that were neither categorically, associatively nor morpho-phonologically related to the target (e.g., *Zahnbürste*, tooth brush; *Billardkugel*; billiard ball). This resulted in a set of 377 compounds (22 targets, 355 potential distractors). Colored pictures for these compounds were taken from the Hemera Photo Objects (n.d.) database, or from the internet.

The material was tested in two pretests: (1) an offline name-agreement test in combination with a semantic rating task and (2) an online name-agreement test. Twenty participants took part in the offline tests, another 15 served in the online test. All participants came from the same population as mentioned above and received a similar compensation. In the offline name agreement test, each distractor picture was presented alongside its target picture, resulting in 355 trials. Participants were asked to write the word that described best the depicted objects and to rate their semantic relatedness, using a 5-point scale (1 = unrelated, 5 = related). The online name agreement test served to assess the preferred naming of the picture under conditions similar to the actual experiment (see **Table 1** for relevant means and SDs). Trials in this test were structured as follows: a fixation cross appeared on a computer screen for 250 ms, followed by the picture that remained on the screen for 600 ms. Time-out was set to 1500. Participants were asked to name the picture as quickly as possible.

We selected all pictures that were predominantly named with a morphologically complex word in the offline (targets mean: 79%, SD: 6, range: 70–85%; distractors mean: 91%, SD: 12, range: 55–100%) as well as in the online naming test (targets: mean: 81%, SD: 14, range: 60–100%; distractors: mean: 84%, SD: 14, range: 53–100%). This resulted in 15 target pictures, each with two different distractor pictures in each of the five distractor conditions. Mean ratings of all pretests for the selected items are provided in **Table 1**. The semantic relatedness judgments were evaluated with the help of a one-way univariate repeated measures ANOVA over items, using semantic relatedness judgments as dependent variable and Condition (+A+M, +C+M, +A–M, +C–M) as factor. The main effect Condition was not significant [ $F(3,42) = 2.172$ ,  $MSE = 0.758$ ,  $p = 0.105$ ,  $\eta_g^2 = 0.117$ ].

Targets and distractors were distributed over five lists, with list order counter-balanced across participants. Participants were presented with all lists. An additional 24 filler trials, each with pictures of three morphologically complex but unrelated words,

were included in each list, to increase the number of unrelated trials (e.g., *Schlittschuh*, ice skate; *Bohrmaschine*, drilling machine; *Sonnenblume*, sunflower). Each block consisted of 39 trials plus six warm-up trials.

## Apparatus

Pictures (ranging from 22 × 245 pixel for “toothbrush” to 241 × 207 pixel for “oil lamp”) were presented on a 21-inch Samsung SyncMaster 1100p plus CRT monitor (1024 × 768 pixel, frame rate: 85 Hz), controlled by a Dell-Dimension 4200 IBM-compatible PC. Participants were seated approximately 60 cm in front of the monitor. Eye-movements were recorded with an EyeLink II (2004) eye-tracker, with a sampling rate of 500 Hz and an eye position resolution of less than 0.5°. The eye-tracker was controlled by a Dell-OptiPlex 280. Onset naming latencies were recorded with a voice key.

## Procedure

Participants were tested individually in a quiet room. They received a written instruction. They were informed that three pictures would appear on the screen and that shortly after picture onset an arrow would signal the picture that they had to name. Participants were asked to name the target picture as quickly and accurately as possible such that the experimenter could correctly identify the target among the other objects on the display (see Bölte et al., 2009). Before the experiment proper, the following steps were taken. First, to minimize target name variation, participants received a booklet with target pictures and names. Second, after having read the booklet, each target picture was presented again for naming on the computer screen. Third, the eyetracker was calibrated and validated using a nine-point calibration type (HV9). Upon successful validation, the experiment started. A drift-correction was applied before each trial using the fixation point.

Trial structure was as follows: a fixation point, centered in the middle of the screen, indicated the beginning of a new trial. After successful fixation, the trial began and three pictures appeared in one of four possible configurations. Either there was one picture left of, one right of (160 pixel away from screen center) and one above (or below) the fixation point (150 pixel away from screen center) or one above, one below and one left (or right) of the fixation point (6.9° apart). An arrow appeared 600 ms after picture onset, signaling the target object. Target position on a list was (nearly) equally distributed (10 top, 10 left, 10 right, 9 bottom). Pictures disappeared with the participants’ voice onset or after 5000 ms. Stimuli were presented as colored photographs on a white background. The experimenter wrote down the participants’ answers.

## Results

Responses different from expected names (1.6%), disfluencies (.8%), voice-key failures (0.1%), and time-outs (1.0%) were excluded from the analyses. Responses given before cue onset were also excluded (2.4%). No item set, but two participants had to be excluded from the analyses due to missing data.

**TABLE 1 | Semantic relatedness rating and name agreement data from off- and online tasks, as a function of distractor condition (SD in parentheses).**

Semantic relatedness rating	Percentage name agreement	
	Offline task	Online task
+A+M	92.0 (11.6)	85.5 (13.6)
+C+M	87.2 (14.5)	82.7 (15.1)
+A–M	90.7 (10.6)	78.9 (14.1)
+C–M	88.4 (13.2)	80.4 (14.4)
Unrelated	97.5 (5.2)	88.4 (10.9)

Mean voice-key latencies measured from cue onset served as dependent variable<sup>4</sup>. (see **Table 2** for mean reaction times (RT) and standard errors; **Figure 1** displays the effects (RT control condition – RT experimental condition) per experiment). Repeated-measurement factors were Presentation (1–5) and Distractor Type (+A+M, +C+M, +A–M, +C–M, unrelated) in an initial two-ways repeated measures ANOVA. Participants named pictures faster toward the end of the experiment, as indicated by a significant linear trend for the factor Presentation [ $F_{(1,37)} = 96.469$ ,  $MSE = 9739$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.723$ ]. There was no significant interaction of Distractor Type and Presentation,  $F < 1$ . Therefore, Presentation was dropped from further analyses. Most importantly, this analysis also yielded a significant main effect of Distractor Type [ $F_{(4,148)} = 5.983$ ,  $MSE = 17894$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.021$ ]<sup>5</sup>.

A two-ways repeated measure ANOVA with the factors Morphological Relatedness (related, unrelated) and Semantic Relatedness (associated, categorically related) using effect as dependent variable (control condition–experimental condition) yielded two significant main effects and a non-significant interaction (Morphological Relatedness:  $F_{(1,37)} = 8.024$ ,  $MSE = 3835$ ,  $p = 0.007$ ,  $\eta_g^2 = 0.029$ ; Semantic Relatedness:  $F_{(1,37)} = 13.810$ ,  $MSE = 2966$ ,  $p = 0.001$ ,  $\eta_g^2 = 0.038$ ; interaction:  $F < 1$ ).

Mean voice key latencies of the +A+M condition were faster than those of the unrelated condition [one-sided  $t$ -tests:  $t_{(37)} = -3.442$ ,  $p = 0.001$ ] and those of the associative condition without morpho-phonological overlap, +A–M [ $t_{(37)} = -2.517$ ,  $p = 0.016$ ]. There was a trend toward significance when comparing the +A+M mean voice key latencies with those of the +C+M condition [ $t_{(37)} = -1.585$ ,  $p = 0.061$ ]. Notice that we did not correct these and all following *post hoc* tests for multiple comparisons. Mean picture naming latencies in the category distractor condition +C–M were numerically longer but did not differ significantly from those in the unrelated condition [two-sided  $t$ -test:  $t_{(37)} = 1.045$ ,  $p = 0.303$ ]. Thus, as in previous

research, same-category members showed no facilitation, but also did not reliably interfere with picture naming in a picture–picture setting (cf. Glaser and Glaser, 1989; La Heij et al., 2003). Note that the main effect of semantic relatedness was significant, showing that an associative relation between distractors and target induced facilitation (37 ms) but a categorical relation did not (2 ms).

Fixations and dwell-time were measured from the onset of the pictures, with the help of the EyeLink Data Viewer program. Dwell-time was defined as the summation of the duration of all fixations on an interest area. Fixations reflect whether a specific item was fixated at all, from picture onset until reaction or trial end.

The eye-tracking data showed that participants fixated only one of the displayed objects in 36.6% of the trials (target: 29.1%, one distractor: 7.5%). Two objects were fixated in 33.9% of the trials (target and one distractor: 31.9%, both distractors: 1.7%). All three objects were looked at in 10.9% of the trials. All other fixations (19.0%) fell outside the objects (see **Table 3** for an overview). The number of gazes shows that participants looked at the target object most often, which does not come as a surprise. As has been known for a long time, fixations – as a measure of overt attention – are not needed for the correct perception of objects or scenes (Fei-Fei et al., 2005). Evidently, targets can be and were named correctly without overt attention, and it is thus very likely distractors can also exert an influence on target naming without overt attention. Thus, overlapping stimulus configurations, as in the variant of Morsella and Miozzo (2002) are not mandatory for obtaining voice-onset latency effects of distractors. However, the visual angle and presentation time used here allow covert attention shifts. Two ANOVAs, one with first fixation onset on the target, the other with dwell time on the target as dependent variable and Distractor Type as factor showed no significant effects ( $F < 1$ ).

## Discussion

To summarize, Experiment 1, with 600 ms time before the target was signaled, revealed both semantic effects (positive and null) as well as facilitation by shared morpho-phonological information with distractor pictures. Distractor pictures that were associatively related to the target picture clearly speeded target naming. Overall, categorically related distractor pictures showed no effect (2 ms). The large and reliable difference between the two semantic conditions, evident in the main effect of semantic relation, with 37 ms facilitation due to associatively related distractors but no effect for categorically related distractors (2 ms), is in fact evidence for an impact of lexical competition on conceptually induced facilitation.

This modulation of conceptual/semantic facilitation by lexical competition fits with discrete models, but not with fully cascading models (Caramazza, 1997), nor with the response-buffer explanation of interference caused by word distractors (Mahon et al., 2007). The main effect of morpho-phonological relatedness, with 33 ms facilitation when morphological relatedness is present but no effect (4 ms) without such overlap, clearly indicates the presence of word-form information of distractor pictures. This replicates the word-form effects with overlapping, colored picture

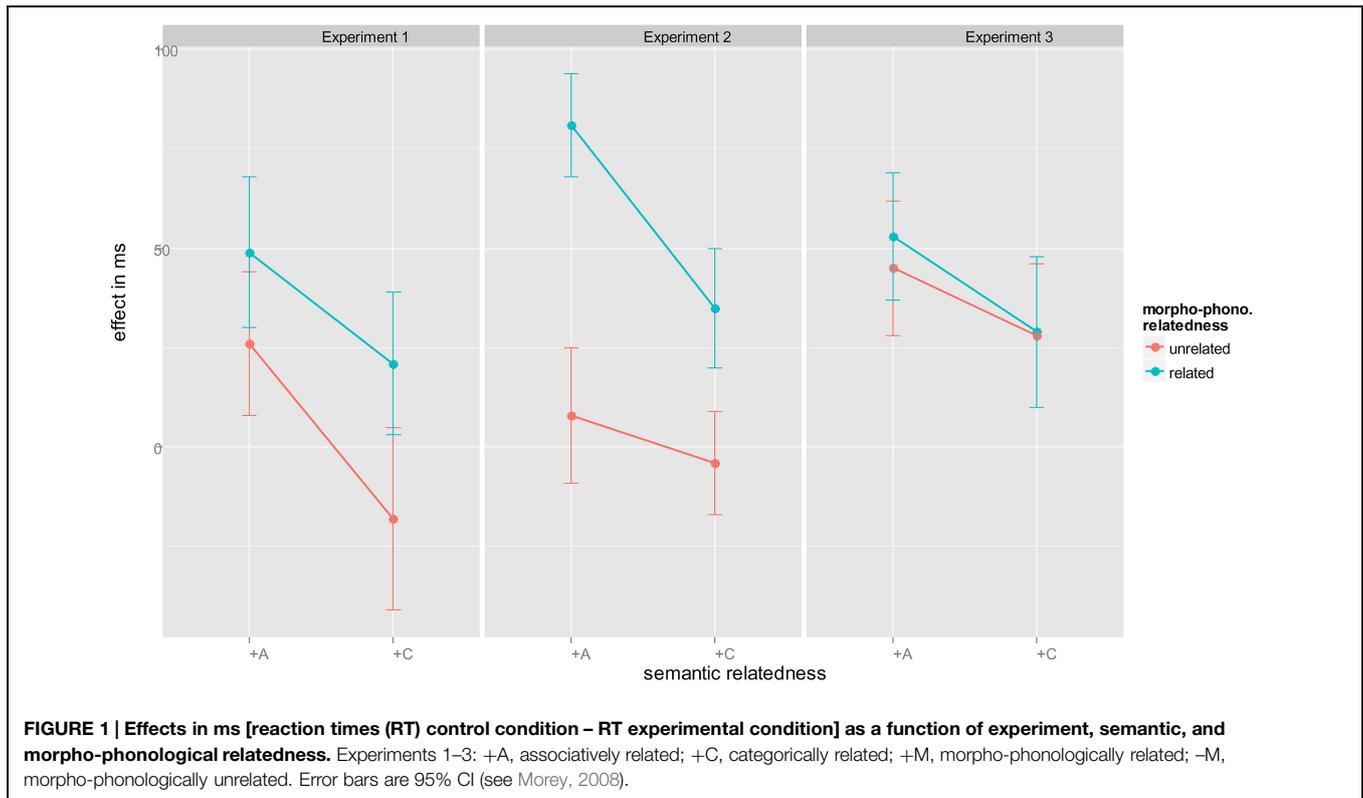
<sup>4</sup>We do not report F2-analyses in this study because targets were repeated over conditions and not nested under conditions (Clark, 1973; Raaijmakers et al., 1999). Linear mixed effects models that have been suggested as alternative to F1 and F2-analyses converged only without random slopes. The LME-analyses corroborated the reported results, but we do not report them here because we felt that the low number of trials per condition renders questionable the results of such analyses (Barr et al., 2013).

<sup>5</sup>We used SPSS to compute  $\eta_g^2$  and a spreadsheet provided by Lakens (2013) to compute  $\eta_g^2$ .

**TABLE 2 | Mean picture naming latencies and standard error (in parentheses) as a function of Distractor Type and Experiment.**

Experiment	Distractor Type				
	+A+M	+C+M	+A–M	+C–M	Unrelated
1	462 (25)*	490 (26)	485 (26)*	529 (27)	511 (30)
2	699 (21)*	745 (20)*	772 (23)	784 (20)	780 (21)
3	730 (24)*	754 (24)*	738 (24)*	755 (26)*	783 (28)

The asterisk (\*) denotes a significant difference to the unrelated condition.



presentation (Morsella and Miozzo, 2002), and provides evidence for full cascading within the language production system.

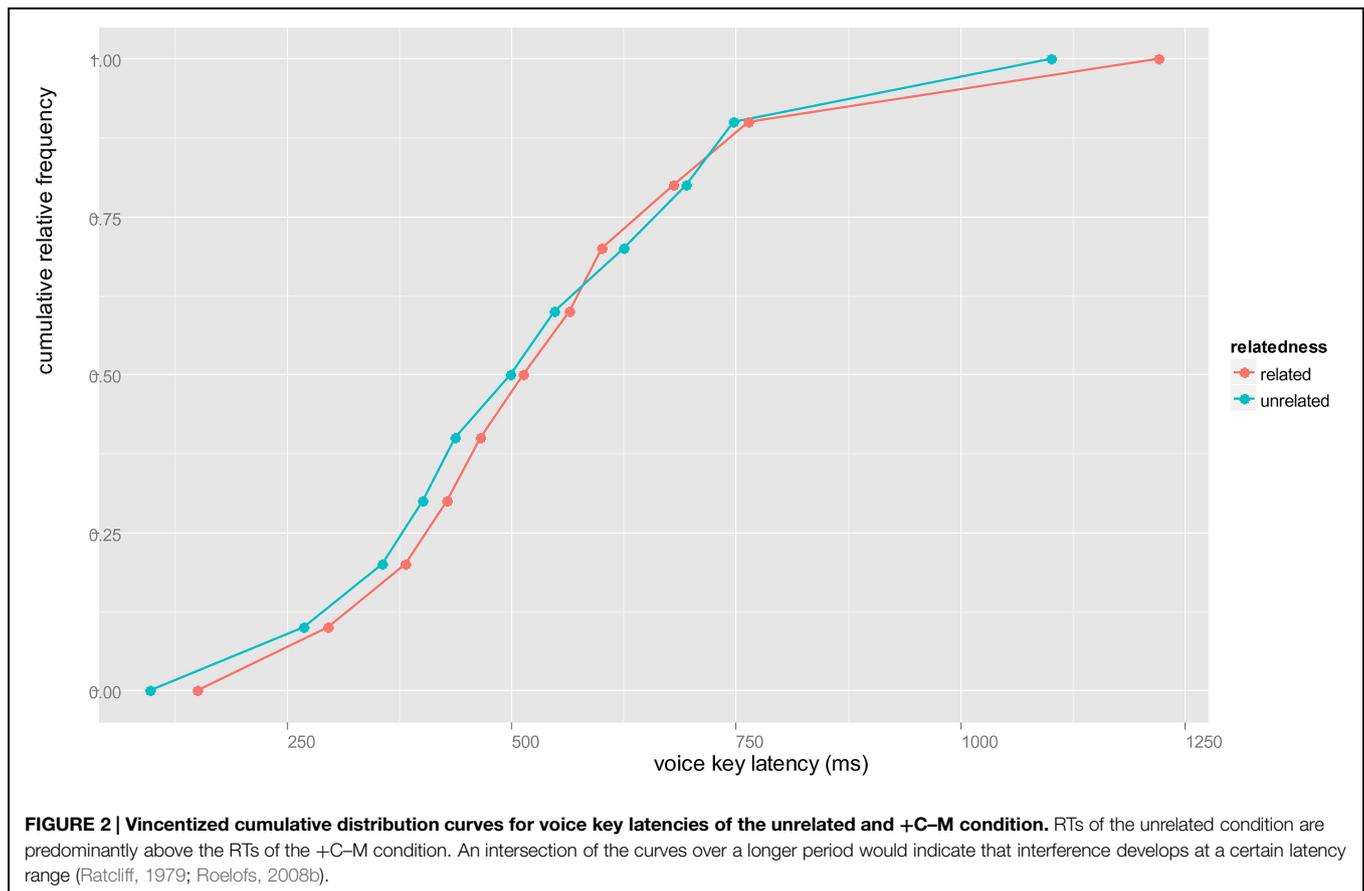
To our knowledge, there are no picture–picture studies with associative relations between distractors and target. Our

participants named target pictures faster in the presence of associatively related distractors. This replicates results from PWI studies (e.g., Bölte et al., 2003, 2005; Costa et al., 2005). Whereas semantic facilitation can be explained by activation at the non-lexical, conceptual level (see also La Heij et al., 2003), the fact that such semantic effects disappear when distractors and targets are from the same semantic category clearly indicates lexical involvement. Unlike others, we obtained no reliable interference relative to the unrelated condition. The closest comparison is a study by Humphreys et al. (1995), who also used a post-cue picture–picture procedure and observed semantic inference for categorically related pairs (e.g., horse–tiger). One difference between our study and Humphreys et al. (1995) is that naming responses were very slow, nearly twice as slow as ours. This suggests that interference might develop over time, but visual inspection of our data does not support this (see Figure 2), as there is no indication of interference at longer RTs.

Let us turn now to the interpretation of the “null effects” for categorical distractors. One argument could be that the distractor pictures never entered the lexical system to start with. But if distractors are not lexicalized, no effects of morpho-phonological relatedness should have been observed. In the absence of associative distractors, it would have been difficult to interpret the null effect, but compared to the clear facilitation for associative stimuli, the null effect seems to indicate that interference occurred, but was canceled out by facilitation due to semantic similarity. Note that according to the pretest, associative, and categorical stimuli were equally related to their targets. The combination of facilitatory conceptual effects, both

**TABLE 3 | Percentage gazes broken down by condition and fixated object for Experiments 1–3.**

Experiment	Fixated object	Distractor condition				
		+A+M	+C+M	+A–M	+C–M	Unrelated
1	Target	6.9	5.1	5.7	5.8	5.6
	Distractor(s)	1.5	1.9	2.1	1.9	1.8
	Target and distractor(s)	7.8	9.1	8.7	8.4	8.7
	Nothing	4.0	3.9	3.4	3.7	4.0
2	Target	11.0	9.7	9.9	9.5	8.7
	Distractor(s)	0.6	0.7	0.6	0.5	0.5
	Target and distractor(s)	6.1	7.6	8.1	8.3	8.7
	Nothing	2.4	1.9	1.4	1.7	2.1
3	Target	12.2	11.2	10.3	11.4	10.6
	Distractor(s)	0.2	0.1	0.1	0.1	0.3
	Target and distractor(s)	5.6	6.6	7.1	6.3	6.6
	Nothing	1.3	1.1	1.6	0.8	1.3



for categorical and associative distractors, with an inhibitory lexical effect for categorical distractors only fits well with the idea of lexical competition implemented in the model proposed by Levelt et al. (1999). Semantic competition due to picture distractors is not predicted by the cascading model by Caramazza (1997), nor is it compatible with the post-lexical explanation of semantic interference that was devised for effects of word distractors (Mahon et al., 2007).

The type of semantic relation and the position of morphological overlap between distractors and target are naturally confounded. Associatively related distractors (e.g., *garden gnome*) overlap with the target name (e.g., *garden chair*) in their onset, sharing their modifier, while categorically related distractors overlap with the target in head position (e.g., *rocking chair*). There are no left-headed compounds in German that would allow separating overlap and semantic relatedness. Given that all three picture names started the same (e.g., *garden gnome*, *garden chair*, *garden fence*), participants could have prepared at least the modifier, in trials with associated stimuli, before even knowing which one was the target. Note, however, that this was not possible for the +A-M condition, which also showed semantic facilitation. Nevertheless, some additional processing advantage in the +A+M condition might result from phonological preparation – which still constitutes a down-stream lexical effect of word-form access and phonological encoding.

Given the SOA of 600 ms, it is quite possible that our participants started the lexical encoding of one or more pictures before the cue appeared. Although in discrete models, a parallel phonological encoding should not happen even in those situations, Experiment 2 was designed to minimize such preparation effects, by reducing the cue onset time to 200 ms.

## Experiment 2: Cue Onset 200 ms

We reduced the SOA between the onset of the three pictures and the cue from 600 to 200 ms. A shorter cue-onset asynchrony provides less time for lexical activation of all pictures, and thus less time for an impact of lexical competition and of word-form similarity. Hence, a phonological preparation effect that might help target naming in cases of onset overlap (as with the associatively related stimuli) could be reduced. As a consequence, overall positive semantic (associative and categorical) effects, if present, might become more pronounced.

## Method

### Participants

Forty participants selected from the same population as before were tested. None had participated in Experiment 1 or in the pretests. They received the same compensations as the

participants of Experiment 1. All had normal or corrected-to-normal vision and were native speakers of German.

## Procedure

The same material and apparatus as in Experiment 1 was used. The only difference to the previous experiment was that the cue signaling the target appeared 200 ms after the onset of the three pictures, instead of 600 ms. All other aspects of the procedure remained the same.

## Results

Responses different from expected names (2.1%), disfluencies (0.5%), voice-key failures (0.6%), time-outs (2.6%) and reactions before cue onset (0.2%) were excluded from the analyses. No item set or participant had to be excluded from the analyses. **Table 2** lists mean RTs and standard errors as a function of Distractor Condition. One difference to Experiment 1 is obvious at first sight: latencies are much longer overall.

Voice-key latencies measured from cue onset were averaged over participants and submitted to separate ANOVAs. We first analyzed the results with Presentation (1–5) and Distractor Type (+A+M, +C+M, +A–M, +C–M, unrelated) as factors. A significant linear trend for the factor Presentation indicated that participants named pictures faster toward the end of the experiment [ $F_{(1,39)} = 105.700$ ,  $MSE = 27091$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.730$ ]. There was no significant interaction between Distractor Type and Presentation,  $F < 1$ . Therefore, the remaining analyses are presented collapsed across this factor. The main effect of Distractor Type was significant [ $F_{(4,156)} = 23.546$ ,  $MSE = 10634$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.044$ ].

In a two-ways repeated measures ANOVA (Morphological Relatedness: related vs. unrelated; Semantic Relatedness: associatively vs. categorical) using effect as dependent variable, there were significant main effects of Morphological Relatedness [ $F_{(1,39)} = 52.617$ ,  $MSE = 2384$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.157$ ] and of Semantic Relatedness [ $F_{(1,39)} = 17.935$ ,  $MSE = 1885$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.048$ ]. Overall, morphologically related distractors yielded facilitation (58 ms), but morphologically unrelated ones did not (–2 ms). Moreover, effects were larger for associatively related (44 ms) than for categorically related distractors (19 ms). The interaction was also significant [ $F_{(1,39)} = 6.810$ ,  $MSE = 2048$ ,  $p = 0.013$ ,  $\eta_g^2 = 0.020$ ]. The interaction was due to the fact that the difference between +A–M and +C–M was only 12 ms, while the difference between +A+M and +C+M was 46 ms. When distractors were morphologically related to their target, associatively related distractors facilitated naming responses more than categorically related ones. When there was no morphological relation, associatively and categorically related distractors were equally ineffective.

Mean voice key latencies were faster of both morpho-phonological conditions relative to the unrelated condition: +A+M [ $t_{(39)} = -7.303$ ,  $p < 0.001$ ] and +C+M, [ $t_{(39)} = -3.391$ ,  $p = 0.001$ ]. Furthermore, there was a significant difference between these two [ $t_{(39)} = -4.789$ ,  $p < 0.001$ ]. Associatively related distractors without morpho-phonological overlap did not differ from the unrelated condition, +A–M [ $t_{(39)} = -0.654$ ,

$p = 0.259$ ], and the same was true for category members without morpho-phonological overlap +C–M [ $t_{(39)} = 0.568$ ,  $p = 0.287$ ].

We had hypothesized that the facilitatory effect of the +A+M condition could be due to a phonological preparation effect. In Experiment 1 participants had approximately 600 ms to prepare the modifier of the compound as first part of naming the target picture. The shortened cue onset (SOA) of Experiment 2 should reduce the influence of this hypothesized effect.

We tested this in an ANOVA with the data from both experiments/SOAs. Given that the overall latencies were quite different, we first *z*-transformed the RT for each SOA (600, 200), and used the effect of the morpho-phonologically related conditions (unrelated condition – related; +A+M, +C+M, respectively) as dependent variable. The ANOVA included the factors SOA (600, 200) and Semantic Relatedness (associated, categorically related). Semantic relatedness did matter [ $F_{(1,78)} = 16.053$ ,  $MSE = 0.241$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.046$ ]. Neither SOA [ $F_{(1,78)} = 1.201$ ,  $MSE = 0.782$ ,  $p = 0.277$ ] nor the interaction [ $F_{(1,78)} = 2.083$ ,  $MSE = 0.241$ ,  $p = 0.153$ ] were significant. Thus, irrespective of SOA, given morpho-phonological overlap between distractor and target pictures, associatively related distractor pictures induced more facilitation (63 ms) than categorically related ones (28 ms). Note again that no effects were found in Experiment 2 in the absence of morpho-phonological overlap.

The eye-tracking data showed that participants fixated only one object in 51.3% of the trials (target: 48.7%, one distractor: 2.6%). Two objects were fixated in 33.9% (target and one distractor: 33.3%, both distractors: 0.3%). All three objects were looked at in 5.3% of the trials. All other fixations (9.6%) fell outside the objects (see **Table 3** for an overview). The number of gazes shows that participants looked at the target object even more often than in Experiment 1.

An ANOVA with first-fixation onset as dependent variable showed a significant effect for Distractor Type [ $F_{(4,152)} = 12.392$ ,  $MSE = 822$ ,  $p \leq 0.001$ ,  $\eta_g^2 = 0.136$ ]<sup>6</sup>. To further investigate this difference, we ran a two-ways repeated measures ANOVA, with Morphological Relatedness (related, unrelated) and Semantic Relatedness (associatively, categorically related) as factors. Targets attracted faster fixation onsets in the presence of morpho-phonologically related distractors than with unrelated ones [ $F_{(1,38)} = 33.566$ ,  $MSE = 919$ ,  $p \leq 0.001$ ,  $\eta_g^2 = 0.136$ +M: 422 ms; –M: 450 ms]. Overall, targets were looked at faster in the presence of associatively related than with categorically related distractors [ $F_{(1,38)} = 5.094$ ,  $MSE = 573$ ,  $p = 0.03$ ,  $\eta_g^2 = 0.015$ ; +A: 432 ms, +C: 441 ms]. A marginally significant interaction qualified the main effects [ $F_{(1,38)} = 3.675$ ,  $MSE = 822$ ,  $p = 0.063$ ,  $\eta_g^2 = 0.015$ ; +A+M: 414 ms, +C+M: 431 ms, +A–M: 451 ms, +C–M: 450 ms]. The interaction showed that the faster fixations to targets in the presence of associatively related distractors only held when the picture names were morpho-phonologically related, not in the absence of morphological relatedness.

The one-way repeated measure ANOVA with the within factor Distractor Type on dwell-time was also significant [ $F_{(4,152)} = 2.477$ ,  $MSE = 6475$ ,  $p = 0.047$ ,  $\eta_g^2 = 0.005$ ].

<sup>6</sup>One participant was excluded due to the number of outliers.

Again, we followed this analysis by the same two-ways repeated measures ANOVA as reported before. Only the factor Semantic Relatedness was significant [ $F_{(1,38)} = 10.326$ ,  $MSE = 3876$ ,  $p = 0.003$ ,  $\eta_g^2 = 0.013$ ]. Associatively related distractors induced shorter dwell times on the targets than categorically related ones. The factor Morphological Relatedness and the interaction were not significant [ $F_{(1,38)} \leq 1.424$ ,  $p \leq 0.240$ ].

## Discussion

Experiment 2 showed similar effects as Experiment 1. Morpho-phonological relatedness between target and distractor pictures (+A+M, +C+M) facilitated picture naming, relative to an unrelated baseline and to morpho-phonologically unrelated distractors. The observed effects do not seem to originate from preparation of the first morpheme shared by distractors and target in the +Ä+M condition, as corroborated by the lack of interaction in the analysis on the data from both SOAs. Different from Experiment 1, associatively related distractors did not facilitate target picture naming when they were morpho-phonologically unrelated. The shortened SOA and/or the absence of a morpho-phonological association may have prevented the full build-up of positive associative as well as negative categorical effects, but apparently did not prevent access to lexical information for the pictures. Numerically, effects of morpho-phonological similarity between the names of distractor and target pictures were even stronger than in Experiment 1. Thus, manipulating SOA seems to differentiate between the strength of semantic-conceptual (associative and categorical) and lexical (presence or absence of morpho-phonological influences) effects. Importantly, as in Experiment 1, the data provide evidence for full cascading to the word-form level, and the substantial difference between effects of associatively related (44 ms) and categorically related distractors (19 ms) at least is compatible with interference due to lexical competition.

In Experiment 3, we changed the way in which it was signaled which picture was the target for naming. Glaser and Glaser (1989) asked their participants to name the first (or the second) picture that appeared on the screen (see also La Heij et al., 2003). We adapted this procedure and signaled the target by a later onset. We wanted to give the distractor pictures a head start, attracting attention by means of their visual onset, to allow for a full impact of conceptual/semantic effects (we argued that it is impossible to avoid semantic processing of visual stimuli). Given that it is clear that the distractor pictures always come first, they may well not be processed lexically at all, because they should not be named. If this holds, there should be no impact of morpho-phonological relatedness, or of lexical competition. Thus, with this presentation manipulation, we investigated whether lexical processing of stimuli that do not have to be named is mandatory, and if so, up to which level. Another important motivation for Experiment 3 is to assess potential differences in the strength of the semantic relation between target and distractor pictures in the four conditions. Although the mean semantic-relatedness judgments (see **Table 1**) did not differ, the small differences between the means could have an impact when online priming effects

are concerned. If the data from Experiment 3 show pure semantic effects, without any lexical competition or morpho-phonological involvement, the priming by the four distractor conditions would be purely conceptual and could be compared directly.

## Experiment 3: Target 200 ms after Distractors

In this experiment, we altered the way the target picture was signaled. In Experiments 1 and 2, we used an arrow that appeared some time after the simultaneous onset of all three pictures, to indicate the target picture. In Experiment 3, the target picture appeared 200 ms after the onset of the distractor pictures. This provides some time for the processing of the distractors, and gives them a head start. This SOA also roughly corresponds to SOAs used in PWI experiments to evoke semantic effects – but note that the processing flow differs for pictures and words. Most importantly, we reasoned that this timing would give rise to positive conceptual-semantic effects, perhaps to competition effects, but not to word-form effects. Positive semantic effects of associative and categorical distractors should be evident because the earlier distractor onset allows activating the relevant conceptual network (Abdel Rahman and Melinger, 2009).

## Method

### Participants

Twenty participants selected from the same population as before were tested in this experiment. None had participated in Experiments 1 and 2.

### Procedure

The same material and apparatus as in Experiments 1 and 2 was used. The difference to the previous experiment was that the target picture appeared 200 ms after distractor-picture onset. Furthermore, we changed the filler conditions. In 12 of the 24 fillers, distractor pictures were replaced by pictures with morpho-phonological overlap either in the first (6) or second constituent (6) of the other distractor picture. Note that the target picture was never morpho-phonologically related to these distractor pictures. However, given the different timing of distractors and target, we wanted to counteract strategic processing induced by the distractor pair (i.e., whenever there is morpho-phonological overlap in the first or second constituent of the distractor pictures, the target picture shares this constituent). The target pictures in the filler condition also had different distractor pictures in each of the five presentations. Additionally, the distractor pictures without morpho-phonological overlap in the filler condition were randomized further within themselves. All other aspects of the procedure remained the same.

## Results

Responses different from expected names (1%), disfluencies (1.3%), voice key failures (0.4%), time-outs (1.8%) and reaction before target onset (0.1%) were excluded from the analyses. No

item set or participant was excluded from the analyses. Voice-key latencies measured from target picture onset were averaged over participants and submitted to an ANOVA. **Table 2** lists RT and standard errors as a function of Distractor Type. Participants named pictures faster toward the end of the experiment, indicated by a significant linear trend [ $F_{(1,19)} = 23.823$ ,  $MSE = 26969$ ,  $p = 0.001$ ,  $\eta_g^2 = 0.556$ ]. There was no significant interaction between Distractor Type and Presentation ( $F < 1$ ). Therefore, the remaining analyses were collapsed across this factor. There was a main effect for the factor Distractor Type [ $F_{(4,76)} = 5.529$ ,  $MSE = 1454$ ,  $p = 0.001$ ,  $\eta_g^2 = 0.030$ ].

Using effect as dependent variable, the main effect of Morphological Relatedness ( $F < 1$ ) was not significant nor the interaction were significant in the two-ways repeated measures ANOVA. The main effect of Semantic Relatedness was marginally significant [ $F_{(1,19)} = 3.154$ ,  $MSE = 2240$ ,  $p = 0.092$ ,  $\eta_g^2 = 0.030$ ].

The significant main effect for the factor Distractor Type was further analyzed using paired one-sided *t*-tests and averaged voice key latencies as dependent variable. Participants were faster in naming the target picture relative to the unrelated condition in the associatively and morpho-phonologically related condition, +A+M [ $t_{(19)} = -4.067$ ,  $p < 0.001$ ] and in the categorically and morpho-phonologically related condition, +C+M [ $t_{(19)} = -2.465$ ,  $p = 0.012$ ]. Facilitation was also significant for both conditions without morphological overlap: +A–M [ $t_{(19)} = -3.644$ ,  $p = 0.001$ ], +C–M [ $t_{(19)} = -2.535$ ,  $p = 0.010$ ].

The eye-tracking data analyzed after target onset showed that participants fixated one object in 57.1% of the trials (target: 56.3%, one distractor: 0.8%). Two objects were fixated in 32.2% (target and one distractor: 22.4%, both distractors: 0.1%). All three objects were looked at in 5.0% of the trials. All other fixations (6.1%) fell outside the objects (see **Table 3** for an overview). Interestingly, the eye-tracking data before target onset demonstrated that both distractors were only fixated in 0.1% of the cases, whereas one of the two distractors was looked at in 36.6%. All other fixations fell outside of the objects (63.3%). First fixation onsets as well as dwell-times did not differ from each other ( $F < 1$ ).

## Discussion

Consistent with our expectation, we observe facilitation in all conditions relative to the unrelated baseline: for both associatively (+A+M = 53 ms, +A–M = 45 ms) and both categorically related distractors (+C+M = 29 ms, +C–M = 28 ms), irrespective of morpho-phonological relatedness. Different from Experiments 1 and 2, category members reliably facilitated picture naming in a picture–picture setting. Thus, we replicate earlier findings that an onset manipulation gives rise to semantic effects (see Glaser and Glaser, 1989; La Heij et al., 2003, for inhibitory effects).

The results from Experiment 3 indicate that distractor pictures were not processed lexically. The effects of the two semantic conditions are statistically the same: both speed up target naming, most probably due to semantic priming through spreading of activation at the conceptual level. There is no evidence for lexical competition with distractor pictures from the same semantic category; both conditions induce significant semantic priming.

The effect in the associative conditions is numerically larger (20 ms) than in the categorical conditions, but note that the main effect of Semantic Relatedness failed significance. The numerical difference might reflect the somewhat larger semantic relatedness scores from the pretest (mean semantic relatedness rating: associatively related: 3.65 vs. categorically related: 3.20). In the absence of lexical competition effects, it is not surprising that distractors are not processed all the way down to the word-form level. Otherwise, we should have observed an additional effect of morpho-phonological overlap, which has proven to be facilitatory over a wide range of material and tasks (Roelofs and Baayen, 2002; Gumnior et al., 2006; Lüttmann et al., 2011a,b), as well as in Experiments 1 and 2.

## General Discussion

This study aimed to test two competing types of model of speech production: the two-step discrete serial model (Levett et al., 1999) and models that allow full cascading of lexical information for multiple concepts, all the way down to word-form and phonological levels (Caramazza, 1997; Peterson and Savoy, 1998). An additional aim was to test differing explanations for semantic interference from word distractors on target processing, which can be ideally tested with picture distractors. We found encoding of distractor names up to the form level, which supports cascading rather than serial models. We also observed interference from distractors on target naming which is not predicted by a post-lexical response buffer explanation (Mahon et al., 2007).

To test the predictions of the competing models, we focused on potential facilitation and inhibition effects of different types of semantic similarity, and on the morpho-phonological relation between distractor and target pictures in picture naming. We manipulated semantic (categorical or associative) relatedness, crossed with morpho-phonological overlap (present or absent). We did so to assess the level up to which distractor pictures, whose names are *not* produced, are lexically encoded. Furthermore, the distractor pictures and the target picture appeared simultaneously or staggered, with the distractor pictures preceding the target picture. In case of simultaneous presentation, the cue signaling the target appeared at different moments in time. The target and cue onset manipulations served to gain insight in the temporal aspects of distractor processing.

Our versions of the picture–picture paradigm show that overlapping pictures and color as signal, as used by Morsella and Miozzo (2002), are not necessary to evoke effects. Varying temporal onsets of distractor and target, or signaling the target by means of a cue, are both effective manipulations and reveal semantic effects that were not observed with the color manipulation (see also Glaser and Glaser, 1989; La Heij et al., 2003). Moreover, it was not necessary to focus attention to the distractor pictures by spatial cueing, as done by Jescheniak et al. (2014) to induce lexical competition by distractor pictures. We simply presented the distractors first, and their onset was enough to attract attention and induce semantic processing. In addition, the cue technique induced form effects from distractors pictures

that shared a morpheme with the target picture (Experiments 1 and 2).

Before moving to the effects observed and their relevance for the predictions made by discrete and cascaded models, we discuss the eye-tracking data. Effects of the experimental manipulations in the eye-tracking data emerged only in Experiment 2, where morphologically related distractors accelerated first-fixation onsets, and associatively related distractors were fixated faster than categorically related ones. Dwell times partly mirrored the pattern of the first-fixation onsets, but without effects of morphological relatedness. One might conclude that the implemented experimental situations were not demanding enough to affect eye-movements, although they effectively affected word-production processes. The role of eye-movements is less well understood in word production than in reading (Rayner, 2009). Meyer et al. (1998) suggested that eye-movements reflect the timing of word-production processes in multi-word utterances. However, this is not mandatory, because speakers can deviate from the observed coupling of eye-movements and word production when the task is easy (Meyer et al., 2012).

In our experiments, eye movements were tracked to investigate whether target fixations are mandatory for accurate naming, and whether fixations on distractor pictures are necessary for effects to emerge. The answer to both seems to be no. In Experiment 1, in approximately 28% of the cases the target was not fixated at all, but naming was very accurate indeed (~94%). Moreover, distractors were rarely fixated alone (9% compared to targets alone: 29%). In about half of the trials (48%) no distractor was looked at but we still get clear effects of distractors on target naming. Overt attentional shifts to distractors, as indicated by eye-movements, are thus not required for their lexical encoding. This replicates our findings with scene stimuli (Dobel et al., 2007).

### Semantic Relatedness

Discrete two-step models implement two lexical levels, lemmas that code syntactic information, and lexemes or word-forms that code morphological and phonological information. Such models allow for the activation of multiple lemmas at the first level, but – with few exceptions – not of multiple word forms.

In two experiments (Experiments 1 and 3) associatively related distractor pictures accelerated target picture naming, even without morpho-phonological similarity. Thus, related concepts such as *lawn mower* and *swimming pool* facilitate the naming of the picture of a *garden chair*. In Experiment 2, with a shorter target-cue onset, facilitation emerged only when distractors (*garden hose*; *rocking chair*) and target (*garden chair*) shared a morpheme. If pictures all belong to – very loosely speaking – the same semantic field, their concepts seem to activate each other, which speeds up conceptual processing and target picture naming.

When the target picture was signaled by means of a cue, categorically related distractors induced neither facilitation nor interference, relative to the unrelated baseline. But facilitation due to categorically related distractors (e.g., *kitchen table* and

*shoe rack*) was only observed when distractor pictures and target picture appeared at different moments in time (Experiment 3). This seems at odds with results by Glaser and Glaser (1989) and La Heij et al. (2003, Experiment 1). Glaser and Glaser (1989) observed interference and argued that this is because distractor and target activate closely related semantic representations. La Heij et al. (2003) proposed that effects observed by Glaser and Glaser (1989) were due to the erroneous selection of distractors as target. Moreover, Glaser and Glaser (1989) used just nine pictures as target and context pictures. When La Heij et al. (2003) reduced distractor-presentation duration from 300 to 50 ms and increased the number of target pictures from 9 to 40, they observed facilitation. We used longer distractor presentation durations than La Heij et al. (2003), had a smaller number of target pictures, but observed facilitation nonetheless (Experiment 3). Thus, it is most likely that neither the number of target pictures nor the distraction presentation duration is the crucial manipulation. Observing semantic facilitation, or interference, rather depends on the ease of target identification. When the target is clearly signaled and distractor pictures are not used as targets in the experiment, effects are facilitatory. In these cases, distractors do not seem to enter the lexical system (as in Experiment 3). If there is (temporal) uncertainty as to which picture is going to be the target, lexical access is initiated for all pictured concepts, rendering lexical selection of the target more difficult when the distractors come from the same semantic category (Glaser and Glaser, 1989; La Heij et al., 2003). This is what we observed in Experiments 1 and 2. We feel that the situation of uncertainty, which concept to express, which environmental stimulus to name, is rule rather than exception during speaking. This is clearly reflected in the fact that all models of speech production adhere to the activation of multiple conceptual representations, and all models allow these non-linguistic representations to activate linguistic ones. As such, certainty as to which pictures are targets for naming and which not (Experiment 3) is the exception, rather than the rule.

A next question is, how “categorical” facilitation (Experiment 3) occurs. As our data show: in a similar manner to associative facilitation. Related concepts activate each other, speeding up target processing at conceptual and subsequent stages, even all the way down to the vocal response. The result challenges the assumptions made by Levelt and Colleagues Roelofs (1992) and Levelt et al. (1999), who claim that conceptual activation always results in the activation of multiple lemmas, which compete for selection. The data from Experiment 3 show that categorically related distractor pictures did activate their conceptual-semantic information, which apparently was not fed forward into the lexical stratum, because we observed no interference. Our data suggest that lexical processing of activated concepts is not mandatory. When there is no uncertainty as to which picture has to be named, the distractors, although activated at the conceptual level, do not enter the lexical system. This in fact also fits with Roelofs (2008a), who argued that task demands determine the presence and direction of semantic effects. When target selection is easy, facilitation occurs, while in case of a difficult selection, inhibition is observed.

When there is uncertainty as to which target should be named (Experiments 1 and 2), we do indeed observe inhibition from distractor pictures that share their semantic category with the target – relative to associatively related distractors. This is evidence for lexical competition and selection (Roelofs, 1992). Lexical selection by competition is not implemented in full cascading models such as the one proposed by Caramazza (1997), Finkbeiner and Caramazza (2006), and Mahon et al. (2007). Consequently, they proposed a different locus for the interference from categorically related distractor words regularly observed in picture naming: the post-lexical response buffer. Given that our interference comes from pictures, not from words, this refutes the response-buffer explanation, because only word distractors can cause havoc in the response buffer, the locus of the interference in their model. (Mahon et al., 2007).

### Morphological Relatedness

In Experiments 1 and 2, the morphological relation between distractors and target modulated the effects obtained for semantic relatedness. This finding shows that distractor pictures were processed all the way “down” to the form level. These data do not agree with results by Meyer et al. (1998), who found separate and sequential processing, from meaning to phonology, for two simultaneously displayed pictures that both had to be named. But note that this is a special situation (Meyer et al., 2012). Importantly, our results argue against views that allow no multiple lexical activation at all (Bloem and La Heij, 2003), and against partially cascaded models that only allow a “limited” flow of activation from conceptual to form representations (Levelt et al., 1991, 1999; Roelofs, 1992, 2003).

According to all models, the conceptual level is “blind” to the morpho-phonological structure of the word belonging to a concept. To observe morpho-phonological facilitation, all concepts must have been processed to a level at which this information is represented, and this must be the lexeme or word-form level. Obviously, the semantic cohort of the target that is set up upon lexical access will incorporate morphologically related words, given that these often are semantically related. However, it has been shown in a variety of tasks that semantic and morphological effects in speech production reflect processing

at different levels, and that morphological similarity without semantic relatedness (as with the “hummingbird” and the “jailbird”) is (almost) as effective as with semantic similarity (“hummingbird” and “blackbird”; e.g., Dohmes et al., 2004; Koester and Schiller, 2008, 2011; Lüttmann et al., 2011b).

Taken together, the morpho-phonological effects fit with fully cascaded models of speech production (Stemberger, 1985; Dell, 1986; Caramazza, 1997; Peterson and Savoy, 1998), in which it is assumed that activation flows continuously from high levels to lower levels. Note that the evidence for competitive lexical selection, based on the nature of the semantic relation between distractor and target pictures, is a challenge to some of these models. The fact that the influence of morpho-phonological relatedness varies from experiment to experiment, and is even absent when there is no uncertainty about the target, suggests that the flow of information depends on specific characteristics of the speaking situation (Roelofs, 2008a). Roelofs and Piai (2011) as well as O’Séaghdha and Frazer (2014) argue that the degree of phonological activation depends on the available attentional capacity. Thus, it is an assignment for any model of speech production to adjust the claims about their basic structure to the requirements of the task, the speaking situation and to the amount of attention paid to stimuli in the environment.

### Ethical Statement

This study was carried out in accordance with the recommendations of ethical guidelines of the Institute for Psychology, Westfälische Wilhelms-Universität, Münster, Germany, with written informed consent from all participants. All participants gave written informed consent in accordance with the Declaration of Helsinki.

### Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01540>

### References

- Abdel Rahman, R., and Melinger, A. (2009). Semantic context effects in language production: a swinging lexical network proposal and a review. *Lang. Cogn. Process.* 24, 713–734. doi: 10.1080/01690960802597250
- Alario, F., Segui, J., and Ferrand, L. (2000). Semantic and associative priming in picture naming. *Q. J. Exp. Psychol. A* 53, 741–764. doi: 10.1080/027249800410535
- Aristei, S., Zwitserlood, P., and Abdel Rahman, R. (2012). Picture-induced semantic interference reflects lexical competition during object naming. *Front. Psychol.* 3:28. doi: 10.3389/fpsyg.2012.00028.
- Baayen, R. H., Piepenbrock, R., and Van Rijn, H. (1993). *The CELEX Lexical Database*. Philadelphia, PA: CD-ROM.
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Belke, E. (2005). Lexical access revisited: evidence from picture-pseudoword interference. *Eur. J. Cogn. Psychol.* 17, 117–150. doi: 10.1080/09541440340000448
- Bloem, I., and La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: implications for models of lexical access in language production. *J. Mem. Lang.* 48, 468–488. doi: 10.1016/S0749-596X(02)0503-X
- Bölte, J., Böhl, A., Dobel, C., and Zwitserlood, P. (2009). Effects of referential ambiguity, time constraints and addressee orientation on the production of morphologically complex words. *J. Cogn. Psychol.* 21, 1166–1199. doi: 10.1080/09541440902719025
- Bölte, J., Jorschick, A., and Zwitserlood, P. (2003). “Reading yellow speeds up naming a picture of a banana: facilitation and inhibition in picture-word interference,” in *Proceedings European Cognitive Science Conference*, eds F. Schmalhofer and R. Young (Mahwah, NJ: Lawrence Erlbaum Associates Inc.), 55–60.

- Bölte, J., Jorschick, A., and Zwitserlood, P. (2005). Categorical inhibition and associative facilitation in language production. *Paper Presented at the 46th Annual Meeting of the Psychonomic Society*, Toronto, ON.
- Butterworth, B. (1989). “Lexical access in speech production,” in *Lexical Representations and Process*, ed. W. Marslen-Wilson (Cambridge, MA: MIT Press), 108–135.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208. doi: 10.1080/026432997381664
- Caramazza, A., Laudanna, A., and Romani, C. (1988). Lexical access and inflectional morphology. *Cognition* 28, 297–332. doi: 10.1016/0010-0277(88)90017-0
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *J. Verbal Learn. Verbal Behav.* 12, 335–359. doi: 10.1016/S0022-5371(73)80014-3
- Costa, A., Alario, F.-X., and Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychon. Bull. Rev.* 12, 125–131. doi: 10.3758/BF03196357
- Costa, A., Caramazza, A., and Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol.* 26, 1283–1296. doi: 10.1037/0278-7393.26.5.1283
- Damian, M. F., and Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychon. Bull. Rev.* 10, 111–117. doi: 10.3758/BF03196474
- Damian, M. F., and Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *J. Exp. Psychol.* 25, 345–361. doi: 10.1037/0278-7393.25.2.345
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283
- Dell, G. S., and Reich, P. A. (1981). Stages in sentence production: an analysis of speech error data. *J. Verbal Learn. Verbal Behav.* 20, 611–629. doi: 10.1016/S0022-5371(81)90202-4
- Dobel, C., Gumnior, H., Bölte, J., and Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychol.* 125, 129–143. doi: 10.1016/j.actpsy.2006.07.004
- Dohmes, P., Zwitserlood, P., and Bölte, J. (2004). The impact of semantic transparency of morphologically complex words on picture naming. *Brain Lang.* 90, 203–212. doi: 10.1016/S0093-934X(03)00433-4
- Eyelink II (2004). [Apparatus, and Software]. Mississauga, ON: SR Research.
- Fei-Fei, L., VanRullen, R., Koch, C., and Perona, P. (2005). Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Vis. Cogn.* 12, 893–924. doi: 10.1080/13506280444000571
- Finkbeiner, M., and Caramazza, A. (2006). Lexical selection is not a competitive process: a reply to La Heij et al. (2006). *Cortex* 42, 1032–1035. doi: 10.1016/S0010-9452(08)70210-7
- Garrett, M. F. (1980). “Levels of processing in sentence production,” in *Language Production Speech and Talk*, Vol. 1, ed. B. Butterworth (London: Academic Press).
- Glaser, W. R., and Dünghoff, F. (1984). The time course of picture-word interference. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 640–654. doi: 10.1037/0096-1523.10.5.640
- Glaser, W. R., and Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *J. Exp. Psychol.* 118, 13–42. doi: 10.1037/0096-3445.118.1.13
- Goldrick, M. (2006). Limited interaction in speech production: chronometric, speech error, and neuropsychological evidence. *Lang. Cogn. Process.* 21, 817–855. doi: 10.1037/0096-3445.21.11.817
- Gumnior, H., Bölte, J., and Zwitserlood, P. (2006). A chatterbox is a box: morphology in German word production. *Lang. Cogn. Process.* 21, 920–944. doi: 10.1080/016909600824278
- Hemera Photo Objects (n.d.). *Volume I – III. [Computer Software]*. Québec: Hemera Publisher.
- Humphreys, G. W., Lloyd-Jones, T. J., and Fias, W. (1995). Semantic interference effects on naming using a postcue procedure: tapping the links between semantics and phonology with pictures and words. *J. Exp. Psychol.* 21, 961–980. doi: 10.1037/0278-7393.21.4.961
- Humphreys, G. W., Riddoch, M., and Quinlan, P. T. (1988). Cascade processes in picture identification. *Cogn. Neuropsychol.* 5, 67–104. doi: 10.1080/02643298808252927
- Janssen, N. (2013). Response exclusion in word–word tasks: a comment on Roelofs, Piai and Schriefers. *Lang. Cogn. Process.* 28, 672–678. doi: 10.1080/01690965.2012.746715
- Janssen, N., Bi, Y., and Caramazza, A. (2008). A tale of two frequencies: determining the speed of lexical access for Mandarin Chinese and English compounds. *Lang. Cogn. Process.* 23, 1191–1223. doi: 10.1080/01690960802250900
- Jescheniak, J. D., Hantsch, A., and Schriefers, H. (2005). Context effects on lexical choice and lexical activation. *J. Exp. Psychol.* 31, 905–920. doi: 10.1037/0278-7393.31.5.905
- Jescheniak, J. D., Matushanskaya, A., Mädebach, A., and Müller, M. M. (2014). Interference from semantically related context pictures in single picture naming – Evidence for competitive lexical selection. *Psychon. Bull. Rev.* 21, 1294–1300. doi: 10.3758/s13423-014-0606-5
- Jescheniak, J. D., Oppermann, F., Hantsch, A., Wagner, V., Mädebach, A., and Schriefers, H. (2009). Do perceived context pictures automatically activate their phonological code? *Exp. Psychol.* 56, 56–65. doi: 10.1027/1618-3169.56.1.56
- Jescheniak, J. D., and Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: further evidence from the coactivation of near-synonyms. *J. Exp. Psychol.* 24, 1256–1274. doi: 10.1037/0278-7393.24.5.1256
- Koester, D., and Schiller, N. O. (2008). Morphological priming in overt language production: electrophysiological evidence from Dutch. *Neuroimage* 42, 1622–1630. doi: 10.1016/j.neuroimage.2008.06.043
- Koester, D., and Schiller, N. O. (2011). The functional neuroanatomy of morphology in language production. *Neuroimage* 55, 732–741. doi: 10.1016/j.neuroimage.2010.11.044
- La Heij, W., Dirx, J., and Kramer, P. (1990). Categorical interference and associative priming in picture naming. *Br. J. Psychol.* 81, 511–525. doi: 10.1111/j.2044-8295.1990.tb02376.x
- La Heij, W., Heikoop, K. W., Akerboom, S., and Bloem, I. (2003). Picture naming in picture context: semantic interference or semantic facilitation? *Psychol. Sci.* 45, 49–62.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. doi: 10.1017/S0140525X99001776
- Levelt, W. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., and Havinga, J. (1991). The time course of lexical access in speech production: a study of picture naming. *Psychol. Rev.* 98, 122–142. doi: 10.1037/0033-295X.98.1.122
- Lüttmann, H., Zwitserlood, P., Böhl, A., and Bölte, J. (2011a). Evidence for morphological composition at the form level in speech production. *J. Cogn. Psychol.* 23, 818–836. doi: 10.1080/20445911.2011.575774
- Lüttmann, H., Zwitserlood, P., and Bölte, J. (2011b). Sharing morphemes without sharing meaning: production and comprehension of German verbs in the context of morphological relatives. *Can. J. Exp. Psychol.* 65, 173–191. doi: 10.1037/a0023794
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., and Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol.* 33, 503–535. doi: 10.1037/0278-7393.33.3.503
- Mahon, B. Z., and Navarrete, E. (2014). The Critical difference in models of speech production: a response to Roelofs and Piai. *Cortex* 52, 123–127. doi: 10.1016/j.cortex.2013.12.001
- Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S., and Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Lang. Cogn. Process.* 11, 257–282. doi: 10.1080/016909696387187
- Melinger, A., and Abdel Rahman, R. (2004). Investigating the interplay between semantic and phonological distractor effects in picture naming. *Brain Lang.* 90, 213–220. doi: 10.1016/S0093-934X(03)00434-6
- Meyer, A. S., and Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Mem. Cogn.* 35, 494–503. doi: 10.3758/BF03193289

- Meyer, A. S., and Schriefers, H. H. (1991). Phonological facilitation in picture-word interference experiments: effects of stimulus onset asynchrony and types of interfering stimuli. *J. Exp. Psychol.* 17, 1146–1160. doi: 10.1037/0278-7393.17.6.1146
- Meyer, A. S., Sleiderink, A. M., and Levelt, W. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition* 66, B25–B33. doi: 10.1016/S0010-0277(98)00009-2
- Meyer, A. S., Wheeldon, L., van der Meulen, F., and Konopka, A. (2012). Effects of speech rate and practice on the allocation of visual attention in multiple object naming. *Front. Psychol.* 3:39. doi: 10.3389/fpsyg.2012.00039
- Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutor. Q. Methods Psychol.* 4, 61–64.
- Morsella, E., and Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *J. Exp. Psychol.* 28, 555–563. doi: 10.1037/0278-7393.28.3.555
- Navarrete, E., and Costa, A. (2005). Phonological activation of ignored pictures: further evidence for a cascade model of lexical access. *J. Mem. Lang.* 53, 359–377. doi: 10.1016/j.jml.2005.05.001
- Oppermann, F., Jescheniak, J. D., and Görge, F. (2014). Resolving competition when naming an object in a multiple-object display. *Psychon. Bull. Rev.* 21, 78–84. doi: 10.3758/s13423-013-0465-5
- Oppermann, F., Jescheniak, J. D., and Schriefers, H. (2008). Conceptual coherence affects phonological activation of context objects during object naming. *J. Exp. Psychol.* 34, 587–601. doi: 10.1037/0278-7393.34.3.587
- O’Séaghda, P. G., and Frazer, A. K. (2014). The exception does not rule: attention constrains form preparation in word production. *J. Exp. Psychol.* 40, 797–810. doi: 10.1037/a0035576
- Peterson, R. R., and Savoy, P. (1998). Lexical selection and phonological encoding during language production: evidence for cascaded processing. *J. Exp. Psychol.* 24, 539–557. doi: 10.1037/0278-7393.24.3.539
- Raaijmakers, J. W., Schrijnemakers, J. C., and Gremmen, F. (1999). How to deal with The language-as-fixed-effect fallacy: common misconceptions and alternative solutions. *J. Mem. Lang.* 41, 416–426. doi: 10.1006/jmla.1999.2650
- Rapp, B., and Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychol. Rev.* 107, 460–499. doi: 10.1037/0033-295X.107.3.460
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86, 446–461. doi: 10.1037/0033-2909.86.3.446
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62, 1457–1506. doi: 10.1080/17470210902816461
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition* 64, 249–284. doi: 10.1016/S0010-0277(97)00027-9
- Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychol. Rev.* 110, 88–125. doi: 10.1037/0033-295X.110.1.88
- Roelofs, A. (2004). Comprehension-based versus production-internal feedback in planning spoken words: a Rejoinder to Rapp and Goldrick (2004). *Psychol. Rev.* 111, 579–580. doi: 10.1037/0033-295X.111.2.579
- Roelofs, A. (2008a). Tracing attention and the activation flow in spoken word planning using eye movements. *J. Exp. Psychol.* 34, 353–368. doi: 10.1037/0278-7393.34.2.353
- Roelofs, A. (2008b). Dynamics of the attentional control of word retrieval: analyses of response time distributions. *J. Exp. Psychol.* 137, 303–323. doi: 10.1037/0096-3445.137.2.303
- Roelofs, A., and Baayen, H. (2002). Morphology by itself in planning the production of spoken words. *Psychon. Bull. Rev.* 9, 132–138. doi: 10.3758/BF03196269
- Roelofs, A., and Piai, V. (2011). Attention demands of spoken word planning: a review. *Front. Psychol.* 2:307. doi: 10.3389/fpsyg.2011.00307
- Roelofs, A., Piai, V., and Schriefers, H. (2013). Selection by competition in word production: rejoinder to Janssen (2013). *Lang. Cogn. Process.* 28, 679–683. doi: 10.1080/01690965.2013.770890
- Schriefers, H., Meyer, A. S., and Levelt, W. J. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *J. Mem. Lang.* 29, 86–102. doi: 10.1016/0749-596X(90)90011-N
- Starreveld, P. A., and La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *J. Exp. Psychol.* 21, 686–698. doi: 10.1037/0278-7393.21.3.686
- Starreveld, P. A., and La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *J. Exp. Psychol.* 22, 896–918. doi: 10.1037/0278-7393.22.4.896
- Stemberger, P. (1985). “An interactive activation model of language production,” in *Progress in the Psychology of Language*, Vol. 1, ed. A. W. Ellis (Hillsdale, NJ: Lawrence Erlbaum Associates Inc.), 143–186.
- Zwitzerlood, P. (1994). “Access to phonological form representations in language comprehension and production,” in *Perspectives on Sentence Processing*, eds C. Clifton, L. Frazier, and K. Rayner (Hillsdale, NY: Lawrence Erlbaum).
- Zwitzerlood, P., Bölte, J., and Dohmes, P. (2002). Where and how morphologically complex words interplay with naming pictures. *Brain Lang.* 81, 358–367. doi: 10.1006/brln.2001.2530

**Conflict of Interest Statement:** The reviewer, Dr. Markus Damian, declares that despite having collaborated with the author Dr. Jens Bölte, the review process was handled objectively. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Bölte, Böhl, Dobel and Zwitzerlood. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The roles of shared vs. distinctive conceptual features in lexical access

Harrison E. Vieth<sup>1\*</sup>, Katie L. McMahon<sup>2</sup> and Greig I. de Zubicaray<sup>1</sup>

<sup>1</sup> School of Psychology, University of Queensland, Brisbane, QLD, Australia

<sup>2</sup> Centre for Advanced Imaging, University of Queensland, Brisbane, QLD, Australia

## Edited by:

Peter Indefrey, University of  
Dusseldorf, Germany

## Reviewed by:

Francesca Peressotti, University of  
Padova, Italy  
Katharina Spalek,  
Humboldt-Universität zu Berlin,  
Germany

## \*Correspondence:

Harrison E. Vieth, School of  
Psychology, University of  
Queensland, Building 24, Brisbane,  
QLD 4072, Australia  
e-mail: harrison.vieth@  
uqconnect.edu.au

Contemporary models of spoken word production assume conceptual feature sharing determines the speed with which objects are named in categorically-related contexts. However, statistical models of concept representation have also identified a role for *feature distinctiveness*, i.e., features that identify a single concept and serve to distinguish it quickly from other similar concepts. In three experiments we investigated whether distinctive features might explain reports of counter-intuitive semantic facilitation effects in the picture word interference (PWI) paradigm. In Experiment 1, categorically-related distractors matched in terms of semantic similarity ratings (e.g., *zebra* and *pony*) and manipulated with respect to feature distinctiveness (e.g., a *zebra* has stripes unlike other equine species) elicited interference effects of comparable magnitude. Experiments 2 and 3 investigated the role of feature distinctiveness with respect to reports of facilitated naming with part-whole distractor-target relations (e.g., a *hump* is a distinguishing part of a CAMEL, whereas *knee* is not, vs. an unrelated part such as *plug*). Related part distractors did not influence target picture naming latencies significantly when the part denoted by the related distractor was not visible in the target picture (whether distinctive or not; Experiment 2). When the part denoted by the related distractor was visible in the target picture, non-distinctive part distractors slowed target naming significantly at SOA of –150 ms (Experiment 3). Thus, our results show that semantic interference does occur for part-whole distractor-target relations in PWI, but only when distractors denote features shared with the target and other category exemplars. We discuss the implications of these results for some recently developed, novel accounts of lexical access in spoken word production.

**Keywords:** lexical access, competition, semantic interference, picture naming, shared features, distinctive features

## INTRODUCTION

A large empirical literature on object naming has demonstrated that speakers are influenced by the activation of concepts related to the object they intend to name. For example, when objects are presented in categorically related vs. unrelated contexts, naming latencies are typically slower (e.g., Rosinski, 1977; Lupker, 1979; Kroll and Stewart, 1994). Virtually all accounts of spoken word production assume that these semantic context effects occur due to the co-activation of conceptual features shared among categorically related objects. However, there is considerable disagreement among accounts as to the consequences of this *conceptual feature overlap* for the production system (e.g., Dell and O'Seaghdha, 1992; Caramazza, 1997; Levelt et al., 1999; Goldrick and Rapp, 2002; Mahon et al., 2007; Rahman and Melinger, 2009).

Semantic context effects are induced successfully in a number of experimental naming paradigms. For example, in the picture-word interference (PWI) paradigm, in which participants ignore a distractor word while naming a target picture, slower naming latencies are observed reliably when distractors (e.g., *wolf*) are category coordinates of the target picture (e.g., DOG) compared to unrelated distractors (e.g., *book*; Schriefers et al., 1990; Levelt

et al., 1991; La Heij and van den Hof, 1995). This effect is known as *semantic interference* and has been interpreted as evidence supporting a competitive lexical selection mechanism in some spoken word production models (Starreveld and La Heij, 1996; Levelt et al., 1999; Rahman and Melinger, 2009). However, non-competitive lexical selection mechanisms have also been proposed to explain the effect (Caramazza, 1997; Mahon et al., 2007).

The *lexical selection by competition* account assumes that naming latencies are a function of the number of active lexical candidates and their activation levels. For instance, if the target concept "HORSE" is activated, related animal category concepts such as *pony*, *cow* etc. also become activated due to conceptual feature overlap, and this activation spreads to their lexical representations (e.g., Collins and Loftus, 1975). This account explains the semantic interference effect in the PWI paradigm in terms of the categorically related distractor increasing the activation level of an existing lexical competitor, slowing target selection compared to an unrelated distractor word that activates a concept that was not activated by the target picture.

Some PWI studies have demonstrated that conceptual feature overlap might not necessarily induce semantic interference.

Costa et al. (2005) reported that naming latencies were *facilitated* using “part-whole” distractor-target pairs (*bumper*-CAR), a result confirmed by Muehlhaus et al. (2013). Further, in two PWI experiments using two different methods of determining semantic overlap, Mahon et al. (2007; Experiments 5 and 7) showed target naming latencies (e.g., HORSE) were facilitated for semantically “close” distractors (e.g., *zebra*) compared to semantically “far” distractors (e.g., *whale*). Mahon et al. (2007) argued that part and semantically close distractors should have higher conceptual-lexical activation levels due to sharing features with the target and thus be stronger competitors according to the competitive lexical selection account. They therefore proposed a post-lexical, non-competitive, *response exclusion* account of lexical selection. According to this account, conceptual feature overlap between distractor and target invariably induces semantic priming. Semantic interference in PWI instead reflects the extent to which the distractor is a relevant response to the task of naming the target picture. If the distractor is a relevant response to the target (e.g., another animal), a post-lexical decision mechanism must take more time to clear it from an articulatory buffer. Further, the account predicts the part-whole facilitation effect in PWI (Costa et al., 2005), as the “part” (e.g., *bumper*) is not a relevant response to the target picture (e.g., CAR).

Rahman and Melinger (2009) modified the competitive lexical selection account to explain part-whole and semantic distance facilitation effects in the PWI paradigm. According to their *swinging lexical network* model, feature-overlap between targets and distractors invariably produces semantic priming *and* interference. A net result of interference or facilitation depends upon the pattern of activation within the network. If shared features between the target and distractor activate a cohort of within-category lexical competitors, this creates one-to-many competition, and the net result is interference. Facilitation results when feature overlap does not spread to many lexical competitors, causing one-to-one competition. As distractors that are parts of whole objects do not spread activation to other related concepts, they produce one-to-one rather than one-to-many competition, and the net result is facilitation. Similarly, facilitation for semantically close distractor-target pairings is attributed to stronger priming due to feature overlap coupled with activation of a narrower category cohort of competitors (e.g., HORSE and *zebra* will activate only members of the equine category), contrasted with weaker priming and activation of a larger cohort for semantically far distractors (e.g., HORSE and *whale* will activate the broader category of animals).

However, more recent research has failed to elicit facilitation effects with similar stimuli. For example, Piai et al. (2011) noted that part-whole facilitation might instead be driven by strong associative links between the part distractor and its corresponding whole. Previous research has shown naming latencies are facilitated when targets are paired with distractors that are associates (e.g., SHIP-port; La Heij et al., 1990; Alario et al., 2000). Muehlhaus et al. (2013) selected part-whole stimuli that were strong associates using cue-target free association norms. Consistent with this explanation, Sailor and Brooks (2014) found that part-distractors produced facilitation only when they were associated with the target. When not associated with the target,

part-distractors produced an interference effect compared to parts unrelated to the target object (Experiments 1 and 3). Further, Sailor and Brooks (2014) were unable to replicate the findings from Costa et al.’s (2005) second experiment using identical materials. In two separate PWI experiments, Vieth et al. (2014) were likewise unable to replicate the facilitation effect reported by Mahon et al. (2007; Experiment 7) using near identical stimuli based on feature production norms (McRae et al., 2005). Instead, they found reliably greater interference for distractors that shared more features with the target.

Might there be another explanation for the (albeit equivocal) reports of feature overlap producing facilitation effects in PWI? To date, all accounts have emphasized feature-overlap between concepts. However, there is considerable behavioral evidence, supported by computational simulations, that distinctive features are activated differentially—and perhaps preferentially—to shared features (Randall et al., 2004; Cree et al., 2006; Grondin et al., 2009). Distinctive features can be defined as features that are (ideally) a perfect cue to a concept, distinguishing it from other related concepts, or in terms of narrowing a contrast set. For instance, the feature “has a *talon*,” is likely to narrow a contrast set to <birds of prey> (see Cree et al., 2006). As Grondin et al. (2009) note, distinctive features “make it easier to respond when the task requires distinguishing an item from among similar items, such as when naming the picture of an object” (p. 6, see also Cree et al., 2006; Taylor et al., 2012).

An examination of the stimuli employed by Mahon et al. (2007) in their Experiment 5 indicates that 17/20 of the *close* target-distractor pairings involved *at least one* distinguishing feature (e.g., HORSE-zebra). These stimuli were selected based on semantic similarity ratings from an independent sample of participants. Past research has shown that similarity ratings tend to emphasize the importance of shared features while deemphasizing distinguishing features (e.g., Medin et al., 1995; Kaplan and Medin, 1997). For example, the *coincidence effect* refers to the finding that two items (e.g., *horse* and *zebra*) that are semantically close due to feature overlap (e.g., *equine animal*, *has legs*, *has a tail*, etc.) yet differ due to a distinguishing feature (e.g., *has stripes*) will tend to receive a higher similarity rating than do two items that share a similar number of semantic features yet only differ modestly (e.g., *horse* and *donkey*). Thus, if distinctive features have a privileged role during conceptual processing (Cree et al., 2006), in that they are activated more quickly and/or strongly than shared features, this might explain why Mahon et al. (2007) (Experiment 5) observed facilitation for their semantically close distractors that contained a high proportion of distinctive features, despite also sharing a number of features with the target pictures.

A similar examination of the part-whole stimuli employed by Costa et al. (2005) indicates that many are distinctive parts of their targets according to published feature production norms (McRae et al., 2005; e.g. PERISCOPE-*submarine*; SINK-*drain*). Other pairings likewise appear distinguishing (e.g., CHURCH-*pew*; AMBULANCE-*stretcher*). As Costa et al. (2005; also Mahon et al., 2007) note, the activation-level of a part distractor should be raised when presented in conjunction with a target picture of the whole object to which it refers, due to feature overlap,

thus making it a more potent competitor according to the lexical selection by competition account. However, a part that is a distinctive feature and so potentially a perfect cue to the target concept should elicit less lexical-level activation than a part that is shared with other objects, due to less activation spreading at the conceptual level. This might explain why some studies observed facilitation with part-whole distractor-target pairings while others observed interference (e.g., Sailor and Brooks, 2014).

Thus, feature distinctiveness might be an important factor influencing the polarity of semantic effects in PWI paradigms. If so, accounts of semantic facilitation effects in spoken word production models would need to be modified to account for preferential processing of distinctive features. Conceivably, both post-lexical and swinging lexical network accounts of PWI effect could be modified to accommodate potential facilitatory effects of distinctive features in terms of stronger semantic priming, the former by assuming that the processing of distinctive distractors is privileged such that they enter the articulatory buffer earlier and are excluded accordingly, while the latter model could assume that distinctive features result in one-to-one rather than one-to-many competition at the lexical level due to their activating only the relevant target concept (see **Figure 1**), and so the net effect is semantic priming.

In this study, we report three PWI experiments examining effects of shared and distinctive distractor features. Experiment 1 manipulated distinctive distractor features while controlling for shared features, with the aim of determining whether the former might be responsible for eliciting a facilitation effect with categorical distractor-target relations (e.g., Mahon et al., 2007; Experiment 5). Experiments 2 and 3 investigated the role of feature distinctiveness with respect to part-whole distractor-target

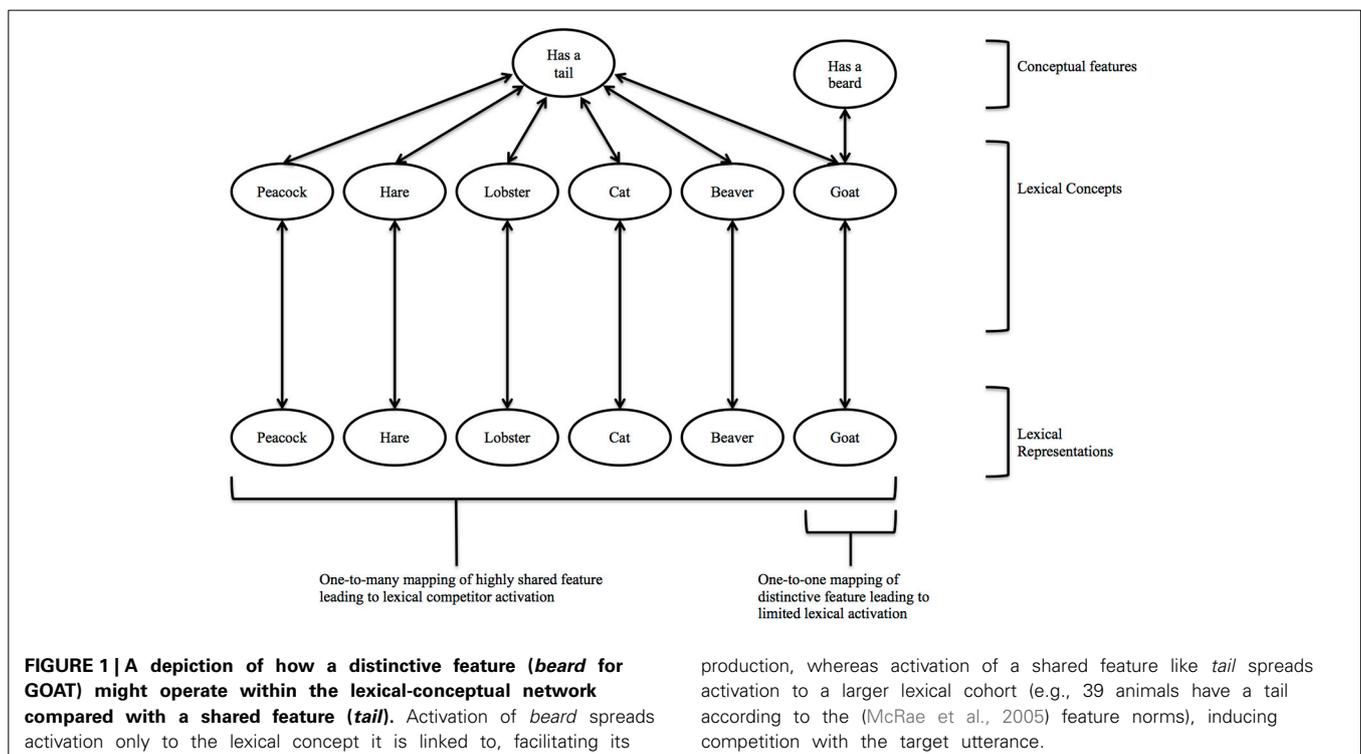
relations (e.g., a *hump* is a distinguishing part of a CAMEL, whereas *knee* is not, vs. an unrelated part such as *plug*). In all three experiments, targets and distractors were constructed so as to have minimal associative relations (e.g., Sailor and Brooks, 2014).

## EXPERIMENT 1

Experiment 1 tested whether *feature distinctiveness* might facilitate naming of categorically-related distractor-target pairings, as they are known to speed simple picture naming (e.g., Grondin et al., 2009). Past research has shown that similarity ratings tend to weight shared features as more important, with two items (e.g., *horse* and *zebra*) matching on one dimension (e.g., *equine animal*) yet differing considerably on another (e.g., *stripes*) tending to receive a higher similarity rating than two items that differ modestly (e.g., *horse* and *donkey*; Medin et al., 1995; Kaplan and Medin, 1997). As we noted in the Introduction to this paper, an examination of the *close* distractor-target pairings in Mahon et al.'s (2007) Experiment 5 revealed the majority involved distinguishing features (e.g., HORSE-zebra) according to feature production norms. Thus, distinguishing features might be responsible for the polarity reversal they observed. Experiment 1 therefore employed a set of target-distractor materials that manipulated distinctive features while controlling for semantic similarity.

## PARTICIPANTS

Participants were 50 students enrolled in first-year psychology courses at the University of Queensland. All were native English speakers. Each participant gave informed consent in accordance with the protocol approved by the Behavioral and Social Sciences



Ethical Review Committee of the University of Queensland and was compensated with course credit.

## DESIGN

Experiment 1 was a  $2 \times 2 \times 2$  mixed design. Independent variables within-participants were semantic relation (*semantically related, unrelated*), and distinctiveness (distinctive, non-distinctive) and SOA between-participants (−160 and 0 ms). These SOAs were selected based on the findings of significant facilitation effects in Mahon et al.'s (2007) Experiments 5 (0 ms) and 7 (−160 ms). Twenty-five participants were randomly assigned to each SOA.

## MATERIALS

Twenty target pictures and 40 distractor words were selected via a ratings study. Pictures were black-and-white line drawings, the majority of which were selected from normative picture databases (Cycowicz et al., 1997; Bonin et al., 2003; Szekely et al., 2004) with remaining items from the internet. The distractors were split into two sets of 20 categorically related items that were matched in terms of semantic similarity to the targets. In one of these sets (*similar-plus-distinctive*), each distractor additionally had at least one feature dimension rated as distinguishing it from the target, despite being matched in overall rated similarity. By way of example, a *semantically similar* pairing was PIGEON-sparrow while the corresponding *similar-plus-distinctive* pairing was PIGEON-canary. In order to reduce the number of related trials in the experiment to approximately 50%, two unrelated distractor conditions were created by re-pairing each distractor with an unrelated target picture (following Mahon et al., 2007; see Appendix A).

In order to create the *semantically similar* and *similar-plus-distinctive* target-distractor pairings, we performed two separate ratings studies. In the first, a group of 37 participants, none of whom participated in the PWI experiment, performed semantic similarity and dissimilarity judgments on a list comprising 72 word triplets, each triplet consisting of a target and two categorically related distractors. Targets were paired with each distractor separately on different trials. Participants were required to rate target-distractor word pairs presented in random order for semantic similarity (“how related are the two concepts denoted by the words”) on a scale of 1 to 7 (1 = unrelated, 7 = highly related) following Mahon et al. (2007). Subsequently, the participants were presented with the word triplets, again in random order, and instructed to select the distractor concept that differed most from the target and nominate the distinguishing feature. In the second ratings study, another group of 11 participants, none of whom participated in the first rating study or the PWI experiment, rated each word for imageability (“the ability to form a picture of the word’s referent in your mind”) following Mahon et al. (2007). Ratings were made on a scale of 1 to 7 (1 = not imageable, 7 = highly imageable).

The sets of 20 *semantically similar* and 20 *similar-plus-distinctive* distractors were thus created using triplets in which both distractors had been rated as highly similar to the target. The *similar-plus-distinctive* distractors were selected according to the consistency with which a distinguishing feature dimension

had been nominated across participants (criterion > 70%). Distractors in both sets were also matched according to imageability ratings, frequency, number of morphemes, syllables, and phonemes, word length, orthographic (OLD) and phonological Levenshtein Distance (PLD) (see Table 1; Balota et al., 2007). A series of *t*-tests demonstrated no significant differences between semantically related conditions on similarity to target  $t_{(38)} = 1.006, p = 0.32$ , imageability  $t_{(38)} = 1.68, p = 0.10$ , word length  $t_{(38)} = 0.21, p = 0.84$ , frequency  $t_{(38)} = 0.17, p = 0.87$ , OLD  $t_{(38)} = 0.17, p = 0.87$ , PLD  $t_{(38)} = 0.71, p = 0.71$ , number of phonemes  $t_{(38)} = 0.61, p = 0.54$ , number of syllables  $t_{(38)} = 0, p = 1$ , and number of morphemes  $t_{(38)} = 1.24, p = 0.22$ . Trials were randomized using Mix software (van Casteren and Davis, 2006) with the constraints that two presentations of the same picture were always interceded by at least five different pictures, and no more than two consecutive trials were of the same distractor type. One unique list per participant was generated.

## APPARATUS

Stimuli presentation, response recording and latency measurement (i.e., voice key) were accomplished via the Cogent 2000 toolbox extension ([www.vislab.ucl.ac.uk/cogent\\_2000.php](http://www.vislab.ucl.ac.uk/cogent_2000.php)) for MATLAB (2010a, MathWorks, Inc.) using a personal computer equipped with a noise-canceling microphone (Logitech, Inc.). The same apparatus was used in all subsequent experiments.

## PROCEDURE

Participants underwent pre-experimental familiarization with the target pictures by naming each three times in random order. The first presentation was accompanied by the target’s proper name printed below, with subsequent presentations only displaying the picture. Each experimental trial commenced with the participant pressing the space bar following the presentation of a question mark (?) at center-screen. Trials began by presenting a fixation cross center-screen for 500 ms, followed by a 50 ms blank screen. The distractor word appeared at −160 or 0 ms SOA relative to target presentation. Distractor words appeared randomly either above or below targets and counterbalanced across trials/conditions. Stimuli remained onscreen for 3000 ms or until the participant made a verbal response. A question

**Table 1 | Matching variables for the stimuli in Experiment 1.**

	Distractor Type	
	Similar	Similar-plus-distinctive
Rated similarity to target	5.33 (0.44)	5.20 (0.39)
Imageability	6.16 (0.65)	6.44 (0.35)
OLD	2.08 (0.74)	2.12 (0.76)
PLD	1.97 (0.82)	2.07 (0.86)
No. Phonemes	4.7 (1.34)	4.95 (1.23)
No. Syllables	1.75 (0.55)	1.75 (0.64)
No. Morphemes	1.1 (0.31)	1.25 (0.44)

Standard Deviations are in parentheses.

OLD, Orthographic Levenshtein Distance; PLD, Phonological Levenshtein Distance.

mark presented centrally then indicated that the participant could proceed to the next trial via space bar press.

## RESULTS AND DISCUSSION

Trials on which the voice key failed to detect a response (0.01%) were discarded as were latencies below 250 ms or above 2000 ms (2.5%). Latencies deviating more than 2.5 standard deviations from within-participant, within-condition means were excluded from analysis (5.7%). Errors were classified according to whether the participant hesitated during naming (i.e., dysfluencies) or misidentified the target, and due to their low frequency (1.6%) were not subjected to analysis. Mean naming latencies and error rates are summarized in **Table 2**.

Data were subjected to repeated measures ANOVAs with participants and items as random factors ( $F_1$  and  $F_2$ , respectively). There was a significant main effect of distractor relation,  $F_{1(1, 48)} = 8.40$ ,  $p = 0.006$ , partial  $\eta^2 = 0.15$ ,  $F_{2(1, 38)} = 14.41$ ,  $p = 0.001$ , partial  $\eta^2 = 0.28$ , yet no significant effect of distinctiveness  $F_{1(1, 48)} < 1$ ,  $p = 0.963$ , partial  $\eta^2 = 0.00$ ,  $F_{2(1, 38)} < 1$ ,  $p = 0.978$ , partial  $\eta^2 = 0.00$ . The effect of SOA was not significant by participants  $F_{1(1, 48)} < 1$ ,  $p = 0.326$ , partial  $\eta^2 = 0.02$ , although was significant by items  $F_{2(1, 38)} = 6.21$ ,  $p = 0.017$ , partial  $\eta^2 = 0.14$ , with naming latencies faster at SOA  $-160$  ms. There were no significant interactions between distractor relation and either distinctiveness,  $F_{1(1, 48)} < 1$ ,  $p = 0.546$ , partial  $\eta^2 = 0.01$ ,  $F_{2(1, 38)} < 1$ ,  $p = 0.469$ , partial  $\eta^2 = 0.01$ , or SOA,  $F_{1(1, 48)} < 1$ ,  $p = 0.561$ , partial  $\eta^2 = 0.01$ ,  $F_{2(1, 38)} < 1$ ,  $p = 0.601$ , partial  $\eta^2 = 0.01$ .

Separate analyses were conducted within each SOA. At  $-160$  ms SOA, there was a significant effect of distractor relation,  $F_{1(1, 24)} = 7.47$ ,  $p = 0.012$ , partial  $\eta^2 = 0.24$ ,  $F_{2(1, 19)} = 9.46$ ,  $p = 0.006$  partial  $\eta^2 = 0.33$ . However, there was no significant effect of distractor distinctiveness  $F_{1(1, 24)} < 1$ ,  $p = 0.537$ , partial  $\eta^2 = 0.02$ ,  $F_{2(1, 19)} < 1$ ,  $p = 0.409$ , partial  $\eta^2 = 0.04$ , or interaction between distinctiveness and relation,  $F_{1(1, 24)} < 1$ ,  $p = 0.760$ , partial  $\eta^2 = 0.00$ ,  $F_{2(1, 19)} < 1$ ,  $p = 0.792$ , partial  $\eta^2 = 0.00$ . At 0 ms SOA, there was no significant effect of distractor relation by participants  $F_{1(1, 24)} = 2.25$ ,

$p = 0.147$ , partial  $\eta^2 = 0.09$ , although the effect was significant by items  $F_{2(1, 19)} = 5.28$ ,  $p = 0.033$  partial  $\eta^2 = 0.22$ . Again, there was no significant effect of distinctiveness,  $F_{1(1, 24)} < 1$ ,  $p = 0.473$ , partial  $\eta^2 = 0.02$ ,  $F_{2(1, 19)} = 1.47$ ,  $p = 0.240$ , partial  $\eta^2 = 0.07$  and no interaction,  $F_{1(1, 24)} = 1.52$ ,  $p = 0.230$ , partial  $\eta^2 = 0.06$ ,  $F_{2(1, 19)} = 1.21$ ,  $p = 0.285$ , partial  $\eta^2 = 0.06$ .

Follow up comparisons investigated the significant main effects of distractor relation found at each SOA. At  $-160$  ms SOA, related distractor-target pairs were named significantly slower than unrelated pairs,  $t_{1(24)} = 2.73$ ,  $p = 0.012$ ,  $M_{diff} = 23$  ms, 95%  $CI = \pm 17$ ,  $t_{2(19)} = 3.08$ ,  $p = 0.006$ ,  $M_{diff} = 21$  ms, 95%  $CI = \pm 14$ . At 0 ms SOA, related distractor-target pairs were named significantly slower than unrelated pairs,  $t_{2(19)} = 2.30$ ,  $p = 0.033$ ,  $M_{diff} = 16$  ms, 95%  $CI = \pm 14$ .

Contrary to our prediction, categorically related distractors with distinguishing features did not influence picture naming latencies differentially: both *similar* and *similar-plus-distinctive* distractors elicited comparable interference compared to the matched unrelated distractors at each SOA. This result indicates that distinguishing features are unlikely to be responsible for semantic facilitation effects observed for categorically related distractors and targets in some PWI experiments using high proportions of distractor-target pairings with distinguishing features (e.g., Mahon et al., 2007; Experiment 5). Moreover, they indicate that conceptual feature overlap is the predominant factor influencing naming latencies in the PWI paradigm when distractors and targets are categorically related. However, the results of Experiment 1 are not informative with respect to the role of distinctive features when distractors are *not* category coordinates of the target picture, as is the case with part-whole relations (e.g., Costa et al., 2005). This latter scenario is explored in Experiment 2.

## EXPERIMENT 2

As noted in the Introduction, Costa et al. (2005) stimuli included distractors that denoted distinctive parts of their targets (e.g., *periscope*-SUBMARINE) according to feature production norms (McRae et al., 2005). In the absence of a categorical relation, part-whole distractor-target pairings represent a context in which a distinctive feature has a one-to-one relationship with a target picture concept that might facilitate its identification via semantic priming (e.g., Taylor et al., 2012), whereas the relationship of a non-distinctive feature is less clear as it is shared among other objects. Experiment 2 therefore employed a set of materials that manipulated distinctive vs. non-distinctive parts of target objects, while ensuring associative relations were minimal (e.g., Piai et al., 2011; Sailor and Brooks, 2014).

## PARTICIPANTS

Twenty-nine students from the University of Queensland participated in this study. All were native English speakers. Each participant gave informed consent and was compensated with course credit.

## RESEARCH DESIGN

Experiment 2 was a  $2 \times 2 \times 3$  repeated measures design, with target picture naming latencies being the dependent variable. The

**Table 2 | Experiment 1: Naming Latencies (in Milliseconds), 95% Confidence Intervals (CI), and Error rates (E%) by Type of Distractor and SOA.**

	Distractor condition			
	Semantically similar	Similar-plus-distinctive	Unrelated (similar)	Unrelated (similar-plus-distinctive)
<b>SOA <math>-160</math> ms</b>				
Mean	784	777	760	756
CI	$\pm 47$	$\pm 44$	$\pm 46$	$\pm 45$
E%	1	2	1.2	2
<b>SOA 0 ms</b>				
Mean	801	813	794	791
CI	$\pm 47$	$\pm 44$	$\pm 46$	45
E%	1.8	1.2	1	2.2

three independent variables were distractor part-relation (related, unrelated), distinctiveness (distinctive, non-distinctive), and SOA ( $-150$ ,  $0$ , or  $+150$  ms), using a within-participants design, following Sailor and Brooks' (2014) findings at SOAs of  $-150$  and  $0$  ms.

## MATERIALS

Twenty-four target pictures and 48 distractors were selected according to published feature production norms (McRae et al., 2005; see Appendix B). Pictures were color photographs sourced from normative databases (e.g., Adlington et al., 2009; Moreno-Martinez and Montoro, 2012) and the internet. Distinctive features were determined via the "distinctiveness" measure in the McRae et al. (2005) norms, defined as the inverse of the number of concepts in which that feature occurs in the norms. Therefore, those features with high scores occur less often between different concepts and are thus more distinct. For each target concept, a part feature was chosen that was high in distinctiveness (values of  $0.5$  and  $1$ ) and low in distinctiveness (values  $< 0.5$ ). The unrelated conditions were created by re-pairing the distinctive and non-distinctive distractor words with unrelated targets following Costa et al. (2005; Experiment 2). Thus, each picture appeared four times, and each distractor word was used twice (with the exception of *stem* that was paired four times with different pictures due to a clerical error; as the results reported below did not differ when this item was removed from analyses, it was retained). Distinctive and non-distinctive distractors were also matched on a number of lexical variables including length, frequency, number of syllables and phonemes, OLD and PLD and word mean bigram frequency (Balota et al., 2007), age of acquisition (Kuperman et al., 2012), and concreteness (Brysbaert et al., 2013), summarized in **Table 3**. None of the objects were associates (probabilities  $< 0.01$  in either direction) according to the University of South Florida Free Association Norms (Nelson et al., 2004) and Edinburgh

Associative Thesaurus (Kiss et al., 1973). Following Costa et al. (2005; p. 127), the part of the object to which the distractor referred was not visible in the target picture (see **Figure 2** for examples). There were no significant differences between distinctive and non-distinctive distractors on word length  $t_{(46)} = 0.12$ ,  $p = 0.91$ , frequency  $t_{(46)} = 0.09$ ,  $p = 0.93$ , OLD  $t_{(46)} = 0.64$ ,  $p = 0.52$ , PLD  $t_{(46)} = 0.79$ ,  $p = 0.44$ , number of phonemes  $t_{(46)} = 0.50$ ,  $p = 0.62$ , number of syllables  $t_{(46)} = 0.32$ ,  $p = 0.75$ , number of morphemes  $t_{(46)} = 0.59$ ,  $p = 0.56$ , bigram frequency  $t_{(46)} = 1.04$ ,  $p = 0.31$ , age of acquisition  $t_{(46)} = 0.30$ ,  $p = 0.77$ , imageability  $t_{(46)} = 1.21$ ,  $p = 0.23$  and concreteness  $t_{(46)} = 0.45$ ,  $p = 0.65$ .

## PROCEDURE

The pre-experimental familiarization and experimental trial delivery were identical to Experiment 1. Participants completed three blocks of picture naming trials, one block at each SOA, with a brief rest period between each block. Participants viewed each picture paired with three distractor types (distinctive, non-distinctive, and unrelated) at each SOA. The order of the trials within each block was pseudorandomized across participants using Mix software (van Casteren and Davis, 2006) such that two presentations of the same picture were always interceded by at least five different pictures, and no more than two consecutive trials were of the same distractor type. The order of the three SOA blocks was counterbalanced across participants according to a Latin square design.

## RESULTS AND DISCUSSION

Data from two participants were excluded as they failed to trigger the voice key on  $> 50\%$  of trials, leaving a final  $N = 27$ . Trials on which the voice key failed to detect a response ( $< 1\%$ ) were discarded as were latencies below  $250$  ms or above  $2000$  ms ( $0.5\%$ ). Latencies deviating more than  $2.5$  standard deviations from within-participant, within-condition means were excluded from analysis ( $3.1\%$ ). Errors were classified according to whether the participant hesitated during naming (i.e., dysfluencies) or misidentified the target, and due to their low frequency ( $0.4\%$ ) were not subjected to analysis.

Data was subjected to a repeated-measures ANOVA by participants and by items, denoted as  $F_1$  and  $F_2$ , respectively. Mean naming latencies,  $95\%$  CIs and error rates are summarized in **Table 4**. There were no significant effects of distractor part-relation,  $F_{1(1, 26)} < 1$ ,  $p = 0.705$ , partial  $\eta^2 = 0.01$ ,  $F_{2(1, 23)} < 1$ ,  $p = 0.659$ , partial  $\eta^2 = 0.01$ , or distinctiveness,  $F_{1(1, 26)} < 1$ ,  $p = 0.462$ , partial  $\eta^2 = 0.02$ ,  $F_{2(1, 23)} < 1$ ,  $p = 0.438$ , partial  $\eta^2 = 0.03$ . There was also no significant effect of SOA by participants  $F_{1(2, 52)} = 1.88$ ,  $p = 0.163$ , partial  $\eta^2 = 0.07$ , although the effect was significant by items,  $F_{2(2, 46)} = 4.56$ ,  $p = 0.016$ , partial  $\eta^2 = 0.17$ . As **Table 4** shows, naming latencies were faster overall at the  $-150$  ms SOA. There was no significant interaction between distractor part-relation and distinctiveness,  $F_{1(2, 52)} < 1$ ,  $p = 0.774$ , partial  $\eta^2 = 0.00$ ,  $F_{2(2, 46)} < 1$ ,  $p = 0.743$ , partial  $\eta^2 = 0.01$ . In addition, there was no significant part-relation  $\times$  SOA interaction,  $F_{1(2, 52)} < 1$ ,  $p = 0.905$ , partial  $\eta^2 = 0.00$ ,  $F_{2(2, 46)} < 1$ ,  $p = 0.772$ , partial  $\eta^2 = 0.01$ . There was also no significant

**Table 3 | Matching variables for the stimuli in Experiment 2.**

	Distractor type	
	Distinctive	Non-distinctive
Distinctiveness	0.88 (0.23)	0.14 (0.09)
Length	4.92 (1.32)	4.88 (1.12)
Frequency	34.09 (62.59)	35.54 (54.06)
OLD	1.59 (0.52)	1.68 (0.44)
PLD	1.50 (0.66)	1.38 (0.42)
Bigram frequency	1715.03 (914.47)	1982.26 (873.99)
Phonemes	3.96 (1.20)	3.79 (1.14)
Syllables	1.25 (0.44)	1.29 (0.46)
Age of acquisition	6.40 (2.02)	6.30 (2.35)
Imageability	5.70 (0.69)	6.01 (0.98)
Concreteness	4.64 (0.34)	4.69 (0.48)
Morphemes	1.08 (0.28)	1.04 (0.20)

Standard deviations are in parentheses.

OLD, Orthographic Levenshtein Distance; PLD, Phonological Levenshtein Distance.



**Table 4 | Experiment 2: Naming Latencies (in Milliseconds), 95% Confidence Intervals (CI), and Error rates (E%) by Type of Distractor and SOA.**

	Distractor condition			
	Distinctive	Non-distinctive	Unrelated (distinctive)	Unrelated (non-distinctive)
<b>SOA -150</b>				
Mean	655	656	654	652
CI	±20	±20	±21	±20
E%	0.2	0.6	0.3	0.3
<b>SOA 0</b>				
Mean	665	669	665	665
CI	±21	±22	±23	±24
E%	0.3	0.7	0.5	0.3
<b>SOA +150</b>				
Mean	661	662	659	659
CI	±25	±23	±24	±19
E%	0.2	0.5	0.3	0.3

distinctiveness  $\times$  SOA interaction,  $F_{1(2,52)} < 1$ ,  $p = 0.716$ , partial  $\eta^2 = 0.01$ ,  $F_{2(2,46)} < 1$ ,  $p = 0.894$ , partial  $\eta^2 = 0.01$ . Finally, there was no significant three-way interaction between distractor relation, distinctiveness, and SOA,  $F_{1(2,52)} < 1$ ,

$p = 0.698$ , partial  $\eta^2 = 0.01$ ,  $F_{2(2,46)} < 1$ ,  $p = 0.918$ , partial  $\eta^2 = 0.00$ .

The results of Experiment 2 can be summarized as follows: part-whole distractor-target relations did not influence naming latencies compared to unrelated parts, irrespective of whether the part was a distinctive feature of the depicted object. The failure to observe an effect of part-whole relatedness is inconsistent with the results of Costa et al. (2005; also Muehlhaus et al., 2013), although consistent with the findings of Sailor and Brooks (2014; Experiments 2 and 3) for non-associate parts at the same SOAs. Thus, associative strength might be a confounding factor for reports of facilitation effects with part-whole relations as proposed by Piai et al. (2011; see also Sailor and Brooks, 2014).

However, it is possible that our failure to obtain an effect of feature distinctiveness for related part distractors reflects the manner in which the stimuli were constructed. Following Costa et al. (2005), the part of the object to which the distractor referred was not visible in the target picture (cf., Sailor and Brooks, 2014, Experiment 2). Feature-distinctiveness effects have been reported in basic level picture naming (e.g., Taylor et al., 2012). As Cree et al. (2006) note, in such tasks it is beneficial to *recognize* a visual feature that is unique to the target. Accordingly, we conducted Experiment 3 to test whether feature distinctiveness will influence picture naming latencies when the distractor refers to a part that is visible in the target object.

### EXPERIMENT 3

Experiment 3 tests whether feature distinctiveness will influence picture naming latencies in the PWI paradigm when the distractor refers to a part that is visible in the target object.

#### PARTICIPANTS

Participants were 27 students from the University of Queensland. All were native English speakers. Each participant gave informed consent and was compensated with course credit.

#### RESEARCH DESIGN

The design was identical to Experiment 2.

#### MATERIALS

Materials were constructed in an identical manner to Experiment 2, although the features that the related-part distractors referred to were now visible in the respective target pictures (see Appendix C). In order to ensure feature visibility, some of the non-distinctive items used in Experiment 2 were replaced. Distinctive and non-distinctive distractors were also matched on a number of lexical variables (see Table 5) including length, frequency, number of syllables and phonemes, OLD and PLD and word mean bigram frequency (Balota et al., 2007), age of acquisition (Kuperman et al., 2012), and concreteness (Brysbaert et al., 2013). None of the objects were associates (probabilities < 0.01 in either direction) according to the University of South Florida Free Association Norms (Nelson et al., 2004) and Edinburgh Associative Thesaurus (Kiss et al., 1973). There were no significant differences between distinctive and non-distinctive part distractors on word length  $t_{(46)} = 1.57, p = 0.12$ , frequency  $t_{(46)} = 0.10, p = 0.92$ , OLD  $t_{(46)} = 0.31, p = 0.76$ , PLD  $t_{(46)} = 1.63, p = 0.11$ , number of phonemes  $t_{(46)} = 0.141, p = 0.16$ , number of syllables  $t_{(46)} = 1.42, p = 0.16$ , bigram frequency  $t_{(46)} = 0.49, p = 0.63$ , age of acquisition  $t_{(46)} = 1.90, p = 0.06$ , imageability  $t_{(46)} = 1.14, p = 0.26$  and concreteness  $t_{(46)} = 1.08, p = 0.28$ .

**Table 5 | Matching variables for the stimuli in Experiment 3.**

	Distractor type	
	Distinctive	Non-distinctive
Distinctiveness	0.88 (0.23)	0.12 (0.09)
Length	5.00 (1.32)	4.50 (1.06)
Frequency	30.89 (60.74)	30.69 (41.17)
OLD	1.63 (0.51)	1.62 (0.40)
PLD	1.53 (0.65)	1.31 (0.32)
Bigram frequency	1706.76 (914.38)	1822.09 (948.50)
Phonemes	4.00 (1.18)	3.58 (1.06)
Syllables	1.29 (0.46)	1.13 (0.34)
Age of acquisition	6.48 (1.91)	5.47 (1.84)
Imageability	5.69 (0.68)	5.96 (0.92)
Concreteness	4.61 (0.34)	4.73 (0.46)

Standard deviations are in parentheses.

OLD, Orthographic Levenshtein Distance; PLD, Phonological Levenshtein Distance.

#### PROCEDURE

The procedure was identical to Experiment 2.

#### RESULTS AND DISCUSSION

Trials on which the voice key failed to detect a response (<1%) were discarded as were latencies below 250 ms or above 2000 ms (<1%). Latencies deviating more than 2.5 standard deviations from within-participant, within-condition means were excluded from analysis (3.2%). Errors were classified according to whether the participant hesitated during naming (i.e., dysfluencies) or misidentified the target, and due to their low frequency (1.2%) were not subjected to analysis. Data were subjected to repeated-measures ANOVAs by participants and by items. Means, CIs, and error rates are reported in Table 6.

The main effect of distractor part-relation was not significant,  $F_{1(1,26)} < 1, p = 0.428$ , partial  $\eta^2 = 0.02$ ,  $F_{2(1,23)} < 1, p = 0.480$ , partial  $\eta^2 = 0.02$ . There was also no significant main effect of distinctiveness,  $F_{1(1,26)} < 1, p = 0.333$ , partial  $\eta^2 = 0.04$ ,  $F_{2(1,23)} < 1, p = 0.330$ , partial  $\eta^2 = 0.04$ . Although the main effect of SOA was not significant by participants,  $F_{1(2,52)} < 1, p = 0.438$ , partial  $\eta^2 = 0.03$ , it was marginally significant by items  $F_{2(2,46)} = 3.15, p = 0.052$ , partial  $\eta^2 = 0.12$ . The interaction between distractor part-relation and distinctiveness was not significant,  $F_{1(1,26)} < 1, p = 0.515$ , partial  $\eta^2 = 0.02$ ,  $F_{2(1,23)} = 1.31, p = 0.264$ , partial  $\eta^2 = 0.05$ . This was also the case for the part-relation  $\times$  SOA interaction,  $F_{1(2,52)} = 2.02, p = 0.144$ , partial  $\eta^2 = 0.07$ ,  $F_{2(2,46)} = 1.72, p = 0.190$ , partial  $\eta^2 = 0.07$ , and distinctiveness  $\times$  SOA interaction,  $F_{1(2,52)} < 1, p = 0.576$ , partial  $\eta^2 = 0.02$ ,  $F_{2(2,46)} < 1, p = 0.649$ , partial  $\eta^2 = 0.02$ . However, the three-way interaction between distractor part-relation, distinctiveness and SOA was marginally significant,  $F_{1(2,52)} = 2.97, p = 0.060$  partial  $\eta^2 = 0.10$ ,  $F_{2(2,46)} = 2.70, p = 0.078$ , partial  $\eta^2 = 0.11$ .

Additional analyses investigated the three-way interaction. At  $-150$  ms SOA, there was a significant effect of part-relation

**Table 6 | Experiment 3: Naming Latencies (in Milliseconds), 95% Confidence Intervals (CI), and Error rates (E%) by Type of Distractor and SOA.**

	Distractor condition			
	Distinctive	Non-distinctive	Unrelated (distinctive)	Unrelated (non-distinctive)
<b>SOA -150</b>				
Mean	634	636	633	622
CI	$\pm 26$	$\pm 25$	$\pm 24$	$\pm 28$
E%	2	2	0.6	1.4
<b>SOA 0</b>				
Mean	634	638	639	635
CI	$\pm 29$	$\pm 29$	$\pm 33$	$\pm 32$
E%	1.5	1.7	1.4	1.1
<b>SOA +150</b>				
Mean	641	634	636	640
CI	$\pm 32$	$\pm 31$	$\pm 29$	$\pm 29$
E%	1.1	1.4	0.5	0.9

by participants,  $F_{1(1,26)} = 8.46$ ,  $p = 0.007$  partial  $\eta^2 = 0.25$ , but was only marginally significant by items  $F_{2(1,23)} = 3.77$ ,  $p = 0.065$ , partial  $\eta^2 = 0.14$ . There was no significant effect of distinctiveness  $F_{1(1,26)} = 1.41$ ,  $p = 0.246$ , partial  $\eta^2 = 0.05$ ,  $F_{2(1,23)} = 1.57$ ,  $p = 0.225$ , partial  $\eta^2 = 0.06$  or interaction by participants  $F_{1(1,26)} = 2.64$ ,  $p = 0.116$ , partial  $\eta^2 = 0.09$ , however the interaction was significant by items  $F_{2(1,23)} = 5.96$ ,  $p = 0.023$ , partial  $\eta^2 = 0.21$ . At 0 ms SOA, there was no significant effect of relatedness  $F_{1(1,26)} < 1$ ,  $p = 809$ , partial  $\eta^2 = 0.00$ ,  $F_{2(1,23)} < 1$ ,  $p = 0.716$ , partial  $\eta^2 = 0.01$ , no significant effect of distinctiveness  $F_{1(1,26)} < 1$ ,  $p = 0.946$ , partial  $\eta^2 = 0.00$ ,  $F_{2(1,23)} < 1$ ,  $p = 0.884$ , partial  $\eta^2 = 0.00$ , and no interaction  $F_{1(1,26)} < 1$ ,  $p = 0.342$ , partial  $\eta^2 = 0.04$ ,  $F_{2(1,23)} < 1$ ,  $p = 0.397$ , partial  $\eta^2 = 0.03$ . At 150 ms SOA, there was no significant effect of relatedness  $F_{1(1,26)} < 1$ ,  $p = 0.775$ , partial  $\eta^2 = 0.00$ ,  $F_{2(1,23)} < 1$ ,  $p = 0.762$ , partial  $\eta^2 = 0.00$ , no significant effect of distinctiveness  $F_{1(1,26)} < 1$ ,  $p = 0.517$ , partial  $\eta^2 = 0.02$ ,  $F_{2(1,23)} < 1$ ,  $p = 0.664$ , partial  $\eta^2 = 0.01$ , and no interaction  $F_{1(1,26)} = 2.38$ ,  $p = 0.135$ , partial  $\eta^2 = 0.08$ ,  $F_{2(1,23)} = 2.54$ ,  $p = 0.125$ , partial  $\eta^2 = 0.10$ .

Paired-samples *t*-tests were conducted to investigate the significant effects found at  $-150$  ms SOA. There were no significant differences between distinctive and non-distinctive distractors  $t_{1(26)} = 0.329$ ,  $p = 0.744$ ,  $M_{\text{diff}} = 2$  ms, 95%  $CI = \pm 11$ ,  $t_{2(23)} = 0.543$ ,  $p = 0.592$ ,  $M_{\text{diff}} = 3$  ms, 95%  $CI = \pm 10$  or between distinctive and unrelated distractors  $t_{1(26)} = 0.355$ ,  $p = 0.741$ ,  $M_{\text{diff}} = 1$  ms, 95%  $CI = \pm 9$ ,  $t_{2(23)} = 0.228$ ,  $p = 0.822$ ,  $M_{\text{diff}} = 1$  ms, 95%  $CI = \pm 12$ . However, there was a significant difference between non-distinctive distractors  $t_{1(26)} = 2.727$ ,  $p = 0.011$ ,  $M_{\text{diff}} = 14$  ms, 95%  $CI = \pm 11$ ,  $t_{2(23)} = 3.383$ ,  $p = 0.003$ ,  $M_{\text{diff}} = 16$  ms, 95%  $CI = \pm 10$  such that non-distinctive distractors were named more slowly than unrelated distractors.

The results of Experiment 3 differ from Experiment 2, in that *non-distinctive* part-whole target-distractor relations slowed picture naming latencies significantly at  $-150$  ms SOA compared to their matched unrelated pairings. This is consistent with the results of Sailor and Brooks (2014, Experiments 1 and 3) who reported interference from non-associated parts.

## GENERAL DISCUSSION

In three experiments using the PWI paradigm, we investigated the roles of distinctive vs. shared conceptual features in lexical access. Experiment 1 employed categorically-related distractor-target pairings manipulated in terms of the presence/absence of a distinctive feature. Experiments 2 and 3 manipulated part-whole related distractor-target pairings in terms of distinctive vs. non-distinctive features and in terms of feature visibility in the target pictures. In Experiments 1 and 2, feature distinctiveness did not influence picture naming latencies differentially. In Experiment 3, non-distinctive part distractors that were visible in the target pictures slowed picture naming latencies significantly compared to their matched unrelated distractors at SOA  $-150$  ms.

Experiment 1 indicates that the presence of a distinctive feature in categorically-related distractor-target pairings does not influence picture naming when those pairings are matched in terms of conceptual feature overlap. Semantically similar-plus-distinctive

distractors slowed picture naming to the same degree as semantically similar distractors *without* a distinctive feature. Therefore, it seems unlikely that distinctive features can explain some facilitation results reported with categorically-related, semantically-close stimuli (Mahon et al., 2007). As we tested more participants (25 at each SOA) than Mahon et al. (2007; 20 and 16 at each SOA in their Experiments 5 and 7, respectively), the null effects are unlikely to be due to lack of power. Why is it that distinctive features facilitate basic-level naming (Grondin et al., 2009; Taylor et al., 2012) and produce priming relative to shared features in word-feature verification tasks (e.g., Cree et al., 2006), yet do not influence naming latencies in PWI? Grondin et al. were careful to emphasize the importance of task variables for determining the relative contributions of distinctive vs. shared features to performance. In Experiment 1, both types of distractor also shared many features with the target. This suggests that distinctive feature activation does not predominate in the presence of activation from many shared features (e.g., Cree et al., 2006), and so does not influence production of the target name. This finding can be accommodated by existing competitive lexical selection (Vigliocco et al., 2004; Rahman and Melinger, 2009; Vieth et al., 2014) and response exclusion accounts (e.g., Mahon et al., 2007). In the former, feature overlap predominates and activates a lexical cohort with the net result being competition; in the latter, identical response relevant criteria result in the post-lexical decision mechanism taking more time to clear both types of distractor from the articulatory buffer.

Experiments 2 and 3 manipulated distinctive features to investigate the part-whole facilitation effect reported by Costa et al. (2005). In Experiment 2, the part distractors were not visible in the target picture in keeping with Costa et al.'s (2005; p. 127) materials. Following proposals that distinctive features need to be visible in order to influence picture naming (Grondin et al., 2009), Experiment 3 ensured that the part the distractor referred to was visible in the target picture. In Experiment 2, we failed to find *any* effect of part-whole related compared to their matched unrelated part distractors. However, when the part denoted by a distractor was visible in the target (Experiment 3), only *non-distinctive* parts slowed picture naming latencies significantly compared to their matched unrelated parts. Sailor and Brooks (2014; Experiment 2) were unable to replicate the facilitation effect reported by Costa et al.'s (2005) Experiment 2 with the same materials and procedure (but see Discussion re part visibility below). However, they demonstrated significant interference with non-associated part distractors in two other experiments.

The results of Experiment 3 are therefore broadly consistent with those of Sailor and Brooks' (2014), in that we also observed interference with non-associated parts. However, they also add to this finding by demonstrating that non-associated part distractors are likely to slow naming latencies in PWI *only* if they do not denote a distinctive feature of the target picture concept. These findings can be accommodated by the lexical selection by competition account. According to this account, activation should spread from the target (e.g., GOAT) to the part distractor (e.g., tail). As non-distinctive parts are shared by many category exemplars (e.g., most animals have tails), spreading activation should therefore result in greater competition at the lexical level. By

contrast, as the target spreads activation only to the distinctive part (e.g., *beard*), less lexical competition occurs due to the one-to-one mapping (see **Figure 1**). A caveat to this interpretation is that the mean naming latencies for distinctive vs. non-distinctive part distractors did not differ significantly<sup>1</sup>. Interestingly, this was the same pattern reported for the mean naming latencies in Experiments 1 and 3 of Sailor and Brooks (2014), i.e., naming latencies for their associated and non-associated part-related distractors were comparable (see their **Tables 1, 3**). Nonetheless, the principal comparisons of interest are between each type of related part and their identically matched unrelated distractors. Although the distinctive and non-distinctive distractor words were matched on a range of variables (see **Table 5**), they were not matched identically as was the case with their respective unrelated distractors.

The results of Experiments 2 and 3 also highlight a potentially important role for feature visibility in determining whether interference will be observed. In conventional PWI experiments with categorically-related distractors, object features are typically visible in the target picture. According to the lexical selection by competition account, the target concept spreads activation to the related distractor due to feature overlap, raising its activation level and that of other lexical competitors. This might explain why distractors denoting visible non-distinctive parts interfered with target picture naming (Experiment 3), compared to non-visible parts (Experiment 2). Cree et al. (2006) had earlier proposed that a feature must be *recognized* in the target object in order for it to be beneficial to picture naming. In terms of PWI, this suggests the target picture concept is able to spread activation to the part distractor once the part is recognized, and this activation then spreads to the lexical level. Thus, feature visibility might be an important factor determining whether interference effects will be elicited with part distractors, and whether facilitation will predominate when associative relations are also present. For example, Costa et al. (2005; Experiment 2) ensured the parts denoted by their distractors (many of which were distinctive *and* associates) were not visible in the target pictures, whereas Sailor and Brooks' (2014) replication of Costa et al.'s experiment did not.

The findings of interference for part-whole related distractors have implications for recently formulated models of lexical access and PWI effects (see Sailor and Brooks, 2014). Both the response exclusion (Mahon et al., 2007) and swinging lexical network (Rahman and Melinger, 2009) accounts were developed to explain reports of semantic facilitation that were deemed problematic for the conventional lexical selection by competition account. Following those earlier reports, both accounts assumed that part distractors facilitate whole object naming via semantic priming. However, it seems that facilitation effects for part distractors in PWI might not be reproducible, unless parts also have an associative relation with the target picture, as proposed by Piai et al. (2011; e.g., Muehlhaus et al., 2013; Sailor and Brooks, 2014). Facilitation with associative part relations can be accounted for by a competitive lexical selection model by assuming the effect occurs at the conceptual level (see La Heij et al., 1990, 2006).

One possible way of modifying the response exclusion account to explain the interference effect observed in Experiment 3 might involve making the additional assumption that visible features of target pictures constitute response relevant criteria, despite the instruction to name the whole object (see also Sailor and Brooks, 2014). However, adopting this modification would first involve abandoning Mahon et al.'s (2007) proposal that conceptual feature overlap does not constitute a response-relevant criterion.

Theoretical accounts of PWI effects have emphasized the semantic relationship between concepts as the determining factor of an effect. However, experimental evidence shows that wide ranges of effects are possible for each type of relationship (i.e., categorical, associative, part-whole). This suggests that variables other than semantic relationship can influence the polarity of PWI effects, and that other reports of semantic facilitation in PWI might be due to task and/or procedural factors. For example, in their Experiment 1, Costa et al. (2005) compared part distractors (e.g., *LAMP-bulb*) to categorical, but unrelated distractors (e.g., *LAMP-wolf*) rather than part distractors at the same level of categorization as in the present and other studies (e.g., Sailor and Brooks, 2014). Costa et al. (2003) had earlier argued that the level of categorization could be used by the semantic system to differentiate the conceptual representations corresponding to the target and distractor. According to their semantic selection account, when target and distractor are from different levels of categorization the semantic system discards the distractor's conceptual representation for further processing, preventing lexical competition from arising. However, the distractor's conceptual representation will enhance the activation of the target, leading to semantic facilitation (but see Kuipers et al., 2006; Hantsch et al., 2012).

Although semantic facilitation in PWI has proved difficult to reproduce in the absence of associative relations, a study by Collina et al. (2013) suggests picture familiarization might also be a possible cause of semantic polarity reversals in PWI. In most PWI studies, participants are typically familiarized with the target pictures two-to-four times prior to performing the experimental series, as was the case in the present study (e.g., Starreveld and La Heij, 1995, 1996; Damian and Martin, 1999; Mahon et al., 2007). In Collina et al.'s study, participants who were familiarized with the target pictures showed interference compared to those who were not familiarized with the target pictures while the latter group showed facilitation. Given that a picture familiarization phase is a standard procedure in PWI experiments (e.g., Starreveld and La Heij, 1995, 1996; Mahon et al., 2007), Collina et al.'s (2013) finding warrants replication and further investigation.

In summary, our findings do not provide empirical support for the proposal that part-whole distractor-target relations facilitate naming in PWI via semantic priming (cf. Costa et al., 2005; Mahon et al., 2007), unless an associative relation is also involved (e.g., Piai et al., 2011; Muehlhaus et al., 2013; Sailor and Brooks, 2014). Instead, our findings indicate that an interference effect can be observed when a non-associated part distractor denotes a conceptual feature shared by the target and other category exemplars. This activation appears contingent on the

<sup>1</sup>We are grateful to an anonymous reviewer for drawing our attention to this.

feature denoted by the part distractor being visible in the target picture. Distinctive features did not influence the level of lexical activation significantly. Together, these findings indicate that semantic interference effects in the PWI paradigm are a product of conceptual feature overlap, consistent with the assumptions of prominent lexical selection by competition accounts (e.g., Roelofs, 1992; Starreveld and La Heij, 1995, 1996; Levelt et al., 1999).

## ACKNOWLEDGMENTS

The authors were supported by an Australian Postgraduate Award (Harrison E. Vieth), Australian Research Council (ARC) Discovery Grant (DP1092619) and Future Fellowship (FT0991634) (Greig I. de Zubicaray). The authors are grateful to Freeda Thong and Megan Barker for their assistance with collecting data, and to the two reviewers for their helpful comments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.01014/abstract>

## REFERENCES

- Adlington, R. L., Laws, K. R., and Gale, T. M. (2009). The HatField image test: a new picture test and norms for experimental and clinical use. *J. Clin. Exp. Neuropsychol.* 31, 731–775. doi: 10.1080/13803390802488103
- Alario, F. X., Segui, J., and Ferrand, L. (2000). Semantic and associative priming in picture naming. *Q. J. Exp. Psychol. A* 53, 741–764. doi: 10.1080/027249800410535
- Balota, D., Yap, M., Hutchison, K., Cortese, M., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behav. Res. Methods* 39, 445–459. doi: 10.3758/bf03193014
- Bonin, P., Peereman, R., Malardier, N., Meot, A., and Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behav. Res. Methods* 25, 158–167. doi: 10.3758/BF03195507
- Brysbaert, M., Warriner, A. B., and Kuperman, V. (2013). Concrete ratings for 40 thousand generally known English word lemmas. *Behav. Res. Methods* 46, 904–911. doi: 10.3758/s13428-013-0403-5
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208.
- Collina, S., Tabossi, P., and De Simone, F. (2013). Word Production and the Picture-Word Interference paradigm: the role of learning. *J. Psycholinguist. Res.* 42, 461–473. doi: 10.1007/s10936-012-9229-z
- Collins, A. M., and Loftus, E. F. (1975). Spreading activation theory of semantic processing. *Psychol. Rev.* 82, 407–428.
- Costa, A., Alario, F. X., and Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychon. Bull. Rev.* 12, 125–131. doi: 10.3758/BF03196357
- Costa, A., Mahon, B., Savova, V., and Caramazza, A. (2003). Level of categorization effect: a novel effect in the picture-word interference paradigm. *Lang. Cogn. Process.* 18, 205–233. doi: 10.1080/01690960143000524
- Cree, G. S., McNorgan, C., and McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: implications for theories of semantic memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 643–658. doi: 10.1037/0278-7393.32.4.643
- Cyowicz, Y. M., Friedman, D., and Rothstein, M. (1997). Picture naming by young children: norms for name agreement, familiarity and visual complexity. *J. Exp. Child Psychol.* 65, 171–237.
- Damian, M. F., and Martin, R. C. (1999). Semantic and phonological codes interact in single word production. [Research Support, U.S. Gov't, P.H.S.]. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 345–361. doi: 10.1037/0278-7393.25.2.345
- Dell, G. S., and O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition* 42, 287–314. doi: 10.1016/0010-0277(92)90046-k
- Goldrick, M., and Rapp, B. (2002). A restricted interaction account (RIA) of spoken word production: the best of both worlds. *Aphasiology* 16, 20–55. doi: 10.1080/02687040143000203
- Grondin, R., Lupker, S. J., and McRae, K. (2009). Shared features dominate semantic richness effects for concrete concepts. *J. Mem. Lang.* 60, 1–19. doi: 10.1016/j.jml.2008.09.001
- Hantsch, A., Jescheniak, J. D., and Mädebach, A. (2012). Naming and categorizing objects: task differences modulate the polarity of semantic effects in the picture-word interference paradigm. *Mem. Cognit.* 40, 760–768. doi: 10.3758/s13421-012-0184-6
- Kaplan, A., and Medin, D. (1997). The coincidence effect in similarity and choice. *Mem. Cognit.* 25, 570–576. doi: 10.3758/bf03201130
- Kiss, G. R., Armstrong, C., Milroy, R., and Piper, J. (1973). "An associative thesaurus of English and its computer analysis," in *The Computer and Literary Studies*, eds A. J. Aitken, R. W. Bailey, and N. Hamilton-Smith (Edinburgh: University Press), 153–166.
- Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008
- Kuipers, J.-R., La Heij, W., and Costa, A. (2006). A further look at semantic context effects in language production: the role of response congruency. *Lang. Cogn. Process.* 21, 892–919. doi: 10.1080/016909600824211
- Kuperman, V., Stadthagen-Gonzalez, H., and Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behav. Res. Methods* 44, 978–990. doi: 10.3758/s13428-012-0210-4
- La Heij, W., Dirx, J., and Kramer, P. (1990). Categorical interference and associative priming in picture naming. *Br. J. Psychol.* 81, 511–525.
- La Heij, W., Kuipers, J.-R., and Starreveld, P. A. (2006). In defense of the lexical-competition account of picture-word interference: a comment on Finkbeiner and Caramazza (2006). *Cortex* 42, 1028–1031. doi: 10.1016/S0010-9452(08)70209-0
- La Heij, W., and van den Hof, E. (1995). Picture-word interference increases with target-set size. *Psychol. Res.* 58, 119–133. doi: 10.1007/bf00571100
- Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. discussion: 38–75.
- Levelt, W. J., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., and Havinga, J. (1991). The time course of lexical access in speech production: a study of picture naming. *Psychol. Rev.* 98, 122–142. doi: 10.1037/0033-295x.98.1.122
- Lupker, S. (1979). The semantic nature of response competition in the picture-word interference task. *Mem. Cognit.* 7, 485–495. doi: 10.3758/bf03198265
- Mahon, B., Costa, A., Peterson, R., Vargas, K., and Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 503–535. doi: 10.1037/0278-7393.33.3.503
- McRae, K., Cree, G., Seidenberg, M., and McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behav. Res. Methods* 37, 547–559. doi: 10.3758/bf03192726
- Medin, D., Goldstone, R., and Markman, A. (1995). Comparison and choice: Relations between similarity processes and decision processes. *Psychon. Bull. Rev.* 2, 1–19. doi: 10.3758/bf03214410
- Moreno-Martínez, F. J., and Montoro, P. R. (2012). An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS ONE* 7:e37528. doi: 10.1371/journal.pone.0037527
- Muehlhaus, J., Heim, S., Sachs, O., Schneider, F., Habel, U., and Sass, K. (2013). Is the motor or the garage more important to the car? *Diff. between semantic associations in single word and sentence production. J. Psychol. Res.* 42, 37–49. doi: 10.1007/s10936-012-9209-3
- Nelson, D. L., McEvoy, C. L., and Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behav. Res. Methods Instrum. Comput.* 36, 402–407. doi: 10.3758/BF03195588
- Piai, V., Roelofs, A., and Schriefers, H. (2011). Distractor strength and selective attention in picture-naming performance. *Mem. Cognit.* 40, 614–644. doi: 10.3758/s13421-011-0171-3
- Rahman, R. A., and Melinger, A. (2009). Semantic context effects in language production: a swinging lexical network proposal and a review. *Lang. Cogn. Process.* 24, 713–734. doi: 10.1080/01690960802597250

- Randall, B., Moss, H. E., Rodd, J. M., Greer, M., and Tyler, L. K. (2004). Distinctiveness and correlation in conceptual structure: behavioral and computational studies. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 393–406. doi: 10.1037/0278-7393.30.2.393
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142.
- Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Dev.* 48, 643–647. doi: 10.2307/1128667
- Sailor, K., and Brooks, P. J. (2014). Do part-whole relations produce facilitation in the picture-word interference task? *Q. J. Exp. Psychol. (Hove)*. doi: 10.1080/17470218.2013.870589. [Epub ahead of print].
- Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production - picture-word interference studies. *J. Mem. Lang.* 29, 86–102.
- Starreveld, P. A., and La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 686.
- Starreveld, P. A., and La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 896–918. doi: 10.1037/0278-7393.22.4.896
- Szekely, A., Jacobsen, T., D'Ámico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *J. Mem. Lang.* 51, 247–250. doi: 10.1016/j.jml.2004.03.002
- Taylor, K. I., Devereux, B. J., Acres, K., Randall, B., and Tyler, L. K. (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. *Cognition* 122, 363–374. doi: 10.1016/j.cognition.2011.11.001
- van Casteren, M., and Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behav. Res. Methods* 38, 584–589. doi: 10.3758/BF03193889
- Vieth, H. E., McMahon, K. L., and de Zubicaray, G. I. (2014). Feature overlap slows lexical selection: evidence from the picture-word interference paradigm. *Q. J. Exp. Psychol. (Hove)*. doi: 10.1080/17470218.2014.923922. [Epub ahead of print].
- Vigliocco, G., Vinson, D. P., Lewis, W., and Garrett, M. F. (2004). Representing the meanings of object and action words: the featural and unitary semantic space hypothesis. *Cogn. Psychol.* 48, 422–488. doi: 10.1016/j.cogpsych.2003.09.001

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 April 2014; accepted: 26 August 2014; published online: 16 September 2014.

Citation: Vieth HE, McMahon KL and de Zubicaray GI (2014) The roles of shared vs. distinctive conceptual features in lexical access. *Front. Psychol.* 5:1014. doi: 10.3389/fpsyg.2014.01014

This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Vieth, McMahon and de Zubicaray. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Processing different kinds of semantic relations in picture-word interference with non-masked and masked distractors

Markus F. Damian<sup>1\*</sup> and Katharina Spalek<sup>2</sup>

<sup>1</sup> School of Experimental Psychology, University of Bristol, Bristol, UK

<sup>2</sup> Department of German Studies and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany

## Edited by:

Peter Indefrey, University of  
Dusseldorf, Germany

## Reviewed by:

Claudio Mulatti, Università degli  
Studi di Padova, Italy  
Vitoria Piai, University of California  
Berkeley, USA

## \*Correspondence:

Markus F. Damian, School of  
Experimental Psychology, University  
of Bristol, 12a Priory Road,  
Bristol BS8 1TU, UK  
e-mail: m.damian@bristol.ac.uk

Spoken production requires lexical selection, guided by the conceptual representation of the to-be-named target. Currently, the question whether lexical selection is subject to competition is hotly debated. In the picture-word interference task, manipulating the visibility of written distractor words provides important insights: clearly visible categorically related distractors cause interference whereas masked distractors induce facilitation (Finkbeiner and Caramazza, 2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stoop-like task. We explored the effect of distractor masking in more depth by investigating its interplay with different types of semantic overlap. Specifically, we contrasted categorical with associatively based relatedness. For the former, we replicated the polarity reversal in semantic effects dependent on whether distractors were masked or not. Post-experimental visibility tests showed that weak semantic facilitation with masked distractors did not depend on individual variability in participants' ability to perceive the distractors. Associatively related distractors showed facilitation with non-masked presentation, but little effect when masked. Overall, the results suggest that it is primarily distractor activation strength which determines whether semantic effects are facilitatory or interfering in PWI tasks.

**Keywords:** spoken production, picture-word interference, lexical access, object naming, competition

## INTRODUCTION

A hotly contested issue within recent research on language production is whether accessing a word in the mental lexicon (i.e., the store of words a speaker knows) is a competitive process or not. Competition is a ubiquitous concept in various aspects of language processing (e.g., Duffy et al., 1988; MacDonald et al., 1994; Green, 1998), and many models of language production likewise assume that lexical access is accomplished via a competitive principle. Most (if not all) models of language production stipulate that word preparation involves the temporary activation of a cohort of semantic alternatives. For instance, according to the influential model by Levelt et al. (1999), a number of competitors are initially co-activated until a winner is chosen—usually the intended word, or, in case of a speech error, often a semantically related word that has accumulated most activation. However, accounts differ in whether they depict the eventual selection of the target item as competitive or not. Competition in this context implies that the time to choose a target is dependent on the number of co-activated competitors and their activation strength. Competition can be implemented either as lateral inhibition (e.g., Cutting and Ferreira, 1999) or by a rule such as Luce's choice ratio (Luce, 1959, 1986; see Roelofs, 1992) in which the time to choose a target word varies as a function of the target word's activation in relation to the activation of its competitors.

A paradigm widely used to study lexical access in spoken word production is the picture-word interference (PWI) task (first introduced by Rosinski et al., 1975): on a given trial, participants see an object which they have to name, and naming latencies are measured. At the same time or in close temporal proximity, a distractor word is presented either visually or auditorily, and participants are instructed to ignore the distractor and focus on object naming. A standard finding in PWI tasks is the *semantic interference effect* (e.g., Glaser and Dünghoff, 1984; Schriefers et al., 1990; Damian and Martin, 1999): participants show slower average object naming latencies when distractor and target belong to the same semantic (taxonomic) category (e.g., *lion-monkey*) than when they are unrelated (e.g., *lion-cupboard*). This finding has been interpreted as evidence for competitive selection: the distractor word increases the activation of a non-target representation, thereby intensifying the underlying competition (see Roelofs, 1992, for computational modeling of this principle). However, this interpretation has recently been challenged based on a number of findings from PWI tasks which are potentially difficult to accommodate within a competitive framework, and alternative, non-competitive accounts of PWI (and more broadly, word production) have been introduced (e.g., Mahon et al., 2007; see Mulatti and Coltheart, 2012; Spalek et al., 2013, for recent overviews).

Semantic interference in PWI tasks arises most reliably when (a) distractor words are clearly visible (assuming visual distractor presentation), as opposed to when they are masked and hence difficult to see, and (b) target and distractor are coordinates of the same semantic category. In the following, we summarize the current state of knowledge with regard to these two aspects, and we then report an experiment which investigates how distractor visibility and semantic relation relate to each other.

### DISTRACTOR VISIBILITY

An important recent observation is that the semantic interference effect demonstrated in numerous previous studies reverses in polarity when distractors are masked. Finkbeiner and Caramazza (2006) compared picture-word interference with clearly visible and with masked distractors. In the latter case, participants were, according to post-experimental interviews, not consciously able to perceive the distractors. With clearly visible distractors, Finkbeiner and Caramazza obtained semantic interference effect (32 ms in their first experiment), but critically, masking of the distractors reversed the polarity of the effect such that it turned into strong and reliable facilitation (32 ms). This pattern was subsequently replicated by Dhooge and Hartsuiker (2010; Experiment 2) with Dutch speakers and materials, although resulting in somewhat smaller effects: semantic interference of 15 ms in the “visible” condition contrasted with semantic facilitation of 12 ms in the “masked” condition (presentation parameters between Dhooge and Hartsuiker’s and Finkbeiner and Caramazza’s studies were largely comparable).

These findings are crucial because they contribute to the wider debate on whether or not lexical retrieval in spoken word production is competitive. For advocates of a competitive view, it is not easy to explain why masking of distractors should reverse the polarity of semantic effects: without additional assumptions, their view would predict that masked distractors either generate semantic interference (perhaps reduced in size), or possibly lead to a null finding. Finkbeiner and Caramazza (2006) argued that the polarity reversal supports an account which dispenses with the notion of lexical competition altogether and instead locates semantic interference effects in PWI tasks at a post-lexical level. According to the “response exclusion hypothesis” (REH; see Mahon et al., 2007, for a detailed outline of this account), lexical access is fundamentally non-competitive, hence spreading of activation at the lexical level (which in the case of PWI arises from distractor processing) generally has facilitatory effects. In addition, however, at a later, post-lexical processing level, semantically related distractors can cause interference which may offset the facilitatory effects arising at the earlier processing levels. A distractor word in a PWI task is thought to temporarily occupy a single-channel prearticulatory “response buffer,” and needs to be removed before target naming can proceed. The time to remove a distractor word from the response buffer mainly depends on its “response relevance.” For instance, if the participant has to name a picture of an animal and the distractor word is the name of an animal (i.e., belongs to the same taxonomic category as the target), it is more difficult to purge the channel than if there is no relationship. In the same vein, if the task is to name an object, semantically related verbs do not interfere with naming, as demonstrated by

Mahon et al.<sup>1</sup> To account for the polarity reversal in semantic PWI effects as a result of distractor masking, Finkbeiner and Caramazza (2006) reasoned that because masking prevents the distractor word from occupying the response buffer, no interference arises. At the same time, masked distractor words are still sufficiently processed to generate semantically based facilitation in the mental lexicon.

Other explanations for the polarity reversal with masked distractor presentation in PWI tasks are possible, however, and crucially, competition need not be abandoned. Piai et al. (2012) suggested that whether or not competition arises might depend primarily on the activation strength of the distractor. Only distractors whose activation crosses a particular threshold will engage in competition with the target and generate semantic interference effects. By contrast, if distractor activation is too low, distractors will not be considered for response selection and hence will not lead to interference; however, such weakly activated distractors might still cause facilitation via overlap with the target at the semantic level. Hence, the polarity reversal demonstrated by Finkbeiner and Caramazza (2006) and Dhooge and Hartsuiker (2010) is explained not with the assumption that semantic facilitation and interference arise at two different loci (at lexical-semantic and response buffer levels, as advocated by the response exclusion hypothesis). Rather, the claim is that only strongly activated distractors will engage in competition with the target (and hence cause interference) whereas weak distractors will merely cause semantically based priming. This “competition threshold hypothesis” shifts the explanatory focus from conscious availability of the distractor identity (as in the response exclusion account) to distractor activation strength. In other words, even distractors which are clearly visible to the participant might not result in interference if they generate only weak activation.

This prediction was tested by Piai et al. (2012) in two experiments. The first experiment manipulated visibility via presence or absence of forward and backward masks around a briefly presented distractor. In a “clearly visible” condition, primes were presented for 53 ms, and following a blank period of 13 ms, the object was presented. Because distractors are not masked, this trial structure renders the distractor relatively easy to perceive. In a “poorly visible” condition, primes were again presented for 53 ms, but now they were preceded by a forward mask consisting of hash signs for 500 ms, and backward masked by a string of consonants for 13 ms before the object was presented (the latter condition is very similar to the masking employed in Finkbeiner and Caramazza, 2006; Dhooge and Hartsuiker, 2010).

<sup>1</sup>On a strict reading of the response exclusion hypothesis, the assumption that target naming is delayed until the distractor has been removed from the response buffer predicts that latencies should exclusively depend on distractor processing, and any effects associated with target processing should be obliterated (Mulatti and Coltheart, 2012). For instance, the well-documented frequency effect in object naming (e.g., Jescheniak and Levelt, 1994) should disappear in a PWI context because processing of both high- and low-frequency target names is delayed until the distractor has been purged; however, that is clearly not the case (e.g., Miozzo and Caramazza, 2003, Experiment 1). One could therefore argue, as Mulatti and Coltheart do, that the REH has already been refuted and no further experimentation is necessary to resolve the issue.

Each target was from a separate semantic category, and distractors never appeared as targets. Under these conditions, results showed a null effect for the “poorly visible” condition, and a *facilitatory* effect of 15 ms in the “clearly visible” condition. A second experiment was very similar to the first one, except that now there were four target exemplars per category, and distractors also appeared as target names. Both aspects should, according to the authors, increase co-activation of multiple entries in the lexicon. Now, results showed 17 ms interference for the “poorly visible” condition, and 13 ms interference in the “clearly visible” condition. According to the authors, these findings demonstrate that strength of distractor activation is the primary variable which determines whether semantically related distractors generate facilitation or interference. Presenting distractors only briefly generally reduces distractor strength, and masking further weakens distractor processing. Other variables (such as response set membership) further influence the degree of co-activation in the lexicon. Overall, Piai et al. suggested that polarity reversals of semantic effects in PWI do not contradict a general principle of competitive lexical access. At the same time, it is clear that the notion of a “competition threshold” represents an important modification of earlier competitive models (e.g., Roelofs, 1992; Levelt et al., 1999).

In our experiment reported below, we further explored the effects of visibility and co-activation on lexical competition. As in the previous studies, we manipulated visibility as a factor with two levels (masked vs. unmasked), but we also assessed individual differences in participants’ ability to extract information from briefly presented distractors. The intention was to explicitly probe the possible relationship between conscious availability of the distractor, and the size and direction of the resulting semantic effect. Finkbeiner and Caramazza (2006) merely asked participants, following the experiment, whether they had noticed any masked distractor words, and reported that only one participant reported being able to see some letters of masked words (this participant was subsequently replaced). Dhooge and Hartsuiker (2010) carried out a more explicit test of visibility: they selected pictures and words which were closely matched to the experimental stimuli, and presented them with the same timing and masking parameters as in the actual experiment. Participants were asked to indicate whether or not they had seen the distractor and if so, to report its name or some of its letters. None of their participants were able to report information on the distractors.

In our experiment we employed a lexical decision task (LDT) as a post-experimental visibility test. Participants were shown the distractor words from the earlier PWI task centered on the screen, using the same masking parameters as in the picture-word interference test. We generated and interleaved an equal number of non-words, and on each trial, participants indicated whether or not they thought the distractor was a word of their language (the experiment was conducted in German). Results from the LDT allowed us to compute individual  $d'$ -scores for each participant. It should be noted that because we used the same materials in the PWI and LDT task, distractors in the LDT had already presented multiple times in the PWI phase of the experiment. For this reason, performance on the LDT might overestimate individuals’ ability of having identified the distractors in the earlier

PWI phase. Nevertheless, we hoped to obtain a relatively wide range of variation in individual  $d'$  scores (and as will be shown below, this was clearly the case). This allowed us to explore the relation between distractor visibility and semantic effects in PWI. If conscious availability is the primary determinant of whether a semantic effect is positive or negative, then for participants with higher visibility scores, the effect should tend toward interference, whereas in participants with lower visibility scores, it should result in facilitation. By contrast, if distractor strength is the primarily important variable, then the masking procedure should generally (and independently of conscious distractor availability) weaken activation strength, and by and large, semantic effects should be facilitatory.

In making these predictions, it is acknowledged that the competition threshold claim makes it difficult to generate precise *a priori* predictions about when semantic interference should turn into facilitation. This is because the threshold itself is not objectively defined, but rather only *post-hoc* via an experimental effect—if semantic interference is found in an PWI task, then distractors must have been strongly enough activated to cross the threshold; if not, they were not.

#### TYPE OF SEMANTIC RELATIONSHIP

A further facet contributing to the recent debate on lexical competition in word production concerns the type of semantic relationship between distractor and target. Interference is generally only obtained with co-hyponyms (targets and distractors belonging to the same taxonomic category); other types of semantic relationships such as part-whole relationships (Costa et al., 2005; but see Sailor and Brooks, 2014), hypernymy-hyponymy (Kuipers and La Heij, 2008; but see Hantsch et al., 2005), and semantically related nouns and verbs (Mahon et al., 2007) tend to generate facilitation. The fact that interference is restricted to categorically related distractors and targets poses potential difficulties for the competitive view: if interference in PWI arises as a result of conceptual overlap, why does interference not extend to forms of overlap other than strict category membership? The REH accounts for this pattern via a principle of “response relevance”: categorically related distractors are response relevant in the sense that they could potentially be plausible target responses, and so take more time to remove from the response buffer. Non-categorically related distractors are not response relevant and so don’t result in interference in the buffer (but might generate facilitation via higher-level overlap with the target).

In the experiment reported below, we manipulated not only distractor visibility (see previous section) but also compared and contrasted the effects of categorically and associatively related distractors. We will briefly summarize previous findings on the effects of associative relationships in the PWI before outlining our motivation for including this form of relatedness in our own experiment.

Whereas taxonomic (from here onwards: categorical) relatedness between target and distractor slows down naming (e.g., Glaser and Döngelhoff, 1984; Schriefers et al., 1990; Damian and Martin, 1999), findings for associatively related items are more mixed, rendering either null results, or facilitation. Lupker (1979) compared the effects of categorical and associative relations

in picture-word interference. While he found that categorical relations caused interference, he did not observe any effect of associative relations. In a second experiment, he tested if there were additive effects of categorical and associative relationships by comparing categorically related distractors with distractors that were both categorically and associatively related, but both types of distractors caused the same amount of interference. Subsequently, however, facilitatory effects of associative relationships were reported. La Heij et al. (1990) manipulated the association strength for categorical distractors. While they found interference for weakly associated categorical distractors, they did not observe any effects for strongly associated categorical distractors. This pattern was explained with the assumption that categorical overlap causes interference whereas an associative relationship generates facilitation, resulting in a null result if both types of relationship are combined. Associatively based facilitation was subsequently demonstrated more explicitly: Alario et al. (2000) contrasted the effects of categorically, non-associated distractors with those of associated, non-categorically related distractors (e.g., dog-bone). They reported interference effects for categorically related distractors and facilitatory effects for associatively related distractors (although possibly following slightly different time courses; this aspect is less relevant for present purposes). Abdel Rahman and Melinger (2007, Experiment 3) found the same pattern, with interference for categorically related distractors and facilitation for associatively related distractors.

The dissociation between associatively and categorically related distractors in PWI was recently further explored via brain imaging by de Zubicaray et al. (2013). They contrasted categorical with “thematic” relations, i.e., associations caused by a common theme (e.g., *mouse* and *cheese* being related through an “eating” event). Behaviorally, they observed facilitation from thematically related distractors, and interference from categorically related distractors, relative to an unrelated condition. In the fMRI data, both types of relationship caused deactivations in the mid portion of the left middle temporal gyrus, but categorical relations also involved the posterior left MTG, while thematic relations involved the left angular gyrus. This finding underscores the assumption that categorical and thematic relations are processed differently.

To sum up, the available evidence suggests that categorical and associative relations cause different effects and should therefore be carefully controlled in studies on picture-word interference. This, however, is not always the case, and “mixed” stimuli might at least partially account for the polarity reversal of semantic effects in PWI tasks outlined in the previous section. Potentially, the categorical relationship asserts itself more strongly in the visible condition and the associative relationship more strongly in the masked condition, making the net effect appear like a polarity reversal of the categorical effect. Note that this result would not be necessarily at odds with the response exclusion hypothesis: This account predicts that masking prevents distractors from entering a buffer, hence, masking eliminates the interference component. Unlike the explanation offered by Finkbeiner and Caramazza (2006), however, we suggest that different items might be responsible for interference in the visible condition and facilitation in the masked condition. Unfortunately, Finkbeiner

and Caramazza (2006) do not provide a list of their items, but Dhooge and Hartsuiker (2010) do. Examining their items, one sees that they used both weakly associated target-distractor pairs (e.g., *spoon-knife*; *monkey -bear*) as well as strongly associated items (*lion-tiger*; *apple-pear*), and furthermore, also pairs that can be thought of as part and whole (*farm-shed*; *pot-lid*).

In order to carefully tease apart the potential influence of categorical and associative relations in both masked and visible distractor presentation, we carried out an experiment which varied both types of relatedness separately. This led to three related experimental conditions: one in which distractors and targets were categorically but not associatively related, one in which they were associatively but not categorically related, and one in which they were both categorically and associatively related. Associative relatedness was determined with subjective ratings in a pre-study, as well as *post-hoc* via participants’ ratings. If it is true that the polarity reversal is mainly due to the associative (facilitatory) component having a stronger effect with masked distractors, and the categorical (interfering) component emerging stronger with non-masked presentation, we should observe the strongest polarity reversal for the combined items. If our hypothesis is correct, the categorical relation mainly causes the interference in visible presentation and the associative relation generates the facilitation in masked presentation. For the categorically related items (without additional association), we should hence observe interference in the visible condition and a null effect in masked presentation. Finally, for associatively related items, we should see an increase of the facilitation effect in masked presentation conditions.

## METHODS

### PARTICIPANTS

Forty-eight students (28 women) from Humboldt-University Berlin took part in the experiment and were paid for their participation. Their mean age was 25 years. All participants were native speakers of German.

### MATERIALS

Twenty line drawings of common objects were used as targets. For each picture (e.g., *orange*), three distractor words were selected: a semantically related word (i.e., a category coordinate, e.g., *banana*), an associatively related word (i.e., a related word from a different category, e.g., *juice*), and a semantically and associatively related word (e.g., *lemon*). Distractor words in the three different conditions were matched on length and frequency. We created three corresponding unrelated conditions by recombining the related distractors within each relatedness type with different pictures. Therefore, for each of the three relatedness types (categorically related, associatively related, combined), the same pictures and words were used in both the related and the unrelated condition. Each participant saw a target word in all six conditions (three critical conditions and three control conditions). See Appendix for a list of all combinations. A different randomization was created for each participant to avoid order effects.

Strength of associative relations was established pre- and *post-hoc*. In a pre-study, we had investigated the association strengths of 22 line drawings, asking 24 participants to rate the association

strength for a target word and the intended distractor word on a scale from 1 (“not related at all”) to 7 (“very strongly related”). These participants came from the same pool as those who participated in the actual experiment, but none were in the experiment. **Table 1** presents the results from the pre-study. As intended, items in the associative condition were rated to be more strongly related than in the categorical condition,  $t_{(21)} = 6.85$ ,  $p < 0.001$ , and similarly, items in the combined condition were rated to be more strongly related than in the categorical condition,  $t_{(21)} = 10.86$ ,  $p < 0.001$ . Associative and combined conditions did not differ in association strength,  $t < 1$ , and none of the baseline conditions differed from each other, all  $ps > 0.20$ .

We also carried out the same rating study, using only the 20 line drawings eventually used in the experiment, after the PWI task and visibility tests (outlined below under the header “Rating study”). **Table 1** presents the results from the *post-hoc* rating as well. The *post-hoc* ratings confirmed the pilot results: the associative and the combined conditions had stronger association strengths than the categorical condition,  $t_{(19)} = 5.51$ ,  $p < 0.001$  and  $t_{(19)} = 8.40$ ,  $p < 0.001$ , respectively, whereas the combined and the associative conditions did not differ in their association strength,  $t < 1$ . The three baseline conditions did not differ significantly from one another [baseline categorical vs. baseline associative:  $t_{(19)} = 1.55$ ,  $p = 0.07$ ; baseline categorical vs. baseline combined:  $t < 1$ ; baseline associative vs. baseline combined:  $t_{(19)} = 1.27$ ,  $p = 0.11$ ].

For the visibility assessment (lexical decision task; see below), the 60 distractors as described above were used as “word” stimuli. Sixty non-words were created by using existing words and replacing one or two letters. These letter changes could occur in any position in the word, and care was taken to change letters in each position equally often. Non-words were matched in length to the word targets. This resulted in 120 target stimuli for the LDT (60 words and 60 non-words).

## PROCEDURE AND APPARATUS

Participants carried out three different tasks: the PWI task, a lexical decision task and a rating task. Within the PWI task, the order of the blocks corresponding to presentation mode (non-masked vs. masked) was counterbalanced across participants. An entire testing session lasted about an hour. PWI and lexical decision tasks were programmed and run with Presentation (NeuroBehavioral Systems). The rating task was carried out with Excel from Microsoft Office. All tasks were presented on a 19” CRT monitor with a refresh rate of 75 Hz (13.33 ms).

**Table 1 | Association strength (means and standard deviations) from the pilot and *post-hoc* experimental ratings.**

	Categorical	Associative	Combined
<b>PILOT RATINGS</b>			
Related	3.96 (0.63)	5.63 (1.09)	5.55 (0.59)
Unrelated	1.45 (0.39)	1.69 (0.69)	1.60 (0.65)
<b>POST-HOC RATINGS</b>			
Related	4.31 (0.67)	5.83 (1.18)	5.67 (0.59)
Unrelated	1.49 (0.41)	1.29 (0.20)	1.56 (0.70)

## Picture-word interference task

Participants were instructed to name objects presented on the computer monitor as quickly and accurately as possible. Trial timing and masking procedure were adopted from Finkbeiner and Caramazza’s (2006) work, as follows: in the *non-masked presentation mode*, a trial started with a fixation cross that was presented for 500 ms in the center of the screen. The distractor word was presented centered on the screen for 53 ms (4 screen refresh cycles). Picture and word were then presented together for 2000 ms. Participants’ responses triggered a voice key, and latencies were measured relative to picture onset. In the *masked presentation mode*, a trial started with a forward mask (#####) presented for 500 ms. The word was presented centered on the screen for 53 ms. It was replaced by the picture and a non-pronounceable mask consisting of a string of 10 consonants presented in the same location as the distractor word. Picture and mask were presented together for 2000 ms. Participants’ responses triggered a voice key, and latencies were measured relative to picture onset. The use of a consonant string as a backward mask was motivated by Finkbeiner and Caramazza (2006) who refer to findings having shown its particular effectiveness in eliminating phonological priming effects.

## Lexical decision task

Because the aim was to assess visibility of distractors in the masked presentation mode of the PWI task, the trial structure was chosen to be very similar. A forward mask (#####) was presented for 500 ms centered on the screen, followed by a letter string presented for 53 ms. The letter string was replaced by a backward mask consisting of a string of 10 consonants presented in the same location as the distractor word. The mask stayed in place until the participant had made a response. Participants were instructed to decide whether or not the briefly presented string was an existing word of their language. They were encouraged to make a guess if they felt they had not seen a stimulus at all. The 120 target stimuli (60 words and 60 non-words) were randomly intermixed, with a new sequence for each participant.

## Rating study

The names of the 20 target pictures and their related distractors were presented as pairs. For each pair, participants were instructed to indicate how strong the association between the two concepts was, using a scale from 1 (“not related”) to 7 (“strongly related”). Items were divided into six blocks, with a given target word occurring only once per block. Each relatedness condition (categorical, associative, combined) and their respective baselines occurred equally often within a given block; the assignment of a particular item in a given condition to a block was counterbalanced across lists. Six different randomizations were created.

## RESULTS

### PICTURE-WORD INTERFERENCE TASK

Fifty-three observations (0.5% of the data) had to be removed due to script errors. Latencies on trials with errors (4.8%) as well as latencies that differed more than three standard deviations from a participant’s conditional mean (1.1%) were excluded

from the analysis. **Table 2** presents mean reaction times and error percentages, split by presentation mode, relatedness and type of relatedness.

**Reaction times**

Latencies were analyzed with analyses of variance (ANOVAs), with the factors presentation mode (non-masked vs. masked), relatedness (related vs. unrelated), and type of relatedness (categorical, associative, combined). The main effect of presentation mode was significant,  $F_{1(1, 47)} = 7.28, MSE = 14, 713, p = 0.010; F_{2(1, 19)} = 49.44, MSE = 782, p < 0.001$ , with 29 ms faster reaction times for the masked than the non-masked condition. The effect of relatedness was also significant, with 13 ms slower reaction times for related than for unrelated items,  $F_{1(1, 47)} = 5.70, MSE = 1248, p = 0.021, F_{2(1, 19)} = 5.08, MSE = 625, p = 0.036$ . The main effect of type of relatedness was not significant,  $F_1 = 1.22, p = 3.01; F_2 = 1.13, p = 0.333$ . The main effects were qualified by a significant interaction of relatedness by presentation mode,  $F_{1(1, 47)} = 7.97, MSE = 1421, p = 0.007; F_{2(1, 19)} = 9.34, MSE = 431, p = 0.006$ , an interaction of type of relatedness by presentation mode,  $F_{1(2, 94)} = 3.88, MSE = 774, p = 0.024; F_{2(2, 38)} = 3.64, MSE = 383, p = 0.040$ , and an interaction of relatedness by type of relatedness [ $F_{1(2, 94)} = 6.57, MSE = 1020, p = 0.002; F_{2(2, 38)} = 3.93, MSE = 734, p = 0.028$ ]. The three-way interaction of presentation mode, relatedness, and type of relatedness was also significant,  $F_{1(2, 94)} = 10.97, MSE = 1160, p < 0.001, F_{2(2, 38)} = 11.15, MSE = 480, p < 0.001$ .

In order to further investigate the significant three-way interaction between presentation mode, relatedness, and type of relatedness, we conducted two additional analyses, as outlined below.

**Simple effects of presentation mode.** First, we investigated effects of relatedness and type of relatedness for each level of presentation mode (non-masked, masked) separately, an analysis which highlights the overall effects of distractor presentation mode on relatedness effects.

For the *non-masked presentation mode*, there was a main effect of relatedness, with slower reaction times for related than for unrelated trials,  $F_{1(1, 47)} = 11.04, MSE = 1649, p = 0.002;$

$F_{2(1, 19)} = 10.74, MSE = 669, p = 0.004$ , an effect of type of relatedness which was significant by participants, but only marginally so by items, [ $F_{1(2, 94)} = 3.99, MSE = 1046, p = 0.022; F_{2(2, 38)} = 2.82, MSE = 795, p = 0.072$ ], and a significant interaction of relatedness and type of relatedness,  $F_{1(2, 94)} = 13.71, MSE = 1302, p < 0.001; F_{2(2, 38)} = 7.64, MSE = 992, p = 0.002$ . We further explored the interaction of relatedness and type of relatedness via paired *t*-tests. The 38 ms interference effect for categorically related items was significant,  $t_{1(47)} = 5.34, p < 0.001; t_{2(19)} = 4.66, p < 0.001; 95\% CI [24, 52]$ . The 15 ms facilitation effect for associatively related items was significant by participants only,  $t_{1(47)} = 2.31, p = 0.025; t_{2(19)} = 1.65, p = 0.116; 95\% CI [2, 27]$ . The 25 ms interference effect for combined items was significant,  $t_{1(47)} = 2.64, p = 0.011; t_{2(19)} = 2.12, p = 0.047; 95\% CI [6, 43]$ .

For the *masked presentation mode*, neither relatedness nor type of relatedness was significant,  $F_1$  and  $F_2 < 1$ . The interaction between relatedness and type of relatedness was not significant by participants,  $F_{1(2, 94)} = 1.80, MSE = 878, p = 0.172$ , and marginally significant by items,  $F_{2(2, 38)} = 2.96, MSE = 222, p = 0.064$ .

**Simple effects of type of relatedness.** Second, we focused on the variable type of relatedness, and investigated for each level (categorical, associative, combined) separately whether presentation mode (non-masked, masked) affected relatedness effects. This analysis specifically aims to identify potential polarity reversals in relatedness effects, as suggested by Finkbeiner and Caramazza (2006) and Dhooge and Hartsuiker (2010).

For *categorically related* items, the effect of presentation mode was significant,  $F_{1(1,47)} = 9.64, MSE = 4539, p = 0.003; F_{2(1,19)} = 51.61, MSE = 295, p < 0.001$ , and so was the effect of relatedness,  $F_{1(1,47)} = 7.99, MSE = 823, p = 0.007; F_{2(1,19)} = 6.73, MSE = 430, p = 0.018$ . Mode and relatedness interacted with each other,  $F_{1(1,47)} = 30.25, MSE = 670, p < 0.001; F_{2(1,19)} = 30.42, MSE = 274, p < 0.001$ . Paired *t*-tests showed the highly significant interference effect of 38 ms for non-masked distractors already reported in the previous section,  $t_{1(47)} = 5.34, p < 0.001; t_{2(19)} = 4.66, p < 0.001; 95\% CI [24, 52]$ . The 11 ms facilitation effect for masked distractors was marginally

**Table 2 | Reaction times (in milliseconds) and errors (in percent) by presentation mode (visible vs. masked distractor presentation), relatedness (related vs. unrelated), and type of relatedness (categorical, associative, combined).**

	Non-masked distractors			Masked distractors		
	Categorical	Associative	Combined	Categorical	Associative	Combined
<b>REACTION TIMES</b>						
Related	711 (132)	672 (119)	702 (77)	654 (77)	662 (78)	660 (78)
Unrelated	673 (116)	687 (125)	677 (80)	665 (80)	660 (89)	656 (86)
Difference	38	-15	25	-11	2	4
<b>ERRORS</b>						
Related	5.4 (5.4)	5.5 (5.7)	7.0 (5.4)	4.2 (4.9)	3.2 (4.4)	4.8 (5.3)
Unrelated	5.3 (6.8)	5.6 (6.0)	5.4 (6.8)	3.8 (3.8)	4.3 (4.3)	3.1 (4.1)
Difference	-0.1	0.1	1.6	0.4	-1.1	1.7

Standard deviations in parentheses.

significant,  $t_{1(47)} = 1.91$ ,  $p = 0.063$ ,  $t_{2(19)} = 1.79$ ,  $p = 0.090$ ;  $t_{2(19)} = 1.86$ ,  $p = 0.078$ ; 95% CI [1, 23].

For *associated* items, the effect of presentation mode was marginally significant,  $F_{1(1, 47)} = 2.95$ ,  $MSE = 4801$ ,  $p = 0.092$ ;  $F_{2(1, 19)} = 8.90$ ,  $MSE = 511$ ,  $p = 0.008$ . Relatedness was not significant,  $F_1 = 2.17$ ,  $p = 0.148$ ;  $F_2 = 1.35$ ,  $p = 0.259$ , nor was the mode  $\times$  relatedness interaction,  $F_1 = 1.79$ ,  $p = 0.188$ ;  $F_2 = 1.60$ ,  $p = 0.221$ .

For *combined* items, the effect of presentation mode was significant,  $F_{1(1, 47)} = 8.12$ ,  $MSE = 3985$ ,  $p = 0.006$ ;  $F_{2(1, 19)} = 20.98$ ,  $MSE = 605$ ,  $p < 0.001$ . The effect of relatedness was significant by participants,  $F_{1(1, 47)} = 5.81$ ,  $MSE = 1341$ ,  $p = 0.020$ , and marginally significant by items,  $F_{2(1, 19)} = 3.56$ ,  $MSE = 875$ ,  $p = 0.075$ . The mode  $\times$  relatedness interaction was not significant,  $F_1 = 1.80$ ,  $p = 0.186$ ;  $F_2 < 1$ ,  $p = 0.344$ .

### Error rates

Error scores are shown in **Table 2**, and were submitted to logistic regression analysis with the factors presentation mode (non-masked vs. masked), relatedness (related vs. unrelated), and type of relatedness (categorical, associative, combined). The results showed a significant effect of presentation mode,  $Wald Z = 2.52$ ,  $p = 0.012$ , with 1.8% more errors in the non-masked than the masked condition. Furthermore, the interaction between relatedness and type of relatedness was significant,  $Wald Z = -2.16$ ,  $p = 0.031$ . Simple effects analysis showed no effect of relatedness for the “categorical” and “associative” conditions,  $Wald Z = -0.24$ ,  $p = 0.811$ , and  $Wald Z = 0.69$ ,  $p = 0.491$  respectively, but a significant effect for the “combined” condition,  $Wald Z = 2.27$ ,  $p = 0.024$ , with 1.6% more errors in the related than the unrelated condition. All other main effects or interactions were not significant,  $Wald Z \leq 1.73$ ,  $p \geq 0.083$ .

### INTERIM SUMMARY

Overall, the latency results from the “non-masked” presentation mode replicated an existing pattern in previous research: a strong categorical semantic interference effect contrasted with a weaker associative facilitation effect. The combined effect of categorical and associative relatedness was almost perfectly additive. In the “masked” presentation mode, effects were much weaker. Most relevant is the 11 ms facilitatory effect in the categorically related condition, which compares with parallel effects in previous research of 32 ms (in Finkbeiner and Caramazza, 2006, Experiment 1) and 12 ms (in Dhooge and Hartsuiker, 2010, Experiment 2). This effect just failed to reach conventional significance (see Section Lexical Decision Task below for further analysis) but numerically, the polarity reversal of the semantic effect dependent on presence or absence of distractor masking which was highlighted by the earlier studies also emerges in the present study. In the associatively and combined relatedness conditions, very little effects emerge under masked conditions.

One possible reason why the masked effects are so small is that the masking procedure may have been too efficient, eliminating (or substantially reducing) distractor processing. The results from the post-experimental visibility test reported in the following section allow some insight into this issue.

### LEXICAL DECISION TASK

**Table 3** summarizes the accuracy results from the lexical decision task. Overall, 71.3% of the masked words were correctly recognized, with an overall false alarm rate (i.e., “word” responses to non-words) of 33.8%. For each participant, we calculated a *d-prime* ( $d'$ ) score based on the hits and false alarm rates for words, using the formula for R suggested by Pallier (2002).  $D'$  scores ranged from 0.25 to 2.93, with a mean of 1.25 and a standard deviation of 0.67, and differed significantly from zero,  $t_{(46)} = 12.80$ ,  $p < 0.001$ . This implies that the masking procedure did not fully prevent distractor visibility.

The latter result may seem surprising, given that we chose our masking procedure to be very similar (in terms of prime durations, nature of mask, etc.) to those used by Finkbeiner and Caramazza (2006) and Dhooge and Hartsuiker (2010). Finkbeiner and Caramazza did not include formal visibility assessments in their study so it is difficult to assess whether their masking had been more stringent than ours. Dhooge and Hartsuiker included, in their second experiment, a visibility test consisting of presence/absence judgments on masked prime, but merely reported that “no distractors were reported” (p. 884). It is worth noting (see our point in the Introduction) that our visibility test possibly overestimated participants’ true ability to access distractor identity in the main experiment. Nevertheless, it is clear from the lexical decision results that distractors were not perfectly masked in our study.  $D'$  scores computed for each participant showed substantial variability, with some participants essentially unable to identify the distractors (those with a  $d'$  close to zero) and others evidently finding it quite easy (those with the highest  $d'$  scores).

The high variability in prime visibility in our study offers a possible explanation for the weak masked effects. According to Finkbeiner and Caramazza (2006), visible distractors will cause interference whereas masked and therefore unconsciously processed distractors will generate semantic facilitation. Perhaps the less-than-perfect masking in our experiment and the associated variability in individual  $d'$  scores (see above) resulted in participants with good visibility generating interference whereas those with poor visibility caused facilitation. If so, the direction of the semantic effect for a particular participant should be predictable based that participant’s ability to see distractors in our post-test. Note that Piai et al. (2012) competition threshold hypothesis by contrast stipulates that masked primes, independently of how well they can be perceived by an individual, should generally create only weak activation which is rarely powerful enough to cross the threshold to engage in competition with target

**Table 3 | Accuracy of lexical decision task, by condition (categorical, associative, combined).**

Condition	Correctly recognized
Categorical	70.9% (17%)
Associative	68.8% (18%)
Combined	74.3% (17%)

*Standard deviation in parentheses.*

name retrieval. Hence, the masked semantic effect in our study should be independent of variability in distractor visibility, as indicated by  $d'$ .

To investigate this issue, we focused on the “categorically related” condition (predictions for the other two types of relatedness are more difficult, as net results might be a combination of interference and facilitation). **Figure 1** shows the masked categorical effect, conceptualized as a percentage change relative to the unrelated baseline condition, and dependent on  $d'$  scores (dots represent individual participants). As can be seen,  $d'$  scores are relatively uniformly distributed within the range, and there is no evident relationship between the experimental effect and individual visibility. A linear regression showed very little effect,  $R^2 = 0.016$ ,  $\beta = -0.13$ ,  $SE = 0.15$ ,  $F_{(1, 45)} < 1$ ,  $p = 0.396$ . In other words, participants with low and high ability to consciously perceive the masked distractors showed very similar experimental effects.

Nevertheless, it is worth noting that the participant with the highest  $d'$  score (2.93) showed the largest semantic interference effect (of  $-10\%$ , or  $-56$  ms; this participant is in the lower right corner of the Figure). Possibly, this participant experienced particularly good visibility of masked distractors in the experiment, which resulted in a correspondingly large interference effect. When this participant was excluded from the analysis as a potential outlier, the overall masked categorical *facilitation* effect rose to 13 ms (cf. Dhooze and Hartsuiker's, 2010, 12 ms effect in the equivalent condition), and was now statistically significant,  $t_{(45)} = 2.19$ ,  $p = 0.034$ ;  $t_{(19)} = 2.11$ ,  $p = 0.048$ . A linear regression between the categorical effect and individual  $d$ -primes, again with this participant excluded, now resulted in an almost perfectly

flat trend line,  $R^2 < 0.001$ ,  $\beta = -0.02$ ,  $SE = 0.15$ ,  $F_{(1, 44)} < 1$ ,  $p = 0.874$ .

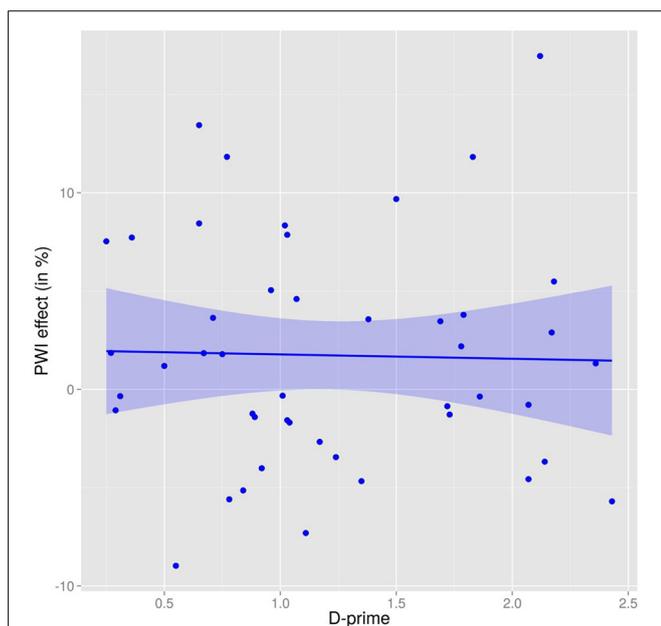
We conclude that despite considerable variability in participants' ability to consciously perceive the masked distractors, categorical relatedness effects in our experiment are clearly not dependent on visibility.

## RESPONSE DISTRIBUTION LATENCIES

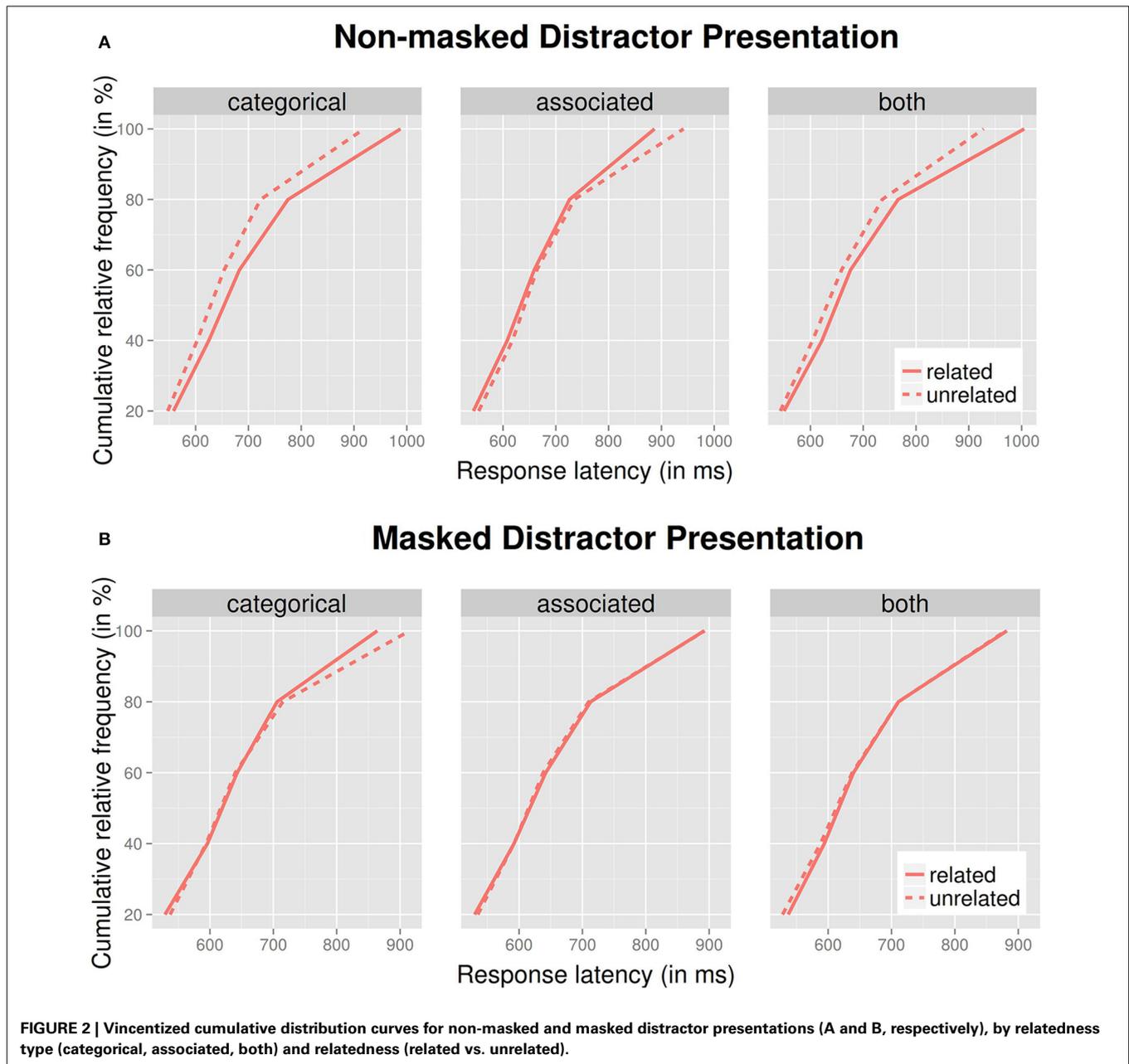
In picture-word interference tasks, mean latencies are generally shorter in masked than in non-masked conditions (Finkbeiner and Caramazza, 2006; Dhooze and Hartsuiker, 2010; Piai et al., 2012). This was also the case in our experiment, reflected in a highly significant main effect of presentation mode. Piai et al. (2012, p. 621) put forward the following line of reasoning: It is plausible to assume that, given that participants are faster under masked conditions, the shortest latencies within the response time distribution should reflect those trials on which the masking procedure was effective, whereas the longer RTs are those in which distractors are not well masked. If so, conditional means of the masked condition might represent a mixture of trials, with the shortest RTs showing facilitation and the longest ones exhibiting interference (and an overall weak effect, as was found in our experiment). We investigated this possibility via computation of Vincentized cumulative distribution curves (Ratcliff, 1979): for each participant and condition, rank-ordered latencies were divided into 20% quantiles, and mean latencies were computed for each quintile. These were then averaged across participants, which preserves the shapes of individuals' latency distributions (cf. Woodworth and Schlosberg, 1954). An analysis of this type provides information about the degree of uniformity with which an effect affects the spectrum of response latencies.

**Figure 2A** shows distribution curves for the “non-masked” presentation mode, and for all three types of relatedness (note that untrimmed latencies were used to generate **Figure 2**; see Heathcote et al., 1991). As expected from previous research (e.g., Roelofs, 2008), effects were spread out across the entire spectrum for the “categorical” and “combined” condition. The facilitatory effect for the “associated” condition emerged to a larger extent in the slower quintile. **Figure 2B** shows curves for the “masked” presentation mode. Intriguingly, the semantic facilitation effect weakly present in the means (cf. **Table 2**) predominantly emerged in the slowest (rightmost) quintile. This is clearly contrary to what one might predict on the assumption that well-masked (and hence fast) RTs should exhibit facilitation whereas poorly masked RTs show interference.

The manner in which Vincentiles are typically computed (for each participant individually, and then averaged) means that each participant equally contributes to all quantiles. Hence, the shown values for, say, the rightmost (slowest) quintile shown in **Figure 2** represent the average of all participants' slowest quintile. Assume a scenario in which a subset of participants had better distractor visibility than others, resulting in slower latencies and semantic interference, whereas a different subset had poor visibility and hence showed faster latencies and semantic facilitation. Because in **Figure 2** all participants are equally represented, this should emerge as an effect spread out across the spectrum (or perhaps no effect at all), but clearly not what is evident in **Figure 2** (an



**FIGURE 1 |** Picture word interference effect (masked, for “categorically related” distractors; as percent relative to unrelated baseline) dependent on  $d$ -prime in lexical decision task. Dots represent individual participants.



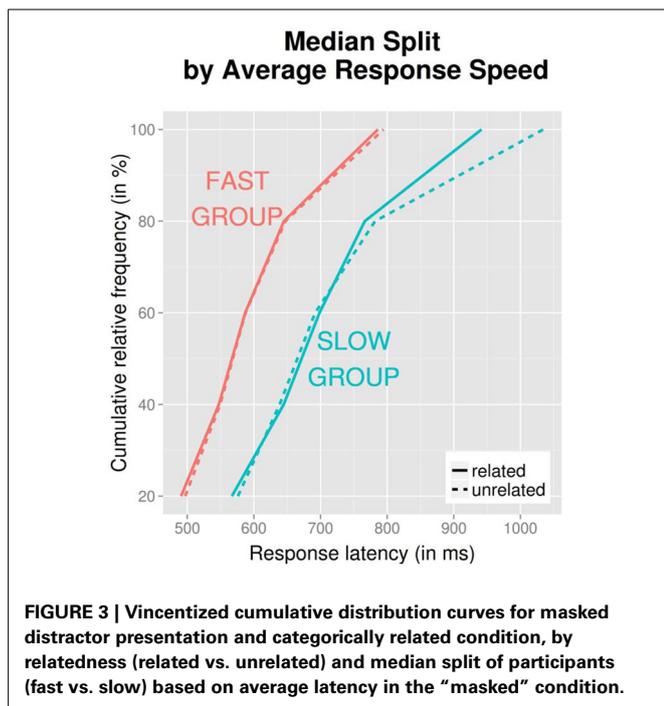
asymmetry). A remaining possibility is that slower subjects, and for those individuals, slower latencies, carried the semantic facilitation effect. To look at this possibility, we performed a median split of participants into a “fast” and “slow” group, based on average latencies in the “masked” condition, and computed quintiles for the categorically related condition for each group. **Figure 3** shows the results. Indeed, it appears that the semantic facilitation effect predominately stems from slower participants (and within the “slow” group, from the slowest quintile).

Given the considerable individual variability in participants’ ability to recognize masked primes (see Section Lexical Decision Task), could it be that visibility is associated with slow latencies? In other words, is there an association between overall response

speed and prime visibility (perhaps because prime processing slows participants down)? A further analysis suggested that this was not the case: a linear regression between overall response speed as assessed by  $d'$  showed no such association,  $R^2 = 0.002$ ,  $\beta = 0.04$ ,  $SE = 0.15$ ,  $F_{(1, 45)} < 1$ ,  $p = 0.747$ .

## DISCUSSION

Recent studies (Finkbeiner and Caramazza, 2006; Dhooge and Hartsuiker, 2010) have suggested that the “classic” semantic interference effect found in numerous picture-word experiments reverses into a facilitatory effect when distractors are masked such that visibility is impaired. This “polarity reversal” has been



interpreted as evidence for the “response exclusion hypothesis” according to which semantic interference effects do not reflect, as commonly assumed, lexical competition between target name and distractor, but rather arise at a post-lexical response buffer level. Masking of distractors presumably prevents them from occupying the response buffer and hence from generating semantic interference. At the same time, masking still allows for some unconscious distractor processing, resulting in conceptually based facilitation. However, other interpretations of the polarity reversal pattern are possible (Piai et al., 2012): perhaps competition only takes place when potential competitors are strongly enough activated (i.e., cross a “competition threshold”). If so, masking, rather than rendering distractors unconscious, simply renders them too weak to engage in competition with target retrieval.

We report an experiment which aimed to contribute to the debate in the following way. Related or unrelated distractor words were either presented such that they were easy to identify, or masked such that they were more difficult to perceive. We additionally manipulated the type of relatedness between distractor and target: they could be either categorically related, associated, or categorically as well as associatively related. Our reasoning was that existing studies may have mixed different types of relatedness, and that semantic interference (with non-masked distractors) and facilitation (with masked distractors) might have arisen from different sets of items, namely categorically and associated pairs, respectively. If so, then the relatedness effects dependent on type of relatedness should emerge differentially with non-masked and masked presentations, and pairs which are categorically as well as associatively related should show the strongest polarity reversal. Furthermore, we added a post-experimental visibility test which allowed some insight into individuals’ differential

ability to perceive masked distractor words. According to Dhooge and Hartsuiker (2010) and Finkbeiner and Caramazza (2006), the directionality of semantic effects should primarily depend on distractor visibility (only visible distractors should be able to enter the “response buffer” and generate interference; invisible distractors should result in facilitation). By contrast, according to Piai et al. (2012) competition threshold hypothesis, masked distractors should largely evoke only weak distractor processing, hence semantic interference effects should generally not induce interference except under certain circumstances (see Piai et al. for details).

For distractors which were presented non-masked and were hence clearly visible to participants, our results showed substantial categorical interference (38 ms), as well as facilitation effect arising from associative relatedness (15 ms). This pattern is generally in line with previous studies on the effects of categorical vs. associative relatedness in PWI tasks (e.g., La Heij et al., 1990; Alario et al., 2000). For distractors which were categorically as well as associatively related, we found an almost perfectly additive pattern, with an empirical interference effect of 25 ms which deviated only 2 ms from the prediction based on additivity. Statistical additivity might imply, based on “additive factors logic” (Sternberg, 1969), that the two effects arise at different processing levels. This is indeed a possibility in line with previous claims. For instance, Cutting and Ferreira (1999) postulated a cascaded model of spoken word production in which phonological word forms are linked to each other via associative links. The broader claim is that lexical entries, at a sub-semantic level, might be organized according to associative relationships (e.g., Fodor, 1983; Shelton and Martin, 1992), perhaps representing co-occurrence in natural discourse (Spence and Owens, 1990). A theoretical model in which semantic interference in PWI reflects lexical-semantic competition whereas associative priming arises due to interlinked word forms could account for our findings from the non-masked presentation condition.

When distractors were briefly presented and sandwiched between forward and backward masks, effects were considerably weaker. For categorically related distractors, the “polarity reversal” predicted from the earlier studies was indeed found, but semantic facilitation in the masked presentation mode was small (11 ms) and failed to reach conventional significance. Masked effects for the associative and combined conditions were not significant. These results allow us to reject the possibility—outlined above—that previous instances of “polarity reversal” may have arisen due to differential sets of items with different types of relationship. Specifically, the predicted strong polarity reversal effect for “combined” items was clearly not present in the current results.

Results from the post-experimental visibility test allowed some further insight into the nature of the shown effects. The overall weak effects under the masked presentation mode might be attributed to too powerful masking: if masks prevent (or largely eliminate) distractor processing, then null or only small effects would be predicted. To the contrary, results from our visibility test showed (a) an overall surprisingly high ability of individuals to recognize the masked letter strings; (b) substantial individual variability in their ability to do so.  $D'$  scores ranged from 0.25 to

2.93, and as visible in **Figure 1** were relatively uniformly spread out within that range. This renders it unlikely that overly strict masking may have caused the weak masked priming effects in our study. An alternative is that masking was insufficient, and indeed, the REH might predict that participants with poor visibility generate semantic facilitation whereas those with good visibility cause semantic interference, plausibly resulting in a small net effect when averaged. However, our analysis which looked at categorically based masked effects in relation to individual differences (reported in **Figure 1**) clearly showed that this was not the case: visibility did not seem to affect polarity, nor size, of the masked semantic effects.

Overall, we interpret these results as more in line with the “competition threshold” claim introduced by Piai et al. (2012) than the “response exclusion hypothesis” favored by Finkbeiner and Caramazza (2006) and Dhooge and Hartsuiker (2010). According to the latter, given the good distractor visibility in the masked condition for at least some (perhaps most) of our participants, distractors should have been prepared for articulation in the response buffer, and semantic interference should have arisen. The fact that **Figure 1** showed no clear dependence of semantic effects on distractor visibility argues against the possibility that conscious processing of distractors is the primary prerequisite for semantic interference in PWI tasks. Piai et al.’s competition threshold can more easily accommodate our results because according to that claim, masking generally reduces activation strength of distractors, and hence independent of how good individuals are at perceiving masked distractors, the overall pattern should be semantic facilitation, or perhaps a null finding.

As highlighted in the Introduction, the competition threshold view makes it difficult to generate precise *a priori* predictions about under which circumstances semantic interference or facilitation effects in PWI tasks should be obtained. To exemplify, **Figure 1** showed that the individual with the highest  $d'$  score showed the strongest semantic interference effect. Perhaps for this individual, distractor visibility was high enough that on most or all of trials, distractors evoked strong enough activation to cross the competition threshold and engage in competition with picture naming. Although this is not implausible, it would clearly be preferable to be able to identify distractor strength—relative to the purported threshold—beforehand in order to generate predictions about the directionality of semantic effects.

We additionally analyzed latencies via cumulative response time distribution plots (see **Figure 2**), and an unexpected pattern which emerged was that the weak semantic facilitation effect in the masked condition mainly emerged for slower participants, and almost exclusively in the slowest quintile of latencies. Although it is common in experimental psychology that effects are more pronounced for slower than for faster latencies, the extreme nature of the pattern found here strikes us as unusual and not easily explained. The lack of an association between overall speed of response and visibility scores certainly argues against the possibility that the slower participants for whom the semantic facilitation effect emerged were those with particularly high distractor visibility. In research on cognitive inhibition which employed response time distribution analyses, the suggestion has been made that under some circumstances, inhibitory effects

may emerge only in slow quintiles because inhibition takes some time to develop (Ridderinkhof, 2002). When applying this line of reasoning to our findings, one would have to speculate that semantic facilitation is so slow to develop that it only emerges in the slower quintiles. But given that picture naming is a conceptually driven task, this suggestion makes little sense—conceptually based effects should emerge swiftly, rather than slowly. Further research is required to resolve this issue.

Overall, our findings add to the extant literature on “polarity reversals” of semantic effects in picture-word interference tasks, and suggest that these effects are genuine and not due to uncontrolled properties of stimuli (such as type of relatedness). However, our findings suggest that visibility of the distractor *per se* is not the primary determinant of whether a semantic effect is positive or negative: visibility tests implied a wide range in individuals’ ability to perceive masked distractors, yet distractor masking generally resulted in weak semantic facilitation. This pattern is more in line with the notion of a “competition threshold” according to which masking generally, and independent of visibility, generally reduces distractor activation strength such that it prevents competition between distractor and target processing. Further research should illuminate the connection between conscious visibility and distractor processing more explicitly, perhaps via studies in which distractor presentation duration is systematically manipulated, and visibility associated with each particular distractor duration is assessed. The response exclusion hypothesis would predict semantic interference only for durations under which visibility tests show conscious access to distractor identity; for shorter durations, semantic facilitation should be found (which of course would disappear with too short durations). The competition threshold account predicts no systematic relation between visibility and polarity of the semantic effects in PWI tasks.

## ACKNOWLEDGMENTS

This research was supported by grant SP1126/4-1 from the Deutsche Forschungsgemeinschaft (DFG) to Katharina Spalek. We would like to thank Carsten Schlieve for technical support, and Hannah Bohle, Julia Knoepke and Annika Labrenz for their assistance in data collection.

## REFERENCES

- Abdel Rahman, R., and Melinger, A. (2007). When bees hamper the production of honey: lexical interference from associates in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 604–614. doi: 10.1037/0278-7393.33.3.604
- Alario, F.-X., Segui, J., and Ferrand, L. (2000). Semantic and associative priming in picture naming. *Q. J. Exp. Psychol. A* 53, 741–764. doi: 10.1080/027249800410535
- Costa, A., Alario, F.-X., and Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychon. Bull. Rev.* 12, 125–131. doi: 10.3758/BF03196357
- Cutting, J. C., and Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 318–344. doi: 10.1037/0278-7393.25.2.318
- Damian, M. F., and Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 345–361. doi: 10.1037/0278-7393.25.2.345
- de Zubicaray, G. I., Hanson, S., and McMahon, K. L. (2013). Differential processing of thematic and categorical conceptual relations in spoken word production. *J. Exp. Psychol. Gen.* 142, 131–142. doi: 10.1037/a0028717

- Dhooge, E., and Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: evidence for response exclusion. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 878–891. doi: 10.1037/a0019128
- Duffy, S. A., Morris, R. K., and Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *J. Mem. Lang.* 27, 429–446. doi: 10.1016/0749-596X(88)90066-6
- Finkbeiner, M., and Caramazza, A. (2006). Now you see it, now you don't: on turning semantic interference into facilitation in a stroop-like task. *Cortex* 6, 790–796. doi: 10.1016/S0010-9452(08)70419-2
- Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
- Glaser, W. R., and Dünghoff, F.-J. (1984). The time course of picture-word interference. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 640–654.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133
- Hantsch, A., Jescheniak, J. D., and Schriefers, H. (2005). Semantic competition between hierarchically related words during speech planning. *Mem. Cogn.* 33, 984–1000. doi: 10.3758/BF03193207
- Heathcote, A., Popiel, S. J., and Mewhort, D. J. (1991). Analysis of response time distributions: an example using the stroop task. *Psychol. Bull.* 109, 340–347. doi: 10.1037/0033-2909.109.2.340
- Jescheniak, J. D., and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 824–843. doi: 10.1037/0278-7393.20.4.824
- Kuipers, J. R., and La Heij, W. (2008). Semantic facilitation in category and action naming: testing the message-congruency account. *J. Mem. Lang.* 58, 123–139. doi: 10.1016/j.jml.2007.05.005
- La Heij, W., Dirks, J., and Kramer, P. (1990). Categorical interference and associative priming in picture naming. *Br. J. Psychol.* 81, 511–525. doi: 10.1111/j.2044-8295.1990.tb02376.x
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75. doi: 10.1017/S0140525X99001776
- Luce, R. D. (1959). *Individual Choice Behavior*. New York, NY: Wiley.
- Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York, NY: Oxford University Press.
- Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Mem. Cogn.* 7, 485–495. doi: 10.3758/BF03198265
- MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676
- Mahon, B., Costa, A., Peterson, R., Vargas, R. A., and Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 503–533. doi: 10.1037/0278-7393.33.3.503
- Miozzo, M., and Caramazza, A. (2003). When more is less: a counterintuitive effect of distractor frequency in the picture-word interference paradigm. *J. Exp. Psychol. Gen.* 132, 228–252. doi: 10.1037/0096-3445.132.2.228
- Mulatti, C., and Coltheart, M. (2012). Picture-word interference and the response-exclusion hypothesis. *Cortex* 48, 363–372. doi: 10.1016/j.cortex.2011.04.025
- Pallier, C. (2002). *Computing Discriminability and Bias with the R Software*. Available online at: <http://www.pallier.org/w/pmwiki.php/Main/Docs>
- Piai, V., Roelofs, A., and Schriefers, H. (2012). Distractor strength and selective attention in picture-naming performance. *Mem. Cogn.* 40, 614–627. doi: 10.3758/s13421-011-0171-3
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86, 446–461. doi: 10.1037/0033-2909.86.3.446
- Ridderinkhof, K. R. (2002). “Activation and suppression in conflict tasks: empirical clarification through distributional analyses,” in *Attention and Performance XIX: Common Mechanisms in Perception and Action*, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 494–519.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F
- Roelofs, A. (2008). Dynamics of the attentional control of word retrieval: analyses of response time distributions. *J. Exp. Psychol. Gen.* 137, 303–323. doi: 10.1037/0096-3445.137.2.303
- Rosinski, R. R., Golinkoff, R. M., and Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Dev.* 46, 247–253. doi: 10.2307/1128859
- Sailor, K., and Brooks, P. J. (2014). Do part-whole relations produce facilitation in the picture-word interference task? *Q. J. Exp. Psychol.* 67, 1768–1785. doi: 10.1080/17470218.2013.870589
- Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990). Exploring the time course of lexical access in production: picture-word interference studies. *J. Mem. Lang.* 29, 86–102. doi: 10.1016/0749-596X(90)90011-N
- Shelton, J. R., and Martin, R. C. (1992). How semantic is automatic semantic priming? *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 1191–1210. doi: 10.1037/0278-7393.18.6.1191
- Spalek, K., Damian, M. F., and Bölte, J. (2013). Is lexical selection in spoken word production competitive? Introduction to the special issue on lexical competition in language production. *Lang. Cogn. Processes* 28, 597–614. doi: 10.1080/01690965.2012.718088
- Spence, D. P., and Owens, K. C. (1990). Lexical co-occurrence and association strength. *J. Psycholinguist. Res.* 19, 317–330. doi: 10.1007/BF01074363
- Sternberg, S. (1969). The discovery of processing stages: extensions of Donders' method. *Acta Psychol.* 30, 276–315. doi: 10.1016/0001-6918(69)90055-9
- Woodworth, R. S., and Schlosberg, H. (1954). *Experimental Psychology*. New York, NY: Holt.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 June 2014; accepted: 29 September 2014; published online: 20 October 2014.

Citation: Damian MF and Spalek K (2014) Processing different kinds of semantic relations in picture-word interference with non-masked and masked distractors. *Front. Psychol.* 5:1183. doi: 10.3389/fpsyg.2014.01183

This article was submitted to *Language Sciences*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Damian and Spalek. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

**APPENDIX**

## Materials used in Experiment.

<b>Target picture</b>	<b>Categorical distractor</b>	<b>Associative distractor</b>	<b>Combined distractor</b>
Apfel (apple)	Birne (pear)	Traube (grape)	Kern (seed)
Bein (leg)	Arm (arm)	Kopf (head)	Hose (trousers)
Biene (bee)	Wespe (wasp)	Fliege (fly)	Honig (honey)
Esel (donkey)	Ochse (oxen)	Pferd (horse)	Knüppel (club)
Fluss (river)	Meer (sea)	Teich (pond)	Angler (fisher)
Fuß (foot)	Hand (hand)	Knie (knee)	Schuh (shoe)
Herz (heart)	Lunge (lung)	Darm (bowel)	Seele (soul)
Huhn (chicken)	Gans (goose)	Schwan (swan)	Ei (egg)
Kuh (cow)	Schwein (pig)	Reh (doe)	Milch (milk)
Löwe (lion)	Tiger (tiger)	Affe (monkey)	Mähne (mane)
Maus (mouse)	Katze (cat)	Hund (dog)	Käse (cheese)
Mond (moon)	Stern (star)	Komet (comet)	Nacht (night)
Nase (nose)	Mund (mouth)	Stirn (forehead)	Schnupfen (cold/flu)
Ohr (ear)	Auge (eye)	Kinn (chin)	Musik (music)
Orange (orange)	Zitrone (lemon)	Banane (banana)	Saft (juice)
Schaf (sheep)	Ziege (goat)	Kamel (camel)	Wolle (wool)
Tisch (table)	Stuhl (chair)	Schrank (wardrobe)	Holz (wood)
Tomate (tomato)	Gurke (cucumber)	Zwiebel (onion)	Nudeln (pasta)
Topf (pot)	Pfanne (pan)	Schüssel (bowl)	Deckel (lid)
Trompete (trumpet)	Pauke (kettledrum)	Flöte (flute)	Schall (sound)



# Semantic gradients in picture-word interference tasks: is the size of interference effects affected by the degree of semantic overlap?

James Hutson and Markus F. Damian\*

School of Experimental Psychology, University of Bristol, Bristol, UK

**Edited by:**

Ian FitzPatrick, Heinrich Heine  
Universität Düsseldorf, Germany

**Reviewed by:**

Ariel M. Cohen-Goldberg,  
Tufts University, USA  
Claudio Mulatti, Università degli  
Studi di Padova, Italy

**\*Correspondence:**

Markus F. Damian, School of  
Experimental Psychology, University  
of Bristol, 12a Priory Road,  
Bristol BS8 1TU, UK  
e-mail: m.damian@bristol.ac.uk

We report two experiments attempting to identify the role of semantic relatedness in picture-word interference studies. Previously published data sets have rendered results which directly contradict each other, with one study suggesting that the stronger the relation between picture and distractor, the more semantic interference is obtained, and another study suggesting the opposite pattern. We replicated the two key experiments with only minor procedural modifications, and found semantic interference effects in both. Critically, these were largely independent of the strength of semantic overlap. Additionally, we attempted to predict individual interference effects per target picture, via various measures of semantic overlap, which also failed to account for the effects. From our results it appears that semantic interference effects in picture-word tasks are similarly present for weakly and strongly overlapping combinations. Implications are discussed in the light of the recent debate on the role of competition in lexical selection.

**Keywords:** spoken production, picture-word interference, lexical access, object naming, competition

## INTRODUCTION

Models of language production which incorporate competitive lexical selection (Roelofs, 1992; Caramazza, 1997; Levelt et al., 1999; La Heij et al., 2006) have recently been challenged by claims that selection occurs without competition (e.g., Mahon et al., 2007). A major source of evidence supporting competitive processing in language production arises from an empirical effect found in picture-word interference (henceforth PWI) experiments. The *semantic interference effect* is characterized by the slower naming of a picture (e.g., *bear*, “target”) when a superimposed written (or simultaneously spoken) word (“distractor”) is related in meaning (e.g., *whale*) in comparison to when it is unrelated (e.g., *house*). The increased difficulty in selecting the picture name in the presence of a semantically related item was long thought to imply competitive lexical selection (e.g., Roelofs, 1992, 1993, 2001, 2003; Humphreys et al., 1995; Starreveld and La Heij, 1995, 1996; Costa et al., 1999; Damian and Martin, 1999; Levelt et al., 1999; Vitkovitch and Tyrrell, 1999; Caramazza and Costa, 2000; Bloem and La Heij, 2003; Damian and Bowers, 2003; Vigliocco et al., 2004; Belke et al., 2005; Hantsch et al., 2005, 2009). However, this view has recently been challenged (e.g., Finkbeiner and Caramazza, 2006; Mahon et al., 2007; Janssen et al., 2008). The work reported in this article investigates the issue via the question whether the degree of semantic overlap between target and distractors is relevant in PWI tasks.

Lexical selection serves the purpose of isolating the single most appropriate item from a cohort of related items. Spoken word production is initiated at the level of conceptual preparation. A cohort of items, often referred to as nodes, is said to become active because activation spreads between related concepts and

their components (e.g., Levelt, 1999). Thus, when presented with a picture of a *bear*, for example, other related items such as *wolf*, *deer*, and *rabbit* (among others), will form the related cohort, with the activation level of each item being determined by the strength of semantic relationship with the target. Activated cohort items at the conceptual level subsequently spread activation to corresponding nodes at the lexical level (Collins and Quillian, 1969; Collins and Loftus, 1975; Dell, 1986; Roelofs, 1992; Caramazza, 1997; Levelt et al., 1999, but see Bloem and La Heij, 2003). Selection of a single item from the activated cohort must then take place at the lexical level prior to further processing.

Semantic interference effects in PWI tasks have long been attributed to the co-activation of the distractor word’s representation at the lexical level, delaying the retrieval of the target’s name (e.g., La Heij, 1988; Glaser and Glaser, 1989; Schriefers et al., 1990; Vitkovitch and Humphreys, 1991; Roelofs, 1992; Levelt et al., 1999). Indeed, semantic interference of this type initially motivated the inclusion of competitive selection principles in models of spoken production. For instance, in one of the most prominent models of language production, WEAVER++ (Roelofs, 1992; Levelt et al., 1999), the speed with which selection of a target item takes place is determined by its activation level in relation to the summed activation level of all other active lemmas. Hence, targets and non-targets compete: the greater the target’s activation level compared to other items, the faster it will be selected, and co-activated non-target items slow response selection. Semantic interference in PWI is accounted for via an exchange of activation between target and distractor at the semantic level (see Roelofs, 1992, for detailed computational simulations).

Recently, however, this interpretation has been challenged, based on results from a variety of methods (see Mulatti and Coltheart, 2012; Spalek et al., 2012, for recent reviews). Instead, an alternative account of PWI effects has been introduced, the “response exclusion hypothesis” (REH, see Mahon et al., 2007, for the most detailed outline of this position). According to this view, it is not lexical competition between target and distractor which causes semantic interference in PWI, but instead, REH advocates a post-lexical, articulatory locus of inhibition. Distractors come to occupy a prearticulatory buffer from which they must be removed prior to the commencement of target naming. In REH, task demands set the rules which govern the ease with which distractors can be removed from the buffer. The semantic component governing task demands is relatively crude and operates on the fulfillment of broad semantic constraints such as category membership, but it is insensitive to more fine-grained semantic properties. Thus, distractor words that share category membership with target pictures are more difficult to remove from the buffer than those which do not. This increased difficulty, according to REH, results in the categorical interference witnessed in numerous picture-word experiments.

The general idea that the distractor needs to be removed from a response buffer before target naming can proceed was initially motivated by the observation that when the frequency of distractor words is manipulated, low frequency distractors generate more interference than high frequency ones (Miozzo and Caramazza, 2003). This finding seems difficult to reconcile with the notion of lexical selection by competition, as the latter view would either predict the opposite pattern, or a null finding of distractor frequency. A further central observation in recent research is that semantic interference appears restricted to categorically related pictures and distractors; other forms of overlap (e.g., part-whole, associations, etc.) tend to generate facilitation (Bloem and La Heij, 2003; Costa et al., 2005). On a competitive account, it is not immediately clear why type of overlap should be of such major relevance. If interference arises as a result of overlap at the semantic level, then all sorts of semantic relations should result in similar effects. As outlined above, the REH advocates a “response buffer” locus of the effect which is sensitive to only very broad semantic criteria. Hence, categorically related distractors have sufficient “task relevance” to delay removal of the distractor word from the response buffer, whereas non-categorically related distractors (e.g., target: *mouse*; distractor: *cheese*) are not interpreted as task-relevant and so create no interference compared to unrelated items. In order to account for the facilitation (rather than interference) typically caused by non-categorical relationships, REH advocates that this reflects priming at the conceptual level. Hence, the net behavioral effect in PWI tasks is taken to result from a combination of semantic priming on the one hand, and response-buffer-based interference on the other hand which is restricted to categorically related pairs (see Blackford et al., 2012, for recent EEG-based evidence that semantic effects in spoken word productions could arise from multiple sources). By abandoning a competitive principle active in lexical selection, REH constitutes a major break with conventional thinking about this issue. However, it must be noted that there are alternative scenarios that maintain the notion of competitive lexical selection

while still accounting for the observation that semantic interference in PWI is restricted to category members. Abdel Rahman and Melinger (2009) have suggested that the flow of activation in the conceptual stratum is very different when items are related categorically and noncategorically. The former is characterized by the activation of a large cohort of items with similarly high levels of activation; the latter, by a comparatively small number of items with a greater range of activation. Thus, interference is restricted to items sharing category membership while other forms of meaning overlap result in conceptual priming.

In the work reported below, we focused on the effects of semantically related distractors, and we asked whether the *degree of semantic overlap between target and distractor* has an influence on the size of the resulting interference effect. As will be shown, competitive and non-competitive theories of lexical selection in word production make opposing predictions in this regard, and indeed, two previous relevant data sets directly contradict each other. We will begin with a brief review of the relevant findings.

Very few studies have directly manipulated the semantic distance between distractor and target using the picture-word paradigm. The first investigation of semantic distance in PWI tasks that we are aware of provided evidence for a gradient such that greater target-distractor semantic relatedness resulted in stronger response inhibition. Vigliocco et al. (2004) introduced the “featural and unitary semantic space” (FUSS) hypothesis, a theory which emphasizes the role of featural representations as essential components of conceptual structure. According to this theory, featural representations are bound into lexico-semantic representations, the organization of which is determined by shared and correlated features between concepts. As such, these computational principles can be used to index the relatedness between individual concepts. The authors gathered featural data for a large number of concepts by asking participants to generate a sufficient number of features for each word, and trained a self-organizing computational model on these features, resulting in a semantic map. Vigliocco et al. then used the semantic distance between words/concepts to predict behavioral effects in various experimental paradigms. In their third experiment, they used relatedness scores derived from FUSS to select targets and distractors for a PWI experiment. Of the four conditions, in one (the “far” condition) pictures and distractors were essentially semantically unrelated, while in the other three they varied in the degree of semantic overlap from “medium” to “very close.” Results showed a graded semantic interference effect, such that interference decreased with semantic distance. Furthermore, category membership as such appeared less relevant, as the results did not significantly change when analysis was restricted to only category coordinates.

The finding of a semantic gradient in PWI tasks, with larger interference caused by strongly than by weakly related distractors, is generally compatible with models of competitive lexical selection: highly related target-distractor pairs should engage in a large degree of activation exchange via the conceptual level, resulting in strong competition between distractor and target at the lexical level; weakly related pairs should result in relatively less competition. More recently, however, results were reported which suggest the opposite pattern. Mahon et al. (2007) carried out a series of

experiments investigating the effect of manipulating the semantic distance between target pictures and distractor words on the speed of naming times. The first relevant study (Experiment 4) attempted to control for semantic distance while manipulating category membership. Semantic distance values were derived from the semantic similarity norms of Cree and McRae (2003). Relatedness values were established through the number of shared features between items. Participants generated lists of features, in a fashion very similar to the method used by Vigliocco et al. (2004). For example, in response to the word *knife*, participants might have generated the features: “is sharp,” “has a handle,” “used for cutting,” and “found in kitchens.” Items from within a particular category will normally share more features than they do with those from other categories. However, within-category items sometimes share very few features, making it possible to match within- and between-category items on semantic relatedness. Thus, stimuli were constructed in which within-category targets and distractors had the same semantic overlap as targets and distractors from separate categories. It was found that categorically related items caused more interference than non-categorically related items, suggesting that category membership exerted an effect over and above that of semantic relatedness, at odds with the findings of Vigliocco et al. outlined above.

In a subsequent experiment which is particularly critical for the current work, Mahon et al. (2007 Experiments 5 and 5b) directly manipulated within-category semantic distance. Rather than semantic relatedness values derived from the extent of feature overlap, ratings of semantic relatedness from human participants were gathered for each pair of items used in the experiment. Each target picture was paired with a closely related distractor word and a distantly related distractor, or with two unrelated distractor words. Surprisingly, results indicated that distantly related target-distractor pairs (e.g., *horse-whale*) interfered *more* than closely related pairs (*horse-zebra*). This finding was replicated in a further study (Experiment 5b) with the same materials but a separate group of participants. Further support for the direction of the effect was provided in Experiment 6 of the series, where naming times of targets in the context of close and far distractors were compared directly rather than with targets with unrelated distractors superimposed. The effect was reliable by participants ( $p \leq 0.05$ ) but only marginally by items ( $p = 0.11$ ). In the final two experiments carried out by Mahon et al. (Experiments 7 and 7b) stimuli were selected such that close and far conditions had a large difference in relatedness according to the norms of Cree and McRae (2003). The nature of the semantic relationships generated from the norms was confirmed through ratings obtained from a group of native English speakers. The interval between onset of picture and distractor (stimulus-onset asynchrony, or SOA) was varied; by varying the “entry time” of the distractor relative to that of the target, the distractor taps into successive stages of target presentation, hence yielding information about temporal patterns. Three separate SOAs were examined:  $-160$ ,  $0$ , and  $+160$  ms. A reliable effect in which distantly related distractors interfered more than closely related distractors was found in both of the staggered presentation conditions but not in the simultaneous presentation condition. A replication of the experiment with only the simultaneous condition again found no

effect of semantic distance on naming times for synchronous presentation.

Although the final two experiments in the series carried out by Mahon et al. (2007) provide somewhat ambivalent results, overall the findings suggest increased interference from distractors that are more distantly related targets, compared to those more closely related. The findings of Experiments 5 and 5b are especially compelling: counterintuitively, strong interference was restricted to the semantically *far* condition, whereas closely related distractors either generated a slight facilitatory effect (in Experiment 5), or an interference effect of less than half that of distantly related distractors (Experiment 5b). The explanation advocated by Mahon et al. is as follows: at the conceptual level, strongly related distractors cause larger facilitation than weakly related ones. At the “response buffer” level, only broad category membership is relevant, so strongly and weakly related distractors generate equivalent interference. The net outcome is that for strongly related distractors, conceptual priming, and response buffer interference largely cancel each other out; for weakly related distractors, relatively less conceptual priming results in a larger behavioral interference effect. PWI tasks should hence exhibit a “reversed semantic gradient.”

Overall, the role and impact of semantic overlap in PWI tasks remains somewhat inconclusive. The two studies by Vigliocco et al. (2004) and Mahon et al. (2007) rendered contradictory results which have never been satisfactorily accounted for. Other than these two key contributions, we are not aware of other studies which would have directly tested the effect of manipulating the semantic distance between categorically related pictures and words. Understanding the nature of within-category semantic gradients in the picture-word paradigm is an important issue which has critical implications for models of language production. A more complete understanding of the processes which contribute to naming response times is vital to progress debate as to whether lexical selection in word production is competitive or not.

In the work below, we contributed to the debate surrounding the nature of within-category semantic gradients in the PWI task by replicating the two key experiments by Vigliocco et al. (2004, Experiment 3) and Mahon et al. (2007, Experiments 5 and 5b). Both studies manipulated within-category distance in a similar manner, yet reported results which directly contradict each other. Hence, we deemed it important to re-run both experiments with the same apparatus and participant pool. Vigliocco et al. and Mahon et al. used the PWI task in a slightly differing format (outlined below); our aim was to run both studies with the same trial format hence our experiments are not exact replications of the original studies. The critical experimental aspects are as follows:

(1) Vigliocco et al. (2004) and Mahon et al. (2007) used picture-distractor SOAs of  $-150$  and  $0$  ms, respectively. With visually presented distractors, an SOA of  $0$  ms is a popular and common choice when targeting semantic effects; in an analysis of the effects of varying SOAs in PWI tasks, Damian and Martin (1999) found the most pronounced semantic effect at this SOA, but reduced interference with  $-100$  ms, and no interference with  $-200$  ms. For this reason, we used SOA =  $0$  ms in both experiments reported below. (2) Both Vigliocco et al. and

Mahon et al. presented target pictures centrally but varied distractor position slightly from trial to trial, although in different ways: Vigliocco et al. presented distractors in a randomly selected location either above or below the fixation cross (the degree of dislocation is not specified) whereas Mahon et al. varied distractor position both horizontally and vertically around the fixation cross by 2 cm. In our reading, the vast majority of published PWI studies with written distractors have used a central presentation of both targets and distractors, therefore we also used this format. We are not aware of findings in the literature suggesting that this procedural variation could be relevant, and indeed, the presence of semantic interference effects in both the previous studies and our own experiments clearly demonstrates that participants accessed the meaning of the distractors. (3) There was some minor variation in trial structure between the earlier studies: Vigliocco et al. presented a continuous series of trials to participants, with a trial sequence of: fixation cross for 500 ms, blank screen for 50 ms, distractor presented, target appearing 150 ms later, and both visible until response. In Mahon et al.'s studies, each trial was initiated by a participant via a key press; on each trial, a fixation cross was presented for 500 ms, followed by the target/distractor combination which was presented until the voice key detected a response. In our two experiments presented below, a continuous series of trials was delivered, and the trial sequence was as follows: a 1000 ms blank screen, followed by a centrally located fixation point presented for 500 ms, immediately followed in the same location by the target picture and distractor word for 2000 ms. There is no reason from the existing literature to suspect that such minor variations could affect results in PWI studies. (4) Neither Mahon et al. nor Vigliocco et al. stated the source of their target pictures for the critical experiments; however, Vigliocco et al. declared for an earlier experiment in their article that "pictures were obtained from Snodgrass and Vanderwart (1980) and supplemented by additional pictures created for the purpose" (p. 445). Indeed, 15/20 targets in Mahon et al., and 21/24 targets in Vigliocco et al. were in the Snodgrass and Vanderwart set. We therefore used these pictures, augmented with a few additional images selected from other object sets.

One aim of our experiments was to investigate whether we could replicate the central (and mutually contradictory) findings from the original studies. A second aim was to investigate the role of semantic overlap in PWI tasks at the item level, by computing interference scores for each target picture which identify the degree of semantic interference associated with a particular target-picture combination. We therefore explored the association between these item-specific interference scores and various measures of semantic relatedness, namely (1) semantic relatedness ratings which we collected from a separate group of participants, (2) semantic distance scores obtained via Latent Semantic Analysis (LSA), (3) Normalized Google Distance (NGD). By regressing item-specific interference effects onto these semantic relatedness measures, we expected to gain more detailed insight into the directionality (if any) of semantic effects in PWI tasks.

## EXPERIMENT 1

The first experiment aimed to provide a replication of Vigliocco et al.'s (2004) Experiment 3, with the procedural variation

outlined above. Pictures were named in the presence of visually presented distractor words, with semantic distance between picture and word manipulated in four conditions: *far*, *medium*, *close*, and *very close*. The *far* condition corresponds to the unrelated condition used in numerous PWI studies. As Mahon et al. (2007; targeted in Experiment 2) used the label *far* to refer to a "weakly related" condition, from here on we will use the common label *unrelated* to avoid confusion. As discussed above, the relatedness between targets and distractors was established through FUSS (Vigliocco et al., 2004), with relatedness values for items in the *unrelated* group >18.5 units on the lexico-semantic map, in the *medium* group ranging from 7.5 to 10.5 units, the *close* group from 4.5 to 7.5 units, and the *very close* group from 1.5 to 4.5 units.

As pointed out by Mahon et al. (2007), some aspects of stimulus selection in this experiment are suboptimal (but justified on the basis of how Vigliocco et al., 2004, generated their materials). Although the same target pictures were used in each condition, the allocation of distractor words between the relatedness conditions was not as controlled, with many, but not all, words appearing in multiple distractor conditions. Further, although the bulk of distractors and pairs were categorically related, some of them were also associatively related (e.g., trousers-belt), some had form overlap (e.g., broom-banana), and indeed, a few pairs were not members of an obvious semantic category (e.g., axe-pencil). Nevertheless, the aim of the current research was primarily to establish the reliability of the original results, and consequently, no alterations were made to the original stimuli. To recap, the results of the original experiment identified a significant linear trend in which semantic interference increased with semantic overlap.

## METHODS

### Participants

Twenty-six undergraduate students at the University of Bristol were recruited as participants in the study and received course credit. For this and the following experiment, ethical approval was granted by the Faculty of Science Human Research Ethics Committee at the University of Bristol. All experiments conformed to the relevant regulatory standards. Informed consent was obtained from all participants prior to testing.

### Materials

Materials were taken from Experiment 3 of Vigliocco et al. (2004), consisting of 24 target pictures paired with 67 distractor words to form *unrelated*, *medium*, *close*, and *very close* pairings. Note that as is typical of PWI studies, the same target pictures were used in all conditions, which excludes the possibility that between-item differences with respect to, e.g., the ability of object names to trigger the voice key, might obscure the results. The semantic distance between target and distractor was established by Vigliocco et al. through FUSS (described above) via semantic feature analysis. Target pictures were largely selected from a set previously shown to have high name agreement (Snodgrass and Vanderwart, 1980). Distractors were matched across the four conditions for frequency and length, and care was taken to minimize phonological overlap with targets, although, as noted above, this was

not achieved across the entire stimuli set. Similarly to the original experiment, 24 filler pictures were selected from semantic categories other than those of the target set. Four distractor words were selected for each filler picture, one of which was selected on the basis of being semantically related to the filler picture while the other three were unrelated to the filler. The total number of experimental trials, including critical and filler items, was 192. Supplementary Material shows the critical combinations.

### Design

Four experimental blocks were created from the 48 critical and filler pictures and the four distractors associated with each picture. In each block every picture was presented once. Across the four blocks, each picture was presented once with each of its four associated distractors and followed each other an equal number of times. Distractors were organized so that each experimental block contained a balanced number of each type, i.e., an equivalent proportion of *unrelated*, *very close*, *medium*, and *close* target-distractor pairs. The order of the blocks was manipulated so that each block was presented the same number of times in each position. Similarly to the original experiment, items were presented in a pseudorandom order with the only constraint stipulating that critical items and filler items alternated.

### Procedure and apparatus

Participants were tested individually. Stimuli were presented using DMDX (Forster and Forster, 2003) running on a PC, with vocal responses captured by a head-mounted microphone. Prior to the commencement of testing participants were presented with two grid-like screens which contained all of the 48 targets pictures and their correct names. Participants were asked to familiarize themselves with the pictures until they felt they could name all of them with the given names (note that a pre-experimental familiarization phase, also used in Experiment 2, is common in PWI experiments, and was also used in the studies of Vigliocco et al., 2004; Mahon et al., 2007). A practice block followed in which each picture was presented once with an unrelated distractor word that was not used in the experiment proper. Participants were instructed to name the target pictures as quickly and accurately as possible, while ignoring written distractor words. In this phase, picture names other than those expected were corrected by the experimenter. Subsequently to the practice, the four blocks of experimental trials were presented. At the end of each block the experiment paused until the participant indicated they were ready to continue.

Targets and distractors were both presented centrally on the screen, and with the same onset (SOA = 0 ms). The same presentation and SOA were also used in our second experiment (see below), rendering them directly comparable. The sequence was as follows: a 1000 ms blank screen, followed by a centrally located fixation point presented for 500 ms, immediately followed in the same location by the target picture and distractor word for 2000 ms. Distractor words were presented in bold 18 pt Courier New typeface. All pictures were clearly visible despite the presence of the superimposed distractor word. DMDX recorded individual naming latencies to the harddrive and determined response latencies via a digital voice key relative to the onset of the target picture.

The entire session including the familiarization and practice process lasted approximately 30 min per participant.

## RESULTS

### Initial analysis

All responses were audiovisually checked for accuracy of the response trigger determined by DMDX, as well as for inaccurate responses using CheckVocal (Protopapas, 2007). Responses were classified as errors if, on a given trial, a name other than that of the target was produced, a correction was made, the response was disfluent, or no response was made within the response window. Latencies faster than 250 ms or longer than 1800 ms (3.6%) were excluded as outliers.

Table 1 shows the results, as well as the original Vigliocco et al. (2004) findings for comparison. For latencies, the results show a semantic interference effect of approximately 40 ms. Surprisingly this effect seems unaffected by the degree of semantic overlap between target and distractor, with very similar interference obtained for the *medium*, *close*, and *very close* conditions. Analyses of variance (ANOVAs) were conducted on the latencies, with either participants ( $F_1$ ) or items ( $F_2$ ) as the random variable, and Condition (*unrelated*, *medium*, *close*, *very close*) as a fixed variable. The results showed a highly significant effect of Condition,  $F_{1(3, 75)} = 11.65$ ,  $MSE = 11, 969$ ,  $p < 0.001$ ;  $F_{2(3, 69)} = 6.34$ ,  $MSE = 10,093$ ,  $p < 0.001$ . A trend analysis performed on the levels of Condition showed a combination of linear [ $F_{1(1, 25)} = 18.71$ ,  $p < 0.001$ ;  $F_{2(1, 25)} = 8.67$ ,  $p = 0.007$ ], quadratic [ $F_{1(1, 25)} = 8.94$ ,  $p = 0.006$ ;  $F_{2(1, 25)} = 12.07$ ,  $p = 0.002$ ] and cubic [ $F_{1(1, 25)} = 4.76$ ,  $p = 0.039$ ;  $F_{2(1, 25)} = 4.02$ ,  $p = 0.056$ ] components (by comparison, Vigliocco et al.'s results were characterized by an exclusively linear trend).

Planned tests which compared the four conditions against each other showed that all three related conditions (*medium*, *close*, *very close*) differed significantly from the *unrelated* condition,  $t_1 \geq 4.63$ ,  $p < 0.001$ ;  $t_2 \geq 2.98$ ,  $p \leq 0.007$ , whereas the three related conditions did not differ significantly from each other,  $t_1 \geq 0.55$ ,  $p \leq 0.585$ ;  $t_2 \geq 0.33$ ,  $p \leq 0.743$ . A further analysis was carried out subsequent to the removal of various potentially problematic stimuli (16 in total) that were either form related, associatively

**Table 1 | Mean response latencies (RT, in ms), error rates (PE, in %), and effects (related minus unrelated) for Experiment 1, and results from Vigliocco et al. (2004).**

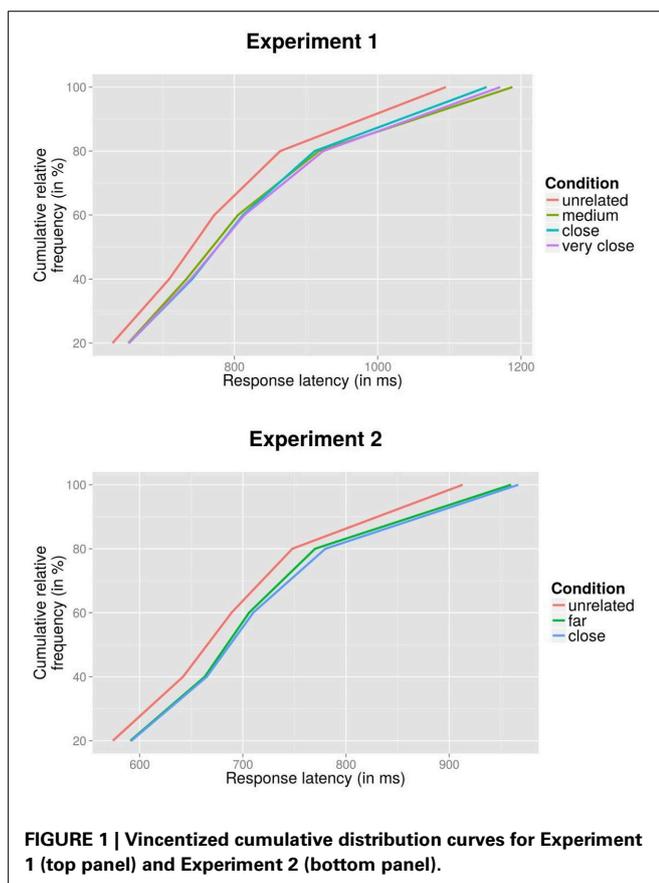
Target-distractor relationship	RT	Effect	PE	Effect
<b>EXPERIMENT 1</b>				
Unrelated	803 (101)		2.2 (3.7)	
Medium	846 (127)	43	2.7 (3.2)	0.5
Close	843 (130)	40	2.8 (5.4)	0.6
Very close	848 (110)	45	3.1 (5.5)	0.9
<b>Vigliocco et al., 2004, EXPERIMENT 3</b>				
Unrelated	642		5.9	
Medium	648	6	6.2	0.3
Close	657	15	6.1	0.2
Very close	671	29	7.5	1.6

Standard deviations are in parentheses.

related, or not categorically related. Removal of these items failed to considerably affect results.

To further explore the effect of relatedness in the naming latencies, Vincentized cumulative distribution curves were computed (Ratcliff, 1979): for each participant and condition, rank-ordered latencies were divided into 20% quantiles, and mean latencies were computed for each quantile. These were then averaged across participants, which preserves the shapes of individuals' latency distributions (cf. Woodworth and Schlosberg, 1954). An analysis of this type provides information about the degree of uniformity with which an effect affects the spectrum of response latencies. This is particularly important in case of a null finding (here, no difference between the three related conditions): perhaps, very strongly related items show conceptual priming (which might manifest itself particularly in fast latencies) but also increased interference (which might particularly affect the right tail of latencies). In this case, the net result might not be visible in conditional means compared to less related conditions, but a pattern would emerge in response time distributions. Indeed, in the Stroop literature, null effects on mean latencies which result from different opposing underlying effects have been highlighted (e.g., Heathcote et al., 1991). **Figure 1** (top panel) shows the results for the four conditions of Experiment 1: for all three related conditions, relatedness exerts a similar effect across the entire spectrum of response times compared to the baseline condition.

Parallel ANOVAs conducted on error proportions showed no effect of Condition,  $F_1 < 1$ ,  $F_2 = 1.06$ .



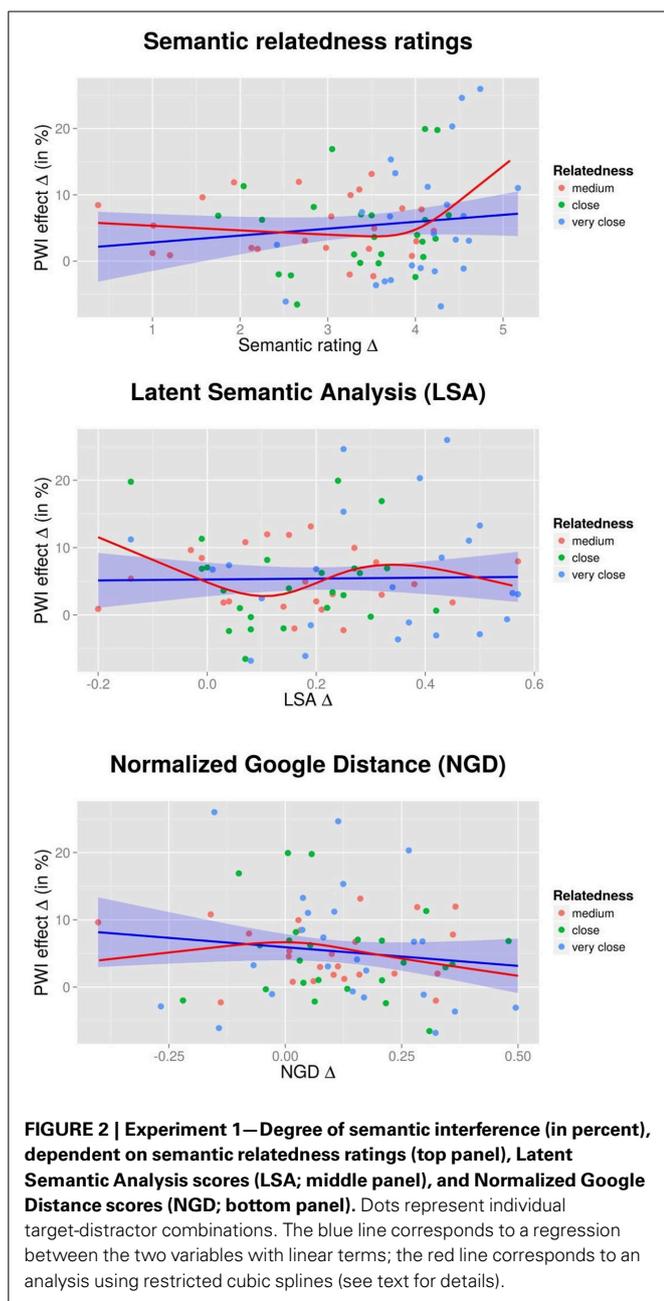
In sum, the latency analysis showed a pattern in which the unrelated condition differed significantly from all three related conditions, but the extent of interference was not affected by the semantic distance between target and distractor. Next, we attempted to further elucidate the results by attempting to account for variability among individual targets-distractor pairs regarding their degree of semantic interference via a number of measures of semantic overlap. For each target and condition, we calculated the associated interference effect in the PWI study, and we assessed the fit between the interference effects and a range of measures of semantic overlap.

### Semantic relatedness ratings

A straightforward way of identifying the degree of semantic overlap between a pair of items is to collect semantic relatedness ratings (e.g., Mahon et al., 2007). We conducted such ratings for all picture-distractor combinations in Experiment 1; items for Experiment 2 (described below) were also included. Twenty-seven individuals, none of whom were participants in the two experiments, were presented with pairs of words corresponding to pictures and distractors in the experiments, and were instructed to rate “how related the two concepts denoted by the words are” (these instructions were taken from Mahon et al.’s ratings). Ratings were carried out on a 1–7 scale, with seven indicating an “very related” pair and one a “not related” pair (a number of examples such as *spider-fly* and *house-bat* were provide as reference points for strongly related and unrelated pairs). A different random order of word pairs was presented to each participant. For the materials in the present study, mean relatedness ratings were 1.7, 4.5, 5.1, and 5.7 for the *unrelated*, *medium*, *close*, and *very close* conditions. From these, we calculated distance scores for each individual target picture by subtracting the unrelated baseline from each of the related ratings. For instance, the target picture “rake” has ratings of 1.9, 2.9, 3.6, and 5.6 when paired with the distractors “carpet,” “sword,” “hatchet,” and “shovel,” so relatedness scores are 1.0, 1.8, and 3.7 for the *medium*, *close*, and *very close* conditions. Hence, higher rating difference scores are associated with a stronger degree of relatedness (or more precisely, a larger difference between the related and the unrelated rating for that item).

Next, for each of the 24 target pictures, we calculated an “interference” score by subtracting the naming latency mean in the unrelated condition from each of the means in the related conditions. For instance, the target picture “rake” generates 51 ms interference when paired with the *medium* distractor “sword” (compared to when paired with the *unrelated* distractor “carpet”), 65 ms when paired with the *close* distractor “hatchet,” and 64 ms when paired with the *very close* distractor “shovel.” To account for variability between targets regarding their overall latencies, values were then converted into percentages relative to the *unrelated* baseline condition; e.g., for the target “rake,” *medium*, *close*, and *very close* conditions resulted in 5.4, 6.8, and 6.7% interference.

**Figure 2** (upper panel) shows a scatter plots, with dots representing individual picture-word combinations (color-coded for condition). PWI interference is on the y-axis, and the ratings effect on the x-axis. If interference increases with growing semantic overlap (as predicted by Vigliocco et al., 2004) this should



result in a positive slope; if interference is stronger for weakly related items (as stipulated by Mahon et al., 2007) a negative slope should emerge. A regression, with the trendline (plus confidence intervals) shown in blue was fitted to the data, but did not result in a significant outcome,  $F_{(1, 70)} = 1.55$ ,  $p = 0.217$ ,  $R^2 = 0.022$ . However, inspection of a density plot of the residuals from the linear model suggested some degree of asymmetry, hence, we attempted to model this potential non-linearity via “restricted cubic splines” (RCS; Harrell, 2001; Baayen, 2008). This technique combines a series of cubic polynomials defined over a series of corresponding intervals. We chose four knots (i.e., three intervals) based on Harrell’s suggestion for samples of our sizes (p. 135). **Figure 2** (top panel) shows the outcome of the RCS analysis with

a red line, suggesting a rise in the right tail end. A RCS model with 4 knots approached significance,  $F_{(3, 68)} = 2.49$ ,  $p = 0.067$ ,  $R^2 = 0.10$ .

The RCS model suggests a possible tendency for a few very strongly related items (see upper right corner of the panel) to provide more interference (>20%) than the other, less related, items. The three combinations which generated interference larger than 20% are coat-suit, cucumber-broccoli, and finger-thumb. Note that the three target pictures *coat*, *cucumber*, and *finger* by themselves are not problematic, as they have average latencies in the unrelated condition which puts them below the overall unrelated mean (753, 788, and 693 ms).

We conclude that overall, semantic relatedness between picture and distractors—as assessed by semantic relatedness ratings—does not appear to affect the degree of semantic interference in the PWI task. The directionality of the trend emerging in the RCS analysis (increased interference for very strongly related items) is generally in line with the predictions made by Vigliocco et al. (2004), but the variance accounted for is low even with the RCS model (~10%).

#### Latent Semantic Analysis (LSA)

Vigliocco et al. (2004) selected their items via relatedness scores generated from FUSS, which are based on the number of shared semantic features. The analysis reported in the previous section showed no clear association between semantic relatedness ratings and the degree of interference in our PWI experiment. But perhaps ratings, based on individuals’ intuitions about semantic relatedness, are not optimal to investigate conceptual structure, and “objective” measures do a better job in predicting PWI results. Although FUSS scores of the individual stimuli were not available to us, an alternative objective measure of semantic distance is Latent Semantic Analysis (LSA; Landauer et al., 1998). LSA applies statistical computations as a means of generating relatedness scores from a large corpus of text. The contextual usage of words is assessed through the aggregation of all the contexts in which a particular word does and does not appear, determining the similarity of meaning through a set of mutual constraints. The degree to which LSA reflects human knowledge has been demonstrated in a number of ways, including category judgment and word sorting (Landauer et al., 1998).

We computed LSA relatedness scores for each picture-target combination, in analog to what was described for the semantic relatedness ratings. As in Vigliocco et al. (2004, p. 448), we used the LSA web-based interface (<http://lsa.colorado.edu/>), using the “General reading up to 1st year of college” topic space and “Matrix comparison.” Then, LSA difference scores were computed in the same way as for the relatedness ratings described in the previous section, and plotted against behavioral interference effects (for three combinations, LSA scores were not available). **Figure 2** (middle panel) shows the relationship between PWI interference and the difference in relatedness. A regression model representing a linear relation between LSA scores and interference did not result in a significant outcome,  $F_{(1, 69)} = 0.02$ ,  $p = 0.884$ ,  $R^2 < 0.01$ , and neither did a RCS model with four knots,  $F_{(3, 67)} = 1.74$ ,  $p = 0.167$ ,  $R^2 = 0.07$ .

### Normalized Google Distance (NGD)

We attempted to predict the amount of interference via “Normalized Google Distance” (NGD), a semantic similarity measure derived from the number of hits returned by the Google search engine for a given set of words (Cilibrasi and Vitanyi, 2006, 2007). The normalized Google distance between two search terms  $x$  and  $y$  is computed as

$$\text{NGD}(x, y) = \frac{\max\{\log f(x), \log f(y)\} - \log f(x, y)}{\log M - \min\{\log f(x), \log f(y)\}}$$

with  $M$  the total number of pages available to Google,  $f(x)$  and  $f(y)$  the number of hits for individual search terms  $x$  and  $y$ , and  $f(x, y)$  the number of hits for joint occurrence. Words which tend to co-occur in the search space take on values close to zero, whereas words which never co-occur take on infinite values: words with similar meaning tend to be close (have lower values) than words with dissimilar meaning. For instance, “coat” and “suit” tend to co-occur (NGD = 0.01) whereas “coat” and “bus” do so less often (NGD = 0.26).

We computed NGD values for all picture-distractor combinations used in Experiment 1<sup>1</sup>. All three related conditions resulted in quite similar average NGD values (0.22, 0.19, and 0.19 for the *medium*, *close*, and *very close* condition) which were slightly lower than those for the unrelated condition (0.31). As for relatedness ratings and LSA measures, we then computed difference scores for all related, relative to the unrelated, combinations. Here we subtracted the related from the unrelated condition (rather than vice versa, as in the previous two analyses) to preserve the directionality of the other two analysis, i.e., higher NGD difference values reflect stronger overlap. **Figure 2** (bottom panel) shows the relationship between PWI interference and the NGD difference scores. A regression model with a linear relation between NGD and interference did not show a significant outcome,  $F_{(1, 70)} = 1.34$ ,  $p = 0.251$ ,  $R^2 = 0.02$ , and neither did a RCS model with four knots,  $F_{(3, 68)} = 0.83$ ,  $p = 0.481$ ,  $R^2 = 0.04$ .

### DISCUSSION

To summarize the results, a clear and strong effect of semantic relatedness was found in this experiment, in line with numerous published studies in the literature. Surprisingly, however, when the unrelated group was compared to the three related conditions, there was no evidence to suggest that response latencies varied as a function of semantic distance. Quantile plots demonstrated that in each of the related conditions, latencies increased uniformly across the entire range of responses relative to the baseline condition, with very little or no difference between them. Consequently, our findings did not suggest the presence of a semantic gradient in which more closely related targets and distractors result in greater interference.

We collected semantic relatedness ratings on our items, and tried to predict the size of the interference effect for a particular

target-distractor combination, depending on their rated relatedness. A marginally significant pattern was found when ratings were modeled onto interference effects via a non-linear technique, with pronounced interference for very strongly related items. More remarkable, however, is the null finding for all but those few items. With two further, alternative measures of semantic distance, namely LSA- and NGD-derived scores of overlap, no systematic pattern was found. Overall, the results are remarkable in their absence of an effect of degree of semantic overlap, despite the presence of strong and significant effects of relatedness when compared to the unrelated baseline.

### EXPERIMENT 2

Experiment 1 provided no compelling evidence for a semantic gradient in PWI tasks. This contrasts with Vigliocco et al. (2004) original experiment where a statistically significant linear trend indicated that more closely related distractors slowed naming more than distantly related items. The second experiment constituted an attempt to replicate (again, with minor procedural variations as described in the Introduction) perhaps the most compelling evidence for a “reversed semantic gradient,” i.e., distantly related distractors interfere *more* with picture naming than closely related distractors. As described in the Introduction, Mahon et al. (2007, Experiments 5 and 5b) compared an *unrelated* baseline to a condition in which items were distantly related (*far*) as well as one in which they were closely related (*close*), and found significant interference only in the *far* condition, but no (Experiment 5) or substantially reduced (Experiment 5b) interference for the *close* condition. In each condition, the same target pictures were named, and the same distractor words used (but differently combined with the targets). Strength of relatedness was established via semantic relatedness ratings. Our Experiment 2 replicates this study, with the only modification other than those outlined in the Introduction the exclusion of a small number of items, for reasons outlined in the Section “Materials.”

### METHODS

#### Participants

Sixty-four undergraduate students at the University of Bristol were recruited as participants and received course credit. None had been in the first experiment.

#### Materials

The stimuli were taken from Mahon et al.’s (2007) Experiments 5 and 5b. The majority of target pictures were from the Snodgrass and Vanderwart (1980) set. Of the original 20 target pictures, two (*boat* and *plane*) were removed because the corresponding related distractor words *submarine* and *helicopter* were relatively long (materials were originally selected with the intention to be included in a study using masked priming, in which long distractor words would be problematic). A further target, *plate*, had a high error rate in a pilot study because it was highly confusable with its corresponding distractor *saucer*, and was therefore omitted. Due to the way in which materials were arranged (see below) this required the removal of an additional target, *glass*.

The remaining 16 target pictures were paired with 16 categorically related distractor words. This set of distractors was

<sup>1</sup>Normalized Google distance (NGD) values were computed in March 2014, and are based on <http://www.google.com>

recombined with targets to manipulate semantic distance. As in Mahon et al. (2007), target pictures were chosen in pairs from a particular semantic category (e.g., furniture: *bed* and *stool*), each paired with a closely related distractor word (*close*: bed-futon; stool-chair), and the distractors reversed to form more distantly related combinations (*far*: bed-chair; stool-futon), and finally paired with unrelated distractors (*unrelated*: bed-pot; stool-zebra) which themselves served as related distractors when combined with other targets. This arrangement allowed the use of the same set of 16 distractor words across all conditions. Note that this design necessitated two unrelated conditions. In total there were 64 target-distractor pairings. See Supplementary Material for all critical combinations.

For targets and distractors in the *close* and *far* conditions, the original arrangement from Mahon et al. (2007) was maintained. As some targets and distractors had been removed from the original stimulus set, a number of the original unrelated pairings were no longer possible, and we recombined items for the unrelated condition. Care was taken to ensure that pairs in this condition were associatively, categorically, and phonologically unrelated.

### Design

The 64 trials were split into two blocks, such that each half contained an equal number of *unrelated*, *close*, and *far* items. There were two instances of each distractor and each target picture in each half of the stimuli. The order in which participants were presented with each half of the stimuli was counterbalanced. Presentation order was randomized, with a minimum distance of three trials between the first and second presentation of each target picture and distractor within each block, as well as a maximum of three consecutive related or unrelated target-distractor pairs.

### Procedure and apparatus

The procedure was very similar to Experiment 1. Prior to the experiment, participants were familiarized with the critical target pictures with via a grid-like screen with the 16 target pictures and their names. A practice block followed in which each picture was presented once with an unrelated distractor word that was not used in the experiment proper; responses other than those expected were corrected by the experimenter. Subsequently, the two experimental blocks, each consisting of 32 trials, were presented. At the end of each block the experiment would pause until the participant was ready for the next block.

The same presentation sequence as in Experiment 1 was used (1000 ms blank screen; 500 ms fixation cross; target and picture simultaneously presented for 2000 ms). Distractor words were presented in bold 18 pt Courier New typeface.

The entire session including the familiarization and practice phase lasted approximately 15 min per participant.

## RESULTS

### Initial analysis

Data were processed in the same way as in the first experiment. Latencies faster than 250 ms or longer than 1800 ms (2.1%) were excluded as outliers. **Table 2** shows the results, together with the corresponding results from Mahon et al. (2007). Latencies

showed approximately 20 ms of semantic interference, and very similar degrees of semantic interference for the *far* and *close* conditions relative to the *unrelated* baseline (for this and all following analyses, the two unrelated baselines were averaged).

ANOVAs applied to latencies, with Condition (*unrelated*, *far*, *close*) as a fixed variable, showed a highly significant effect of Condition by participants,  $F_{1(2,126)} = 9.07$ ,  $MSE = 12028$ ,  $p < 0.001$ , which was marginally significant in the analysis by items,  $F_{2(2,30)} = 2.85$ ,  $MSE = 2456$ ,  $p = 0.074$ . We did not perform a trend analysis as in the first experiment, due to the low number of conditional means. Planned tests which compared the two conditions against each other showed that the two related conditions (*far*, *close*) differed significantly from the unrelated condition in the analysis by participants,  $t_1 \geq 3.35$ ,  $p \leq 0.001$ , and marginally in the analysis by items,  $t_2 \geq 1.78$ ,  $p \leq 0.095$ . By contrast, the two related conditions did not differ significantly from each other,  $t_1 = 0.67$ ,  $p = 0.503$ ;  $t_2 = 0.48$ ,  $p = 0.636$ . **Figure 1** (bottom panel) shows cumulative response time distributions which suggest a similar effect for the two related conditions compared to the unrelated condition across the entire spectrum of response times.

Parallel ANOVAs conducted on error proportions showed an effect of Condition which was significant by participants,  $F_{1(2,126)} = 3.12$ ,  $MSE = 34.3$ ,  $p = 0.048$ , but not by items,  $F_{2(2,30)} = 2.32$ ,  $MSE = 8.6$ ,  $p = 0.116$ . Planned tests showed that the *close* condition differed significantly from the *unrelated* condition in the analysis by participants,  $t_{1(63)} = 2.16$ ,  $p = 0.034$ , and just failed to reach significance by items,  $t_{2(15)} = 2.12$ ,  $p = 0.051$ . The *far* condition also differed significantly from the *unrelated* condition in the analysis by participants,  $t_{1(63)} = 2.05$ ,  $p = 0.045$ , but not by items,  $t_{2(15)} = 1.46$ ,  $p = 0.164$ . The two related conditions did not differ significantly from each other,  $t_{1(63)} = 1.04$ ,  $p = 0.300$ ;  $t_{2(15)} = 0.86$ ,  $p = 0.403$ .

### Semantic relatedness ratings

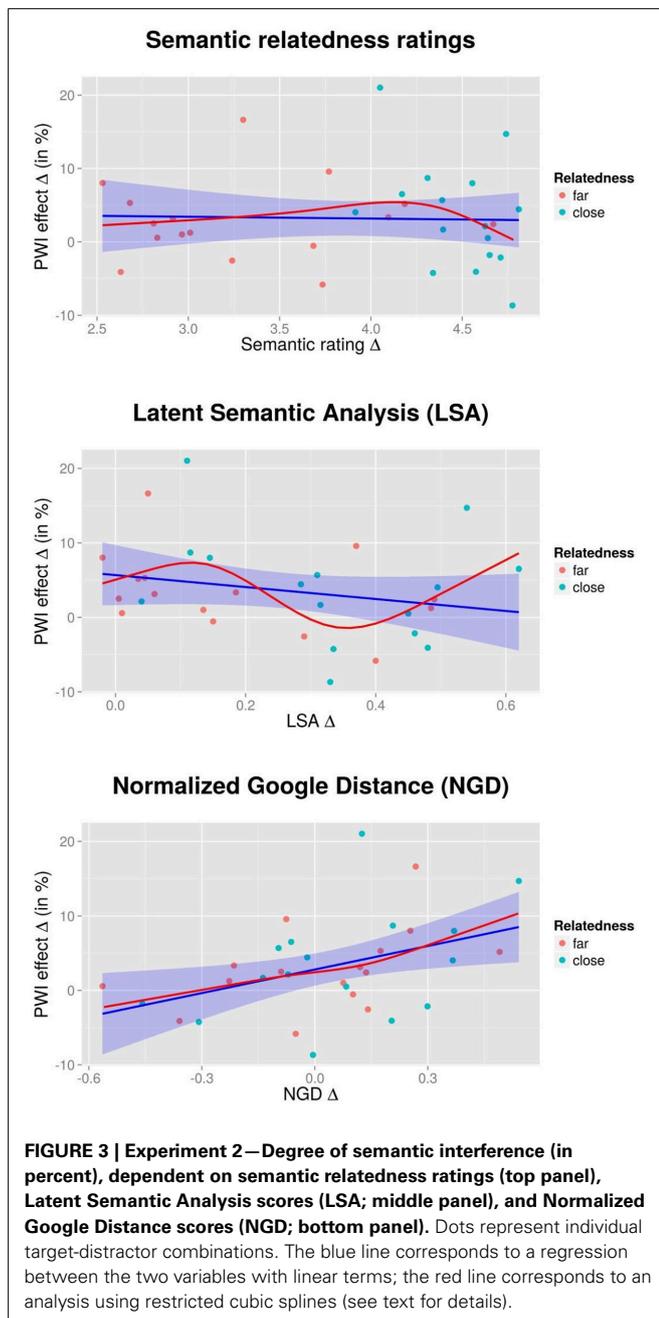
The materials used in this experiment had been included in the semantic relatedness ratings outlined in Section Semantic relatedness ratings. The results showed average ratings of 1.6 for the *unrelated* condition, of 4.9 for the *far* condition, and of 6.1 for

**Table 2 | Mean response latencies (RT, in ms), error rates (PE, in %), and effects (related minus unrelated) for Experiment 2, and results from Mahon et al. (2007).**

Target-distractor relationship	RT	Effect	PE	Effect
<b>EXPERIMENT 2</b>				
Unrelated	708 (94)		1.5 (3.0)	
Far	727 (112)	19	2.2 (3.6)	0.7
Close	731 (109)	23	2.4 (4.2)	1.4
<b>Mahon et al., 2007, EXPERIMENT 5</b>				
Unrelated	728		1.2	
Far	765	37	1.8	0.6
Close	724	-4	1.9	0.7
<b>Mahon et al., 2007, EXPERIMENT 5b</b>				
Unrelated	709		1.6	
Far	746	37	2.0	0.4
Close	726	17	2.4	0.8

the *close* condition. This compares well with the relatedness rating results reported in Mahon et al. (2007), which had shown means of 1.3, 3.9, and 5.3.

As in the first experiments, the association between ratings and corresponding PWI effects were investigated. Rating and PWI effects for each picture-word combination were computed in the same way as described in Section Semantic relatedness ratings. **Figure 3** (top panel) shows the results. A regression model with a linear relation between ratings and PWI effects resulted in no significant outcome,  $F_{(1, 30)} = 0.03$ ,  $p = 0.873$ ,  $R^2 < 0.01$ , and neither did a RCS model with four knots,  $F_{(3, 28)} = 0.59$ ,  $p = 0.627$ ,  $R^2 = 0.06$ .



### Latent Semantic Analysis (LSA)

LSA scores were computed for each target-distractor combination, and difference scores were calculated in the same way as outlined in Section Latent Semantic Analysis (LSA). **Figure 3** (middle panel) shows the association between LSA difference scores and behavioral PWI effects. A regression model with a linear relation resulted in no significant outcome,  $F_{(1, 28)} = 1.69$ ,  $p = 0.204$ ,  $R^2 = 0.06$ . A RCS model with four knots resulted in a marginally significant outcome,  $F_{(3, 26)} = 2.92$ ,  $p = 0.053$ ,  $R^2 = 0.25$ .

### Normalized Google Distance (NGD)

We computed NGD scores for the current stimuli, as described under Section Normalized Google Distance (NGD). Average values were 0.35 for the *unrelated* condition, and 0.33 and 0.28 for *far* and *close* combinations, respectively. **Figure 3** (bottom panel) shows the relationship between NGD difference scores and behavioral PWI effects. A regression model with a linear relation resulted in a significant outcome,  $F_{(1, 30)} = 6.61$ ,  $p = 0.015$ ,  $R^2 = 0.18$ . By contrast, a RCS model with four knots was not significant,  $F_{(3, 28)} = 2.20$ ,  $p = 0.111$ ,  $R^2 = 0.19$ .

### DISCUSSION

In this experiment, an effect of relatedness emerged both in the latency and error analyses: in line with numerous published results in the literature, categorically related distractors interfered with target naming. As in the first experiment, however, no clear pattern emerged with regard to a semantic gradient: *close* and *far* conditions had quite similar average latencies, suggesting that the degree of overlap was largely irrelevant. Cumulative response time plots confirmed this pattern, with both related conditions slowing down responses across the entire spectrum. Furthermore, we explored whether the size of the interference effect generated by a particular picture-distractor combination could be predicted based on various measures of semantic relatedness. Relatedness ratings had no predictive power. LSA scores did not predict interference in a linear analysis, but showed a marginally significant result when modeled via RCS. NGD scores significantly predicted interference in the linear analysis. Overall, the pattern is that as in the first experiment, remarkably little variability in the PWI task is explained by measures of semantic overlap between target and distractor.

### GENERAL DISCUSSION

In two experiments we sought to replicate mutually contradictory previous findings regarding the possibility of a “semantic gradient” in PWI tasks: do strongly related targets and distractors induce more semantic interference than weakly related ones, as previously reported by Vigliocco et al. (2004), or does the opposite hold, as reported by Mahon et al. (2007)? Answering this question is of critical importance for the recent debate on whether lexical selection in spoken production is competitive or not. We duplicated the two previous key empirical studies with only very minor modifications (see Introduction) and found a general semantic interference effect: in both experiments, semantically related distractors slowed picture naming times, relative to unrelated distractors. However, in neither study did we find an

additional semantic gradient; the size of the semantic interference was not influenced by whether items were strongly or weakly related.

Our attempts to model a linear or non-linear relationship between interference effects and various measures of semantic relatedness rendered mixed results. In Experiment 1, only the non-linear relationship between semantic relatedness measures and interference approached statistical significance. **Figure 2** (top panel; red line) suggests that the degree of relatedness is largely irrelevant for the measured extent of PWI interference, except for a few items which are particularly strongly related and entail large interference effects. However, the corresponding analysis on the items from Experiment 2 show a different outcome, with (if at all) strongly related items creating *less* interference (although the RCS analysis was very far from significance). Could it be that the most strongly related target-distractor pairs in Experiment 2 were more related than the most strongly related pairs in Experiment 1? The relatedness ratings that we collected do not suggest that this was the case: average ratings were 5.7/7 for the *very close* condition in Experiment 1, and 6.1/7 for the *close* condition in Experiment 2. Direct comparison of the topmost panels of **Figures 2** and **3** also shows that strength of semantic overlap for the most strongly related pairs is quite similar.

In Experiment 2, two further results from the regression analysis were significant (or close to significance): first, an RCS model with LSA scores as the predictor rendered a marginally significant result. However, the pattern (red line in the middle panel of **Figure 3**) is not easily interpretable, and in any case does not resemble the likewise curve from Experiment 1 (red line in the middle panel of **Figure 2**). Second, a regression with a linear relation between NGD and interference returned a significant result (bottom panel of **Figure 3**). This pattern (more interference for items with stronger overlap as measured by NGD) goes against the predictions from the REH, but a similar pattern was not found in the first experiment (bottom panel of **Figure 2**). Overall, the most striking aspect of the regression results is how little of the variance is accounted for by any of the semantic relatedness measures, with all  $R^2 \leq 0.25$ .

These results admittedly are somewhat perplexing. Not only did the two existing key studies by Vigliocco and Mahon report contradictory results, but our own attempts to replicate them rendered results which are not compatible with either of the earlier findings (but results were consistent across our two experiments, such that in both experiments semantic interference was found, coupled with little additional effect of semantic relatedness). For the time being tentatively accepting our null finding concerning the effects of semantic distance in PWI tasks, how can the results be interpreted, and what are the implications for the current debate on whether or not lexical selection in spoken production is competitive?

A central component of REH is the claim that the response buffer is sensitive only to categorical membership but not to the degree of semantic similarity. Hence at this level both weakly and strongly related picture-word distractors should generate identical relatedness effects, as was found in Experiment 2 where close and far conditions had very similar latency means. Additionally, REH assumes that for strongly related items, conceptual priming

counteracts response buffer-based interference<sup>2</sup>. Given our null finding concerning a semantic gradient, a possibility for advocates of REH would be to drop the claim that there is conceptual priming, and state that semantic effects in PWI are exclusively response buffer-based. This is possible, but would then leave unexplained why some forms of semantic overlap (associatively related, part-whole, etc.; see Introduction) tend to generate facilitation effects in PWI tasks, given that conceptual priming was hypothesized to be the source of such effects.

Similarly, however, the results are not straightforwardly explained within the typical assumptions of competitive lexical selection. Although we are not aware of attempts to simulate the effects of varying semantic overlap in models of this type (such as WEAVER++), intuitively strongly related distractors should cause more interference than weakly related ones. Given the possibility that effects in PWI could reflect a combination of conceptual priming and lexical competition, perhaps models of this type could be specified such that the two contradictory forces cancel each other out: strongly related distractors cause substantial conceptual priming which facilitates target retrieval, yet also induce powerful competition which slows down target retrieval; weakly related distractors cause relatively less conceptual priming, but also less lexical competition. Only detailed computational simulations will show whether this possibility is feasible.

Finally, our results do not provide clear support for an alternative account of distractor interference which predicts greater interference from distantly related category members while maintaining the assumption of competitive lexical selection (Abdel Rahman and Melinger, 2009). Following this account, more closely related distractors should lead to a smaller cohort of items becoming activated during lexical selection, and consequently, less interference should be observed in comparison to the larger cohorts activation when more distantly related category members are presented. However, our results do not show the predicted negative semantic gradient.

As outlined in the Introduction, when setting up our studies we modified a number of relatively minor procedural details compared to the original experiments, mainly in order to render the two studies more similar and therefore comparable to each other. Based on the extensive literature on the PWI technique, there is no reason to believe that these variations could have critically influenced the results. For instance, Vigliocco et al. (2004) used an SOA of  $-150$  ms, whereas we used 0 ms. Could it be that under the negative SOA, a semantic gradient is present (as suggested by Vigliocco et al.), whereas under a simultaneous SOA

<sup>2</sup>On a strict reading of the response exclusion hypothesis, the assumption that target naming is delayed until the distractor has been removed from the response buffer predicts that latencies should exclusively depend on distractor processing, and any effects associated with target processing should be obliterated (Mulatti and Coltheart, 2012). For instance, the well-documented frequency effect in object naming (e.g., Jescheniak and Levelt, 1994) should disappear in a PWI context because processing of both high- and low-frequency target names are delayed until the distractor has been purged; however, that is clearly not the case (e.g., Miozzo and Caramazza, 2003, Experiment 1). Likewise, if target processing is speeded up due to conceptual priming, this should not affect latencies in a PWI task because the target will still have to wait until the distractor has been purged from the buffer.

degree of relatedness is not relevant (as suggested by our results)? Although this is not impossible, it would be a challenge to account for such a pattern within the theoretical frameworks currently available to explain PWI effects. Nevertheless, future work should perhaps aim to replicate the original studies to a closer extent than we accomplished.

A central component of our research approach was the attempt to predict semantic effects for individual picture-word pairs, based on a range of measures of semantic overlap. Human semantic relatedness ratings showed a reasonable degree of association with values obtained from LSA (Experiment 1:  $r = 0.564$ ,  $p < 0.001$ ; Experiment 2:  $r = 0.521$ ,  $p = 0.003$ ). However, one aspect of the findings which came as a surprise is that there was particularly low convergence between Normalized Google Distance (NGD) and the other two measures: NGD did not correlate with ratings (Experiment 1:  $r = 0.035$ ,  $p = 0.772$ ; Experiment 2:  $r = 0.183$ ,  $p = 0.316$ ) nor with LSA scores (Experiment 1:  $r = -0.098$ ,  $p = 0.415$ ; Experiment 2:  $r = -0.044$ ,  $p = 0.819$ ). Given the claim that NGD allows the automated discovery of meaning (Cilibrasi and Vitanyi, 2006, 2007), it is surprising that these correlations are so low. We conclude that NGD evidently gauges a different construct from the other two types of measures, one which is primarily sensitive to co-occurrence rather than to overlap in terms of semantic properties (indeed, the three related conditions in our first experiment had virtually identical NGD scores; the two related conditions in the second experiment were also relatively similar to each other).

A few other aspects of the results deserve discussion. In our second experiment, overall latencies (708 ms in the unrelated condition) were very similar to those reported by Mahon et al. (2007; 719 ms across their Experiments 5 and 5b). By contrast, there is a relatively large discrepancy between the unrelated mean of our first experiment (802 ms) and the one reported by Vigliocco et al. (642 ms). One possible contributing factor is that Vigliocco et al. used an SOA of  $-150$  ms, whereas we used one of  $0$  ms. It is well known in PWI studies that the mere presence of a visually presented distractor word (irrespective of semantic or form overlap with the target) tends to maximally interfere with target naming when onset of both target and distractor coincide. For instance, Damian and Martin (1999, Experiment 1) reported unrelated means of 700, 714, and 744 ms for SOAs of  $-200$ ,  $-100$ , and  $0$  ms. This gradient is most likely attributable to attentional interference when two stimulus dimensions are presented in extremely close succession or simultaneously.

A further aspect worth noting is that we observed a considerable difference in the latencies between our two experiments (an unrelated mean of 708 ms in Experiment 1, and one of 802 ms in Experiment 2). Given that we obtained target pictures mostly from the Snodgrass and Vanderwart (1980) set, recruited participants from the same pool, and aimed to render the two experiments procedurally as similar as possible, the reason for this speed difference is currently unclear. Finally, it must be noted that the size of the semantic effect in Experiment 1 (ignoring the degree of relatedness) was 40 ms, but in Experiment 2 it was only 21 ms. Again, the reasons for this variation are unclear. It is not implausible that overall slower latencies (as in the first, compared to the second, experiment) should be associated with

larger semantic interference effects, but this is unlikely to account for the observed size difference.

The lack of an effect of strength of relatedness in both experiments could be considered a null finding: the amount of interference generated by semantically related distractors in PWI is not influenced by semantic relatedness. Given the general difficulty in interpreting null findings, we sought to further explore the results of Experiment 1 via calculation of Bayes factors (Dienes, 2011), with the following line of reasoning. In Vigliocco et al., compared to the full-blown interference effect in the “very close” condition (29 ms), effects in the “close” and “medium” condition (15 and 6 ms, respectively) were reduced by 14 and 23 ms, or by 48 and 79% of the maximal value in the “very close” condition. In our own results, the “very close” condition resulted in interference of 42 ms; a reduction of this effect for the “close” and “medium” condition of the same size as found in Vigliocco et al. would predict values of 22 and 33 ms. Given the empirically obtained effects (reduction of 5 ms for the “close,” and of 0 ms for the “medium” condition) and corresponding standard errors, we were able to calculate Bayes Factors for the two conditions. We used the effects predicted from Vigliocco et al. as the mean of a normal distribution, with a standard deviation half that size as suggested by Dienes (2011, Supplementary Material). This resulted in Bayes factors of 0.77 for the “close” condition, and 0.37 for the “medium” condition. Based on the convention that Bayes factors at or below  $1/3$  are considered as substantial evidence supporting the null hypothesis, we argue that this is definitely the case for the “medium” condition, and less conclusively so for the “close” condition. A similar analysis was conducted on the results of Experiment 2: a Bayes Factor was calculated for the null difference between the effects generated by the “far” and the “close” condition. In Mahon et al. (averaged across their Experiments 5 and 5b), the effect for the “close” (7 ms) condition was reduced by 30 ms, or 82%, relative to the “far” condition (37 ms). For our own results this would predict an effect of 3 ms for the “close” condition (we found 23 ms). Given the same assumptions as in the corresponding analysis in Experiment 1, we obtained a Bayes Factor of 0.13, which constitutes strong evidence supporting the null hypothesis.

Additionally, critics might argue that the evidence provided here is inconclusive due to potential power issues. For the omnibus ANOVAs, the calculated *post-hoc* power to detect a medium-sized (0.25) effect given the included number of participants and conditions was 0.54 in Experiment 1, and 0.88 in Experiment 2. The regression analyses described in Sections Semantic relatedness ratings, Latent Semantic Analysis (LSA), and Normalized Google Distance (NGD) corresponded to analyses by items, hence, power is determined by the number of combinations which were included in the design, rather than the number of participants tested. For Experiment 1, an analysis returned a power value of 0.73 to determine a medium-sized (0.3) effect with this sample size. In Experiment 2, fewer items were included, and the power was 0.38, which is admittedly low. Note that as these two experiments sought to replicate the earlier studies by Vigliocco et al. (2004) and Mahon et al. (2007), we were restricted to using the original materials. Future work on the potential role of semantic distance effects in PWI tasks would be well advised to

substantially enlarge the included number of items (in addition to testing a large number of participants) in order to minimize the chances of type II errors.

Given that we were not able to fully resolve the issue of the role of semantic overlap in PWI tasks, the question arises of what could be the next step in tackling the problem. On balance, the design of Experiment 2 in which the same distractor words are used across all conditions is clearly preferable to the one in Experiment 1 in which only some, but not all, distractor words re-occurred across conditions. However, the number of items which we included in our Experiment 2 (see Section Materials) was admittedly low, and it would be desirable to replicate the design of this study with a considerably increased number of items. We believe that researchers should also consider an approach involving multiple regression, in which a large number of targets and distractors are shown in various combinations, category membership is binarily coded, and the question is to what extent residual variance in latencies can be attributed to semantic overlap measures derived from ratings etc. once category membership is taken into account. Overall, the issue of whether degree of semantic overlap matters in PWI tasks remains remarkably elusive, and we do not believe that the currently available results should be used to constrain theorizing on the nature of lexical selection in spoken word production.

## ACKNOWLEDGMENTS

This research was supported by postgraduate studentship PSYC.SC1801.6525 from the Economic and Social Sciences Research Council (ESRC) to James Hutson.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00872/abstract>

## REFERENCES

- Abdel Rahman, R., and Melinger, A. (2009). Semantic context effects in language production: a swinging lexical network proposal and a review. *Lang. Cogn. Process.* 24, 713–734. doi: 10.1080/01690960802597250
- Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511801686
- Belke, E., Meyer, A., and Damian, M. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Q. J. Exp. Psychol. A* 58, 667–692. doi: 10.1080/02724980443000142
- Blackford, T., Holcomb, P. J., Grainger, J., and Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition* 123, 84–99. doi: 10.1016/j.cognition.2011.12.007
- Bloem, I., and La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: implications for models of lexical access in language production. *J. Mem. Lang.* 48, 468–488. doi: 10.1016/S0749-596X(02)00503-X
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208. doi: 10.1080/026432997381664
- Caramazza, A., and Costa, A. (2000). The semantic interference effect in the picture-word interference paradigm: does the response set matter? *Cognition* 75, 51–64. doi: 10.1016/S0010-0277(99)00082-7
- Cilibrasi, R., and Vitanyi, P. (2006). “Similarity of objects and the meaning of words,” in *Theory and Applications of Models of Computation* eds J.-Y. Cai, S. B. Cooper and A. Li (Berlin; Heidelberg; Chicago, IL: Springer), 21–45. doi: 10.1007/11750321\_2
- Cilibrasi, R. L., and Vitanyi, P. M. (2007). The Google similarity distance. *IEEE Trans. Knowl. Data Eng.* 19, 370–383. doi: 10.1109/TKDE.2007.48
- Collins, A., and Quillian, M. (1969). Retrieval time from semantic memory. *J. Verb. Learn. Verb. Behav.* 8, 240–247. doi: 10.1016/S0022-5371(69)80069-1
- Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407
- Costa, A., Alario, F. X., and Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychon. Bull. Rev.* 12, 125–131. doi: 10.3758/BF03196357
- Costa, A., Miozzo, M., and Caramazza, A. (1999). Lexical selection in bilinguals: do words in the bilingual’s two lexicons compete for selection? *J. Mem. Lang.* 41, 365–397. doi: 10.1006/jmla.1999.2651
- Cree, G. S., and 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). *J. Exp. Psychol. Gen.* 132, 163–201. doi: 10.1037/0096-3445.132.2.163
- Damian, M. E., and Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychon. Bull. Rev.* 10, 111–117. doi: 10.3758/BF03196474
- Damian, M. F., and Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 345–361. doi: 10.1037/0278-7393.25.2.345
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283
- Dienes, Z. (2011). Bayesian versus orthodox statistics: which side are you on? *Perspect. Psychol. Sci.* 6, 274–290. doi: 10.1177/1745691611406920
- Finkbeiner, M., and Caramazza, A. (2006). Now you see it, now you don’t: on turning semantic interference into facilitation in a Stoop-like task. *Cortex* 42, 790–796. doi: 10.1016/S0010-9452(08)70419-2
- Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124. doi: 10.3758/BF03195503
- Glaser, W. R., and Glaser, M. O. (1989). Context effects in stroop-like word and picture processing. *J. Exp. Psychol. Gen.* 118, 13–42. doi: 10.1037/0096-3445.118.1.13
- Hantsch, A., Jescheniak, J. D., and Schriefers, H. (2005). Semantic competition between hierarchically related words during speech planning. *Mem. Cogn.* 33, 984–1000. doi: 10.3758/BF03193207
- Hantsch, A., Jescheniak, J. D., and Schriefers, H. (2009). Distractor modality can turn semantic interference into semantic facilitation in the picture-word interference task: implications for theories of lexical access in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1443–1453. doi: 10.1037/a0017020
- Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer. doi: 10.1007/978-1-4757-3462-1
- Heathcote, A., Popiel, S. J., and Mewhort, D. J. (1991). Analysis of response time distributions: an example using the Stroop task. *Psychol. Bull.* 109, 340–347. doi: 10.1037/0033-2909.109.2.340
- Humphreys, G. W., Lloyd-Jones, T. J., and Fias, W. (1995). Semantic interference effects on naming using a postcue procedure: tapping the links between semantics and phonology with pictures and words. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 961–980. doi: 10.1037/0278-7393.21.4.961
- Janssen, N., Schirm, W., Mahon, B. Z., and Caramazza, A. (2008). Semantic interference in a delayed naming task: evidence for the response exclusion hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 249–256. doi: 10.1037/0278-7393.34.1.249
- Jescheniak, J. D., and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 824–843. doi: 10.1037/0278-7393.20.4.824
- La Heij, W. (1988). Components of Stroop-like interference. *Mem. Cogn.* 16, 400–410. doi: 10.3758/BF03214220
- La Heij, W., Kuipers, J.-R., and Starreveld, P. A. (2006). In defense of the lexical-competition account of picture-word interference: a comment on Finkbeiner and Caramazza (2006). *Cortex* 42, 1028–1031. doi: 10.1016/S0010-9452(08)70209-0
- Landauer, T., Foltz, P., and Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Process.* 25, 259–284. doi: 10.1080/01638539809545028

- Levelt, W. J. M. (1999). Models of word production. *Trends Cogn. Sci.* 3, 223–232. doi: 10.1016/S1364-6613(99)01319-4
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. doi: 10.1017/S0140525X99001776
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., and Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 503–535. doi: 10.1037/0278-7393.33.3.503
- Miozzo, M., and Caramazza, A. (2003). When more is less: a counterintuitive effect of distractor frequency in the picture-word interference paradigm. *J. Exp. Psychol. Gen.* 132, 228–252. doi: 10.1037/0096-3445.132.2.228
- Mulatti, C., and Coltheart, M. (2012). Picture-word interference and the response-exclusion hypothesis. *Cortex* 48, 363–372. doi: 10.1016/j.cortex.2011.04.025
- Protopapas, A. (2007). CheckVocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behav. Res. Methods* 39, 859–862. doi: 10.3758/BF03192979
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86, 446–461. doi: 10.1037/0033-2909.86.3.446
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F
- Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: retrieval of verbs. *Cognition* 47, 59–87. doi: 10.1016/0010-0277(93)90062-Z
- Roelofs, A. (2001). Set size and repetition matter: comment on Caramazza and Costa (2000). *Cognition* 80, 283–290. doi: 10.1016/S0010-0277(01)00134-2
- Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychol. Rev.* 110, 88–125. doi: 10.1037/0033-295X.110.1.88
- Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *J. Mem. Lang.* 29, 86–102. doi: 10.1016/0749-596X(90)90011-N
- Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215. doi: 10.1037/0278-7393.6.2.174
- Spalek, K., Damian, M. F., and Bölte, J. (2012). Is lexical selection in spoken word production competitive? Introduction to the Special Issue on lexical competition in language production. *Lang. Cogn. Process.* 68, 69–84. doi: 10.1080/01690965.2012.718088
- Starreveld, P. A., and La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 686–698. doi: 10.1037/0278-7393.21.3.686
- Starreveld, P. A., and La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 896–918. doi: 10.1037/0278-7393.22.4.896
- Vigliocco, G., Vinson, D. P., Lewis, W., and Garrett, M. F. (2004). Representing the meanings of object and action words: the featural and unitary semantic space hypothesis. *Cogn. Psychol.* 48, 422–488. doi: 10.1016/j.cogpsych.2003.09.001
- Vitkovitch, M., and Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: it's in the links. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 664–680. doi: 10.1037/0278-7393.17.4.664
- Vitkovitch, M., and Tyrrell, L. (1999). The effects of distractor words on naming pictures at the subordinate level. *Q. J. Exp. Psychol. A* 52, 905–926. doi: 10.1080/713755854
- Woodworth, R. S., and Schlosberg, H. (1954). *Experimental Psychology, Rev. Edn.* New York, NY: Holt.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 April 2014; accepted: 22 July 2014; published online: 12 August 2014.

Citation: Hutson J and Damian MF (2014) Semantic gradients in picture-word interference tasks: is the size of interference effects affected by the degree of semantic overlap? *Front. Psychol.* 5:872. doi: 10.3389/fpsyg.2014.00872

This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Hutson and Damian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Bilinguals implicitly name objects in both their languages: an ERP study

Katie Von Holzen<sup>1,2\*</sup> and Nivedita Mani<sup>2</sup>

<sup>1</sup> Laboratoire Psychologie de la Perception UMR 8158, Université Paris Descartes, Paris, France

<sup>2</sup> Psychology of Language Research Group, Georg-Elias-Müller Institute of Psychology, University of Göttingen, Göttingen, Germany

## Edited by:

Ian FitzPatrick, Heinrich Heine  
Universität Düsseldorf, Germany

## Reviewed by:

Judith F. Kroll, Pennsylvania State  
University, USA  
Jasmin Sadat, Royal Holloway,  
University of London, UK

## \*Correspondence:

Katie Von Holzen, Laboratoire  
Psychologie de la Perception UMR  
8242, Université Paris Descartes,  
75006 Paris, France  
e-mail: katie.m.vonholzen@  
gmail.com

Upon being presented with a familiar name-known image, monolingual infants and adults implicitly generate the image's label (Meyer et al., 2007; Mani and Plunkett, 2010, 2011; Mani et al., 2012a). Although the cross-linguistic influences on overt bilingual production are well studied (for a summary see Colomé and Miozzo, 2010), evidence that bilinguals implicitly generate the label for familiar objects in both languages remains mixed. For example, bilinguals implicitly generate picture labels in both of their languages, but only when tested in L2 and not L1 (Wu and Thierry, 2011) or when immersed in their L2 (Spivey and Marian, 1999; Marian and Spivey, 2003a,b) but not when immersed in their L1 (Weber and Cutler, 2004). The current study tests whether bilinguals implicitly generate picture labels in both of their languages when tested in their L1 with a cross-modal ERP priming paradigm. The results extend previous findings by showing that not just do bilinguals implicitly generate the labels for visually fixated images in both of their languages when immersed in their L1, but also that these implicitly generated labels in one language can prime recognition of subsequently presented auditory targets across languages (i.e., L2–L1). The current study provides support for cascaded models of lexical access during speech production, as well as a new priming paradigm for the study of bilingual language processing.

**Keywords:** bilingualism, implicit naming, phonological priming, lexical access, ERP

## INTRODUCTION

Research on speech production has awarded considerable attention to the stages involved in a speaker's selection of an appropriate lexical item(s) to communicate her message. Among other issues, this work has examined how a speaker selects one word among other appropriate partially activated words for production, whether these other activated words interact with the speakers' choice and production of the chosen word, and the extent to which the phonological and semantic features of these other activated words are retrieved during speech production. Most models of speech production agree that the search for the appropriate lexical item in production also lends activation to items semantically related to the chosen word, either through activation of semantic features shared by the words or through activation of the corresponding lexical nodes of the semantically related words (Dell, 1986; Levelt, 1989; Roelofs, 1992; Caramazza, 1997). Models of speech production disagree, however, with regard to the extent to which the phonological features associated with these competing lexical nodes are retrieved in speaking. Discrete models of speech production suggest that while semantically related lexical nodes are simultaneously activated, phonological activation is restricted to the selected lexical node alone (Levelt, 1989; Levelt et al., 1999). In contrast, cascaded models of lexical access assume that the phonological properties of semantically related lexical nodes are all simultaneously activated (Dell, 1986; Caramazza, 1997; Dell et al., 1997).

Particularly useful for resolving the discrepancies between cascaded and discrete models is the study of bilingual speech

production. Between their two languages, bilinguals have many translation equivalents with varying levels of phonological overlap. One class of words, cognates, contain similar orthographic-phonological forms across languages. If, as argued by discrete models of speech production, only the phonological information for the corresponding node is activated, bilingual speech production should be similar for both cognate and non-cognate words. Studies investigating bilingual cognate and non-cognate picture naming, however, demonstrate a difference in naming latency between cognates and non-cognates (Costa et al., 2000; Hoshino and Kroll, 2008; Colomé and Miozzo, 2010; Strijkers et al., 2010; Poarch and van Hell, 2012). These results have overwhelmingly demonstrated that bilinguals activate phonological information from the non-target language, providing support for cascaded models of lexical access by showing that both selected and non-selected lexical nodes activate their corresponding phonological codes. Due to their special status across languages, however, the presence of cognate words may induce a bilingual processing mode (Wu and Thierry, 2010b). Stronger support for cascaded models of lexical access is therefore provided by studies not examining cognate word stimuli, yet still show that phonological information from both languages is activated during production in one language (Hermans et al., 1998; Colomé, 2001; Kaushanskaya and Marian, 2007; Hoshino and Thierry, 2011; Wu and Thierry, 2011). For example, Spalek and colleagues (Spalek et al., 2014) had German-English bilinguals produce adjective-noun pairs that either contained (e.g., *green goat*) or did not contain (e.g., *green skirt*) overt phonological onset overlap in English. Some trials,

however, although they did not overlap in English, did contain phonological onset overlap once translated to German (e.g., *blue flower*, “blaue Blume”). Trials that overlapped overtly in English and covertly in German modulated the ERP (event-related potential) response in comparison to non-related trials, suggesting that German translation equivalents of the English words were simultaneously activated and influenced production, despite the entire experiment being conducted in English.

In the current study, we examine a special instance of the retrieval of phonological and semantic information in the selection of lexical nodes for production. Unlike the body of research described previously, which focused on overt production, we turn to covert production, or implicit label generation. This allows for the study of the information bilinguals use to name, or activate upon visual fixation of, objects, before information is ultimately chosen for production. Specifically, we examine whether bilingual speakers implicitly produce the labels of visually fixated images and whether they do so in both their languages.

Upon viewing an image, studies suggest that the label for this image is implicitly generated, and that this implicitly generated label can prime recognition of a subsequently presented, related target. In Meyer et al.’s study (Meyer et al., 2007; see also Jescheniak et al., 2002; Meyer and Damian, 2007), for example, adults were presented with an unlabeled prime image (i.e., *boy*) followed by a visual display of four images, one of which was a homophone of the unlabeled picture prime (i.e., *buoy*). There was no semantic overlap between the homophone target and the prime image. Indeed, the only overlap between the two images lies in the labels for the two images—any preference for looking toward the homophone image could, therefore only be explained as a result of participants’ implicitly generating the label for both images and this implicitly generated label subsequently priming recognition of the related target image. Indeed, consistent with this explanation, participants were more likely to fixate the phonologically related target image compared to phonologically unrelated distractor images. This finding has been taken to show that participants implicitly generate the labels for visually fixated images, i.e., that they retrieve the phonological properties associated with the labels for visually fixated images. Such implicit label generation has also been found in infants using phonologically related prime-target pictures (Mani and Plunkett, 2010, 2011; Mani et al., 2012a), suggesting that auditory and visual information are integrated as young as 18-months-of-age.

Implicit label generation has also been explored using the ERP method in a cross-modal priming paradigm. Desroches et al. (2009) presented participants with picture prime—spoken target word pairs that were either identical, onset-overlapping, rhymes, or unrelated, while simultaneously measuring participants’ ERP responses to the spoken target words (see Mani et al., 2012b for similar studies with infants). ERPs (event-related potentials) are averaged waveforms of electrical brain activity (EEG) time-locked to the presentation of stimuli and can provide a measure of speech processing with a high temporal resolution of brain activity. One ERP component investigated by Desroches et al. (2009), the N400 (Kutas and Hillyard, 1984), described as a negative inflection peaking at approximately 400 ms after stimulus onset, indexes the integration of a stimulus into the context set by a preceding stimulus:

the larger the N400 amplitude, the more difficult the integration process between stimuli. Although Desroches et al., reported some variation in component latency, N400 amplitude was reduced for both onset overlapping and rhyming prime-target pairs. Using a cross-modal priming procedure, the authors argued, ensures that any priming effects were the result of top-down processes resulting from connections at the phonological and lexical levels instead of bottom-up influence due to acoustic overlap between prime and target. This conclusion also assumes that participants implicitly generated the label for the picture primes, which ultimately primed recognition of the spoken target word.

Unlike monolingual speakers, however, bilingual speakers have at least two labels for every object, one in one language (e.g. *dog*, English) and one in the other (e.g. *Hund*, German). When viewing objects, therefore, bilinguals may implicitly generate the label in one or both of their languages. In terms of overt speech production in bilinguals, cross-language effects have been found when participants are tested in both their L1 and L2 and immersed in their L1 (Costa et al., 2000; Colomé, 2001; Hoshino and Kroll, 2008; Colomé and Miozzo, 2010; Strijkers et al., 2010; Poarch and van Hell, 2012) or their L2 (Hermans et al., 1998; Costa et al., 2000, 2009; Kaushanskaya and Marian, 2007; Hoshino and Thierry, 2011; Spalek et al., 2014). Explicit naming or overt production may tap into different processes compared to implicit naming or covert production, particularly to do with the later stages of the speech monitoring process involved in overt production and due to delays introduced by the actual production of muscle movements (see Wu and Thierry, 2011 for similar suggestions) which may allow the required time for effects of L2 access to appear. It is, therefore, important to distinguish between findings of studies examining explicit and implicit naming. With regard to covert production, however, some evidence suggests that bilinguals may implicitly generate the labels for objects in both their languages, but that this depends on the language of testing, as well as whether they are immersed in their L1 or L2 (see Wu and Thierry, 2010b for a discussion of context in bilingual experiments): Previous studies have demonstrated that bilinguals implicitly label objects in both languages when they are immersed in their L2 and tested in L2, but results differ when participants are tested in their L1 (Spivey and Marian, 1999; Marian and Spivey, 2003a,b; Wu and Thierry, 2011). In the current study, using the cross-modal priming paradigm of Desroches et al. (2009), we examine whether bilinguals implicitly generate the label for objects in both of their languages when they are tested in their L1 and immersed in an L1 environment, a context that has previously failed to yield this effect (Weber and Cutler, 2004).

For bilinguals immersed in an L2 environment, successful L2 performance may come at the cost of L1 fluency. Linck et al. (2009) compared English learners of Spanish who were either immersed in a Spanish, L2 environment, or remained in their native, L1 English environment. When tested on both comprehension and production, an interesting asymmetry appeared between the two groups of participants: although the L2 performance was better for learners immersed in an L2 environment than their L1 environment counterparts, these participants showed decreased L1 access. This pattern of results led the authors to suggest that when immersed in their L2, the learners inhibited

activation of their L1. Within a group of participants tested before and after L2 immersion, however, Baus et al. (2013) found similar results, although only for low frequency, non-cognate words, suggesting that the decrease in L1 access during L2 immersion is the result of decreased L1 usage and not L1 inhibition. Although the purpose of the current paper is not to resolve which mechanisms are at work during L2 or L1 immersion, these studies highlight the effects that immersion can have on L1 and L2 access and performance.

To our knowledge, only one study has specifically investigated the question of whether bilinguals implicitly generate the label in one or both of their languages (although others have indirectly addressed it, see below). Wu and Thierry (2011) presented Chinese-English participants with pairs of pictures and asked them to judge whether the labels of the two pictures rhymed in L2 English (Experiment 1) or shared a character in L1 Chinese (Experiment 2). The stimuli were manipulated such that some of the picture pairs overlapped in one language (e.g. rhymed in L2 English), while others overlapped in the other language (e.g. character overlap in L1 Chinese). EEG data was recorded throughout the experiment to examine the neurocognitive indices of cross-language lexical access. Consistent with the standard N400 priming effect (Kutas and Hillyard, 1984), when asked to evaluate L1 Chinese overlap, participants found picture pairs whose labels overlapped in Chinese easier to process relative to unrelated picture pairs. Similarly, when asked to evaluate L2 English overlap, participants found picture pairs whose labels rhymed in English easier to process relative to unrelated pictures whose labels did not rhyme. Critically, when asked to evaluate L2 English overlap, picture pairs whose labels overlapped in L1 Chinese were also easier to process, suggesting that Chinese-English bilinguals activated both the L1 and L2 labels for the pictures. However, an effect of L2 English overlap was not found when participants were making rhyme judgments in L1 Chinese. Wu and Thierry attribute this asymmetric effect to the possibility that L2 word forms are not implicitly generated while making judgments in L1. In contrast, L1 word forms were activated during L2 processing and Wu and Thierry suggest that this is the result of bilinguals' inability to prevent interference from their L1 during L2 speech planning (Green, 1998).

Experiments using the visual world paradigm, however, suggest that both languages are activated even when bilingual participants are tested in their L1 and immersed in an L2 environment. In a series of experiments, Marian and Spivey (2003a,b) and Spivey and Marian (1999) presented Russian-English bilinguals with a visual display containing several objects. In one version of the experiment, participants were instructed in L1 Russian to move a target object (e.g. *marka*, "stamp"). Although they were tested in Russian, participants were more likely to look at a distractor object that had a phonologically related label in L2 English (e.g., *flomaster*, "marker") than a distractor object with an unrelated label (e.g., *lineka*, "ruler"). The results suggest that the word form of objects were also activated in L2 English, causing the English label for the phonologically related distractor object (i.e., *flomaster*, "marker") to compete for activation with the Russian label for the target object (i.e., *marka*, "stamp"). In another version of the experiment, where participants were instructed in L2

English, they were more likely to look at a target that was phonologically related in L1 Russian, than a distractor object with an unrelated label. In other words, bilinguals implicitly generated the label for the objects in both of their languages, regardless of the language of testing.

Nevertheless, participants in the Marian and Spivey experiments were immersed in their L2, where they heard their L2 every day in their surrounding environment. Using a similar paradigm to that of Marian and Spivey, Weber and Cutler (2004) extended the results of Marian and Spivey to participants tested in their L2 while immersed in their L1. Interestingly, Weber and Cutler did not find evidence of L2 activation when participants were tested in their L1 while immersed in an L1 environment. Weber and Cutler suggest that these results may reflect a difference in the background of their participants and the testing environment in comparison to the bilinguals tested by Spivey and Marian: The bilinguals tested by Spivey and Marian were immersed in their L2, English, perhaps increasing the likelihood that English would be co-activated. In contrast, the Dutch-English bilinguals tested by Weber and Cutler lived in the Netherlands and used their L1, Dutch, in their everyday life, making L2 English less relevant for activation when participants were tested in L1. It may, therefore, be more likely for L2 words to be activated during L1 processing when the participants are immersed in their L2.

It is of interest, however, that the Chinese-English bilinguals tested by Wu and Thierry (2011) do not show effects of L2 activation in L1 processing, despite being immersed in their L2 (similar to the language environments of Spivey and Marian, 1999; Marian and Spivey, 2003a,b). A potential explanation for this difference might come from nature of the task performed by participants. Spivey and Marian did not explicitly ask participants to judge the phonological overlap between target and distractor object labels in either of the languages of the participants, while Wu and Thierry focused participants' attention on phonological overlap for the picture pairs in one language, i.e., either their L1 or their L2. It is possible that this conscious focus on phonological overlap in the one language may reduce the influence of the "other" language, especially when the other language is the less dominant L2. The current study, therefore, does not focus participants' attention on phonological overlap in either of their languages. Instead, we asked participants to perform a non-linguistic task (a picture matching task), which drew their attention away from the relationship between the prime and target. We suggest that this provides a more accurate measure of whether bilinguals implicitly generate the labels of visually presented images in both their languages by not biasing their attention to linguistic relationships.

Using a cross-modal priming paradigm, participants were presented with visual picture primes (presented in silence) followed by auditory L1 targets. Although it differs from the picture-picture task of Wu and Thierry, it is similar to the visual world paradigm (Spivey and Marian, 1999; Marian and Spivey, 2003a,b; Weber and Cutler, 2004) where the target is a spoken word. This paradigm, shown to elicit an N400 component for both onset and rhyme related picture prime—target word pairs (Desroches et al., 2009), allows not only for the study of implicit label generation, but also removes the potential role of acoustic overlap between prime and target. This also allowed for an unbiased

investigation of cross-linguistic priming on auditory word recognition (i.e., L2 picture prime label—L1 auditory target). Although studies that have investigated auditory word recognition in bilinguals are increasing (Sinai and Pratt, 2002; Ju and Luce, 2004; Weber and Paris, 2004; Blumenfeld and Marian, 2005, 2007; Cutler et al., 2006; Marian et al., 2008; Rueschemeyer et al., 2008; Canseco-Gonzalez et al., 2010; FitzPatrick and Indefrey, 2010; Altvater-Mackensen and Mani, 2011; Lagrou et al., 2011; see also Shook and Marian, 2012; Weber and Broersma, 2012; Von Holzen and Mani, 2012; FitzPatrick and Indefrey, 2014), there are relatively few studies that have specifically investigated whether cross-linguistic priming can influence auditory word recognition (Phillips et al., 2006; Pratt et al., 2013). These studies, however, used both languages within their experiment, possibly creating an artificial bilingual environment (Grosjean, 1997). In the current study, we address this problem by testing participants exclusively in their L1.

In the current study, implicit generation of both language labels was examined by manipulating the relationship between the L1 and L2 labels for the picture prime and the L1 auditory target words. Thus, we included four conditions in the experiment: (1) picture prime labels and L1 target words were identical in L1 German; (2) L1 German labels for the picture primes and L1 target words were phonologically related in German; (3) L2 English labels for the picture primes and L1 target words sounded similar<sup>1</sup>; (4) L1 and L2 labels for the picture prime and L1 target words were phonologically, orthographically, or semantically unrelated (see **Figure 1** for examples). Similar to Wu and Thierry (2011), in trials where picture prime labels and L1 target words were related within- or between-language, the relationship was offset-overlap, which has already been demonstrated to elicit a N400 priming effect in a similar study with monolingual participants (Desroches et al., 2009). In summary, the current study aims to answer two questions:

1. When viewing images, do bilinguals implicitly generate the labels for these images in one or both of their languages?
2. Can implicitly generated labels in L2 influence processing of auditory words in L1?

Our index of these effects is obtained from Event-related potential (ERP) recordings of participants' brain activity as they heard the auditory target labels, focusing on the N400 component. In the study of bilingualism, the N400 waveform has been used as a measure of priming between prime-target pairs that are related across languages semantically (Alvarez et al., 2003; Phillips et al., 2004; Martin et al., 2009; Palmer et al., 2010), orthographically (De Bruijn et al., 2001), phonologically (Altvater-Mackensen and Mani, 2011), as well as through their translations (Thierry and Wu, 2007; Wu and Thierry, 2010a, 2011; see Moreno et al., 2008 for a review of ERP use in bilingualism). We suggest that,

<sup>1</sup>As complete phonological overlap between words across languages rarely occurs (Dijkstra et al., 1999), reducing our stimuli set to prime-target pairs that overlap completely across languages would have severely reduced our word choice for prime-target pairs for the between-language related condition. As a result, we chose prime-target pairs that sounded similar to one another in English and German, respectively.

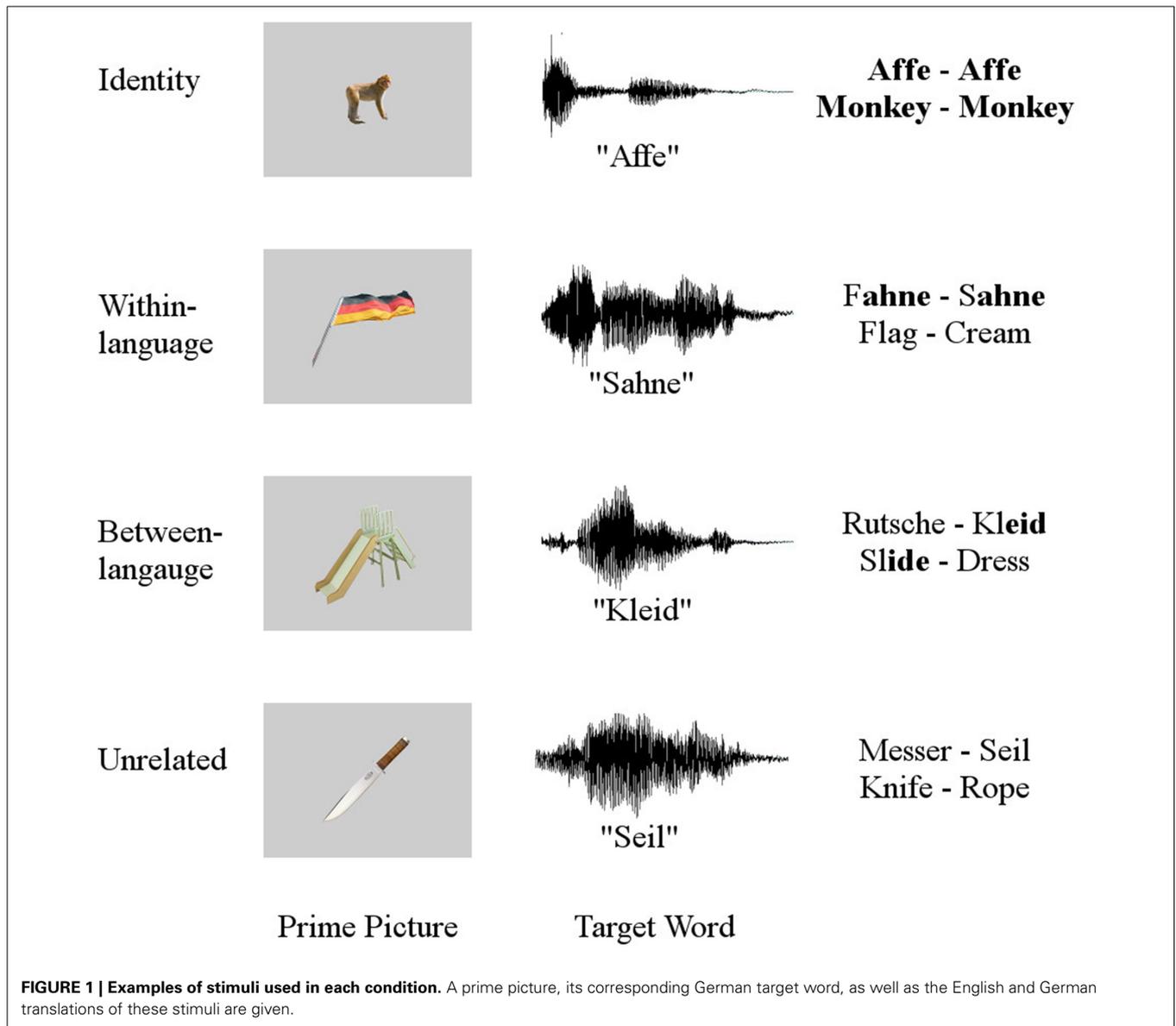
in contrast to the visual world paradigm, participants' ERPs may be a more sensitive index of the subtle differences involved in bilingual language processing (Mueller, 2005), given that participant eye gaze may be influenced by the number of objects in the visual display (Sorensen and Bailey, 2007) or simply not reflect competition effects (Dahan and Tanenhaus, 2004). In combination with the cross-modal priming paradigm, the use of ERPs may help clarify the differences in L2 activation during L1 processing when bilinguals are immersed in their L2 (Spivey and Marian, 1999; Marian and Spivey, 2003a,b) or their L1 (Weber and Cutler, 2004).

If participants were to implicitly generate the labels for the picture primes in their L1, this would be reflected by a significant reduction in N400 amplitude when the target word is either the same as the L1 label for the picture prime or phonologically related to the L1 label for the picture prime. Crucially, if bilinguals also implicitly generate the L2 labels for the picture primes, then N400 amplitude should also be reduced when the target word sounds similar to the L2 label for the picture prime. Furthermore, our manipulations allow for a comparison of priming effects for prime-target pairs that are related within- (e.g., L1 to L1) and between-languages (e.g., L2 to L1). This relates to a question within the study of bilingualism which asks whether a bilinguals' two languages are organized into two separate, but connected lexicons or are integrated into one large lexicon and would present evidence of cross-language priming in an experiment conducted entirely in one language.

## MATERIALS AND METHODS

### PARTICIPANTS

The current study tested a group of 16 German-English bilinguals (age  $M = 27.63$ ; age  $SD = 7.82$ ; age range = 20 – 48). Participants were recruited as bilinguals from the local population in a middle-sized German city and told that the purpose of the experiment was to examine their visual perception and were afterwards debriefed on the full purpose of the experiment. After the experiment, participants filled out a language proficiency questionnaire (adapted from Rueschemeyer et al., 2008). In this questionnaire, participants indicated the age at which they began learning both German and English. All participants had been exposed to both German and English before age 10. Participants also indicated their proficiency in reading, writing, listening, speaking, and syntax in German and English. These proficiency scores were averaged to create a combined proficiency score for both German and English. Participants reported an average combined proficiency score that was similar in both German ( $M = 9.19$ ;  $SD = 1.02$ ) and English ( $M = 9.05$ ;  $SD = 0.99$ ;  $p > 0.05$ ). In addition, participants also took part in a picture-naming task where we could assess the accuracy with which they labeled images in German and English. The results of these tests are reported in the Results section, showing that participants correctly and equally quickly labeled images in both German and English. All but three participants reported German to be their mother tongue and English their second language. Of these three, two participants reported German and English to be their mother tongue while one participant reported English to be her mother tongue, having learned German before she was 3 years of age.



Therefore, we consider German to be the L1 of the participants and English, their L2, although these two languages are balanced. Participants were living in Germany at the point of testing, immersed in their L1. Before the experiment participants signed an informed consent form approved by the ethics committee of the University Göttingen and received 15 Euros afterwards for their participation.

**STIMULI**

The stimuli consisted of 120 primes and 120 targets, resulting in 120 prime-target pairs. Primes were visually presented in silence, i.e., were presented as unlabeled, familiar images. Targets were presented auditorily, i.e., the picture prime was followed by an auditorily presented target word. A female, native speaker of German recorded all target words. The relationship between the labels for the prime image and the auditory target labels was manipulated to create four conditions: identity—picture prime

labels and L1 target words were identical in German (e. g., prime picture *monkey* “Affe”—target word *Affe*), within-language condition—L1 German labels for the picture primes and L1 target words were phonologically related in German (e. g., prime picture *flag* “Fahne”—target word *Sahne* “cream”), between-language condition—L2 English labels for the picture primes and L1 target words sounded similar (e. g., prime picture *slide* “Rutsche”—target word *Kleid* “dress”), or unrelated—L1 and L2 labels for the picture prime and L1 target words were phonologically, orthographically, or semantically unrelated (e. g., prime picture *knife* “Messer”—target word *Seil* “rope”). The 120 prime-target pairs were distributed across the four conditions with 30 pairs per condition. Prime and target words across languages were matched on the frequency of the words as well as the number of syllables and phonemes in the words ( $p$ 's > 0.05). **Figure 1** contains example stimuli from each condition. A list of stimuli can be found in the Supplementary Material.

## PROCEDURE

### Main experiment

Participants were seated in a dimly lit, quiet experimental room, facing a 92 cm wide and 50 cm high TV screen at a distance of 100 cm from the screen. All instructions given to participants, including the written instructions presented on an instruction sheet, were in German. Participants were presented with 120 trials distributed across the four conditions, with 30 trials per condition. Each trial began with a fixation cross displayed in the center of the screen for 1000 ms. Following the offset of the fixation cross, participants were presented with the prime image centrally located on the screen. The prime image remained on screen for 500 ms (from 1000 to 1500 ms into the trial) followed by a blank, black screen. At 1550 ms into the trial, 50 ms after the offset of the prime picture, participants were presented with an auditory target word. At 3000 ms (i.e., 1450 ms after the onset of the target label), a second image was displayed that was either identical to the prime image or was different to the prime image. This image remained on-screen for 500 ms (from 3000 to 3500 ms into the trial) and was followed by a blank, black screen for a further 1000 ms (from 3500 to 4500 ms into the trial). Participants were instructed to indicate in this interval (from 3500 to 4500 ms into the trial) whether the second image matched the first image presented or not, by pressing one of two buttons in front of them. Participants were informed that the experiment investigated the mechanisms underlying their perception of the similarity between the two images presented. They were informed that they would hear spoken words during the experiment but that their task was to ignore these spoken words. This was done in order to avoid any overt attention to the relationship between the labels for the prime images and the target words.

### Production task

Following the main experiment, participants also completed a production task, where they were asked to label a series of 60 images aloud in both German and English. Half of the participants labeled the images in German first, while the other half labeled in English first. Stimuli used in the production task were the prime images from the within- and between-language conditions. The answers provided by the participants were automatically recorded via a microphone, time-locked to the appearance of the image on-screen. Production data was analyzed offline to determine whether participants labeled the prime images from these conditions with the label we had chosen for each picture. If the label provided by a participant was different from the chosen label for an image, then the trial containing this image was removed from the main experiment. This was to ensure that the labels implicitly generated by individual participants were related to the target words in the two critical conditions. For example, in the between-language condition, the picture prime “beagle” could also be given the label “dog.” However, the label “dog” does not sound similar to the target word “Igel” and therefore no longer fulfills the between-language manipulation. For the between-language condition, this removed 13% of trials (71 trials) and for the within-language condition 10.34% of trials (60 trials)<sup>2</sup>.

<sup>2</sup>An analysis on the data without removing trials that were produced correctly showed the same pattern of results as the analysis reported.

## EEG RECORDING AND ANALYSIS

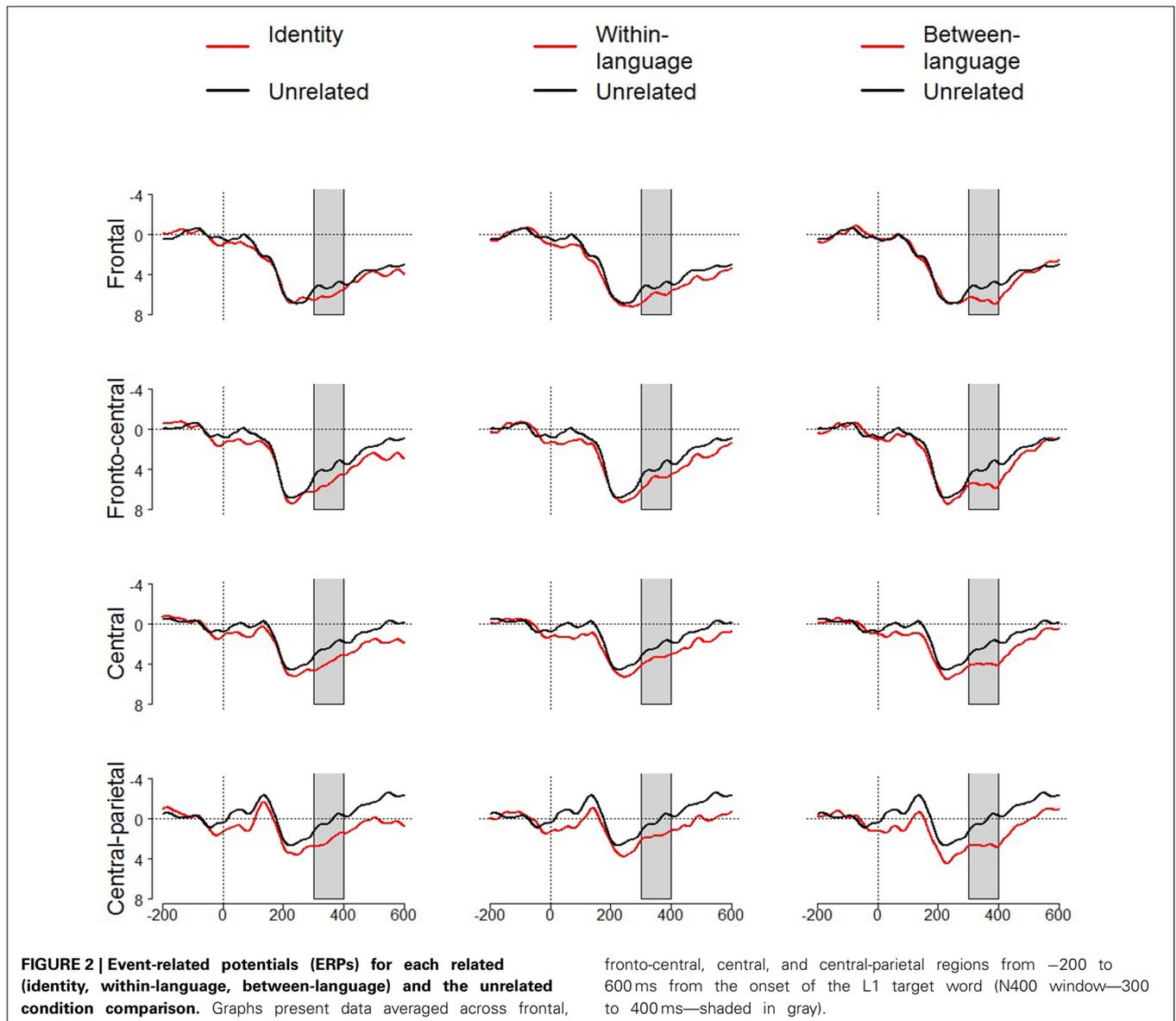
Electrophysiological data was recorded using the Biosemi Active Two Amplifier system at a sampling rate of 2048 Hz from 32 Ag/AgCl electrodes placed according to the 10–20 convention. Electrode offsets were kept at less than 25  $\mu$ V. Electroencephalogram was re-referenced offline to the averaged mastoid reference. EEG data was then filtered off-line using a 0.1 Hz high-pass forward filter and a 20 Hz low-pass, zero-phase shift filter.

Averaging and artifact rejection were carried out using the BESA software (Version 5.3). Blink and movement artifacts were automatically rejected using a 100 Hz amplitude cut-off across all electrodes. Epochs were defined from  $-200$  to 1000 ms from the onset of the auditory target word. We then analyzed the data in 50 ms time windows (from 0 to 1000 ms) to determine the onset and offset of significant differences between conditions. Based on this analysis, and the known onset of the N400 (Kutas and Hillyard, 1984), we focused our analyses on the time window between 300 and 400 ms (the N400 window; Desroches et al., 2009).

Average ERP waveforms were quantified by computing mean amplitudes per subject, electrode and condition in the selected time windows. ERP waveforms were baseline corrected by subtracting the mean amplitude for the baseline time window ( $-200$  to 0 ms) from the selected time window. For purposes of data reduction, a selection of electrode locations was entered into data analysis, 16 electrodes divided into four regions and four lateral columns: frontal (F7, F3, F4, F8), fronto-central (FC5, FC1, FC2, FC6), central (T7, C3, C4, T8), and centro-parietal (CP5, CP1, CP2, CP6). Our analysis was based on specific planned comparisons between related conditions (identical, within- and between-language conditions) and the unrelated condition instead of a general condition effect; we therefore do not report the omnibus ANOVA (see Abelson and Prentice, 1997). Factors included in the repeated measures ANOVA were region (frontal, fronto-central, central, and centro-parietal), electrode laterality (4), and condition (2; related, unrelated).

## RESULTS

**Figure 2** displays ERP waveforms aggregated across electrodes, separated by region (frontal, fronto-central, central, centro-parietal) for each of the three condition comparisons (identity vs. unrelated/within-language vs. unrelated/between-language vs. unrelated). We first examined the difference in the mean N400 amplitude of the brain potentials following identity and unrelated prime-target pairs. Here, a significant main effect of condition revealed that mean N400 amplitude was reduced for identity prime-target pairs relative to unrelated prime-target pairs,  $F_{(1, 15)} = 6.20$ ,  $p = 0.025$ ,  $\eta_p^2 = 0.29$ . No other interactions with condition were significant ( $ps > 0.25$ ). Planned *post-hoc* analyses revealed that mean N400 amplitude for identity prime-target pairs was reduced relative to unrelated prime-target pairs across all regions, i.e., at frontal,  $t_{(15)} = 2.23$ ,  $p = 0.042$ ,  $d = 0.25$ , fronto-central,  $t_{(15)} = 2.95$ ,  $p = 0.01$ ,  $d = 0.37$ , and central  $t_{(15)} = 2.46$ ,  $p = 0.026$ ,  $d = 0.43$ , regions, and approached significance for the centro-parietal region,  $t_{(15)} = 1.92$ ,  $p = 0.074$ ,  $d = 0.48$ . As expected, complete match between the label for the prime image and the target word in the identity condition resulted



in easier processing of the target word, relative to when the prime label was unrelated to the target word.

Next, we examined the difference in mean N400 amplitude for within-language related prime-target pairs and unrelated prime-target pairs. A repeated-measures ANOVA revealed a significant main effect of condition,  $F_{(1, 15)} = 4.54, p = 0.05, \eta_p^2 = 0.23$ , suggesting less negative N400 amplitude to L1 targets preceded by primes whose L1 labels were phonologically related to the L1 target words, relative to targets preceded by unrelated primes. No other interactions with condition were significant ( $p$ 's > 0.56). Planned *post-hoc* analyses revealed that mean N400 amplitude for within-language prime-target pairs was reduced, relative to unrelated prime-target pairs, across all regions, significant at frontal,  $t_{(15)} = 2.16, p = 0.047, d = 0.23$ , and fronto-central,  $t_{(15)} = 2.25, p = 0.04, d = 0.24$ , regions and approached significance at the centro-parietal region,  $t_{(15)} = 1.83, p = 0.088, d = 0.32$ , but not the central region ( $p > 0.12$ ). In line with

predictions, participants implicitly generated the label for the prime image in their L1, which speeded processing of the phonologically related L1 target word.

Finally, we examined the difference in mean N400 amplitude for between-language related prime-target pairs and unrelated prime-target pairs. An ANOVA comparing N400 amplitude across between-language related pairs and unrelated prime-target pairs revealed a near-significant main effect of condition,  $F_{(1, 15)} = 4.25, p = 0.057, \eta_p^2 = 0.22$ . No other interactions with condition were significant ( $p$ 's > 0.18). Planned *post-hoc* analyses revealed that mean N400 amplitude for between-language prime-target pairs was reduced, relative to unrelated prime-target pairs, across all regions, significant at central,  $t_{(15)} = 2.37, p = 0.032, d = 0.40$ , and centro-parietal,  $t_{(15)} = 2.13, p = 0.05, d = 0.53$ , regions, but not frontal or fronto-central regions ( $p$ 's > 0.12). The only way that the prime image in between-language related prime-target pairs could influence recognition of the target,

would be if participants were to *also* implicitly generate the label for the prime image in their L2, and for this implicitly generated L2 label to speed processing of the phonologically related L1 target word.

To examine whether there was any differences in the magnitude of the priming effect across related conditions, further analyses compared ERPs to targets between the related conditions. ANOVAs comparing N400 amplitude for identity and within-language, identity and between-language, and within- and between-language prime-target pairs revealed no significant main effect of condition ( $p > 0.38$ ) or interactions with condition ( $p$ 's  $> 0.15$ ), except for a significant interaction between condition X electrode laterality for the comparison between the identity and between-language conditions,  $F_{(3, 45)} = 3.76$ ,  $p = 0.031$ ,  $\eta_p^2 = 0.20$ . Paired-samples  $t$ -tests, however, revealed that there was no significant difference between the identity and between-language conditions at each of the lateral columns ( $p$ 's  $> 0.3$ ). The lack of a significant difference between the related conditions suggests a similar magnitude in the priming effect independent of a within- or between-language relationship between prime labels and target words.

In addition, we also measured the accuracy with which participants labeled the within and between-language related primes in German and English as well as the latency to name the images. In this task, conducted after the main experiment, participants were asked to overtly name the prime pictures from the within- and between-language conditions in German and English. Trials with no response were considered technical errors (German: 1.67%; English: 3.13%) and not included in the analysis. We considered responses correct if they accurately labeled the picture, regardless of the label we chose for the picture (e.g., *cat* or *kitty* for the prime picture *kitten*). Participants gave an incorrect response for 2.50% of trials in German and 2.70% of trials in English. Trials containing images whose labels participants incorrectly labeled were excluded from the analysis. There was no difference in accuracy for participants when they labeled pictures in German (96.83%) or English (94.17%;  $p > .3$ ). In addition, there was no difference in reaction time for pictures named in German ( $M = 712.20$  ms;  $SD = 188.24$  ms) and English ( $M = 728.00$  ms;  $SD = 185.05$  ms;  $p > 0.75$ ). Taken together, participants showed no difference between German and English in the production task.

## DISCUSSION

The current study asked whether bilingual adults implicitly generate the label for words in one or both of their languages, and whether implicitly generated L2 labels can prime L1 auditory words. These findings suggest that bilingual adults implicitly generate the label for visually fixated images in both of their languages, and that this implicitly generated label can, subsequently influence recognition of an auditorily presented, similar sounding L1 target word. These results significantly extend previous findings to strongly support suggestions that (a) implicit generation of the labels of visually fixated images in both languages of bilinguals immersed in their L1 environment and (b) L2 prime labels influencing recognition of L1 target words, despite the experiment being carried out in an L1 environment with only L1 stimuli being used in the experiment. This presents a robust

test of the extent to which bilinguals implicitly generate picture labels in their L2 as well as their L1 in an experiment situation that discourages such activation.

### PICTURE LABELS ARE IMPLICITLY GENERATED IN L1 AND L2

Prime-target pairs whose labels were either identical (identity; i.e., prime picture *monkey* “Affe”—target word *Affe*) or phonologically related within L1 (within-language; i.e., prime picture *flag* “Fahne”—target word *Sahne* “cream”) elicited a reduction in N400 amplitude, suggesting that participants implicitly generated the label for prime pictures in their L1. That is, the implicitly generated L1 label subsequently primed the L1 target word as a result of the complete or phonological overlap between prime and target. This replicates previous studies with monolingual adults (Meyer et al., 2007; Desroches et al., 2009) and infants (Mani and Plunkett, 2010, 2011) that show that prime pictures presented in silence activate their labels and corresponding phonological information, priming subsequently presented identical or phonologically-related targets.

Critical to the current study's research questions, the reduction in N400 amplitude for L1 target words preceded by prime pictures whose label in L2 English was similar sounding to the L1 target word (between-language; i.e., prime picture *slide* “Rutsche”—target word *Kleid* “dress”), suggests that bilingual participants also implicitly generated the prime picture label in their L2. This demonstrates that bilinguals implicitly generate the labels for objects in not one, but both of their languages. In a previous study, Wu and Thierry (2011) asked Chinese-English bilinguals to preform rhyme judgments on picture pairs, some of which were rhyme pairs in either Chinese or English. Priming effects were elicited for rhyme pairs in both languages when participants were tested in L2 English, but when participants were tested in L1 Chinese, only Chinese rhyme pairs elicited priming effects. The results of Wu and Thierry suggest that whether participants implicitly generated picture labels in one or both languages, therefore, depended on the language they were tested in. Wu and Thierry conclude that this asymmetry shows that spoken language planning (i.e., implicit label generation) in L1 proceeds without activating L2 word information, but that bilinguals are unable to prevent L1 interference during L2 speech planning.

In the current study, however, we find that when participants are tested in their L1, they implicitly generate the label not only in L1, as Wu and Thierry found, but also in L2. We suggest that the difference between the current study and that of Wu and Thierry is the result of the tasks which participants completed during the experiment. The bilingual participants tested by Wu and Thierry were instructed to make rhyme judgments for picture pairs. This task required participants to focus on the linguistic relationship between the prime and target pictures and narrow their focus to one language in order to successfully complete the task. When tested in L1, participants were better at narrowing their focus and preventing interference from L2, but this was not the case when participants were test in L2 and as a result L1 words were also activated. In the current study, bilingual participants judged whether the picture prime and a subsequently presented picture (after the target word) were the same or different. This task did not require participants to pay attention to the relationship between picture

prime and target word. Unlike the participants tested by Wu and Thierry, the participants tested in the current study did not need to narrow their focus to one language in order to successfully complete their task. We suggest that this difference in task is the reason we find implicit label generation in both languages when participants were tested in L1, while Wu and Thierry did not. It is especially useful for future research that implicit label generation in bilinguals can be studied without using a task that calls attention to the relationship between prime and target.

We note that our results also contrast with previous work by Weber and Cutler (2004) who report that bilinguals do not activate their L2 in L1 processing, when immersed in an L1 dominant environment. In our study, similar to Weber and Cutler (2004), participants were immersed in their L1 and tested in their L1. In this situation, the only relevant language is L1. Yet, our results demonstrate that participants implicitly generated the label for prime pictures in both L1 and L2. One possible explanation is the potential difference in L1 and L2 use between the Dutch-English bilinguals tested by Weber and Cutler, and the German-English bilinguals tested in the current study. Previous studies have found that an L2 immersion context has a positive influence on L2 proficiency and performance compared to L2 classroom exposure, although this comes at the cost of L1 fluency (Linck et al., 2009; see also Baus et al., 2013). Faced with the task of L2 usage every day, bilinguals may inhibit their L1 (Green, 1998) in order to perform successfully in their L1 or this may simply be the result of reduced frequency of L1 use (Gollan et al., 2005, 2008). In the context of L1 immersion, L2 fluency may also experience a reduction, which would account for the findings of Weber and Cutler but not those of the current study. The difference, then, would lie in the usage of L2 English in the different L1 contexts of Dutch and English. Although comparisons of English proficiency and frequency of use across cultures are at the moment anecdotal at best, such considerations are crucial to the future study of bilingual language processing. Alternatively, it is possible that the use of a more sensitive paradigm to assess participants' access to L2 words, i.e., the cross-modal ERP priming paradigm, may have allowed us to tap into cross-language effects that could not be observed in Weber and Cutler. Indeed, a number of studies have shown that such subtle cross-language effects do not lead to noticeable differences in responding while triggering different patterns of neural activity (Kotz and Elston-Güttler, 2004; McLaughlin et al., 2004; Tokowicz and MacWhinney, 2005; Thierry and Wu, 2007; Wu and Thierry, 2010a; for a review see Mueller, 2005).

We suggest that, while future studies are needed to compare different language combinations and environments, the cross-modal ERP priming paradigm used in the current study is useful tool for measuring bilingual co-activation and allows us to tap into subtle effects of other language activation in bilingual language processing. This finding, taken together with other studies using different methods (Spivey and Marian, 1999; Wu and Thierry, 2011; Von Holzen and Mani, 2012) suggests that activation of both languages during processing in one language is a powerful phenomenon. While viewing pictures, bilinguals activate the corresponding label as well as its phonological information in both of their languages. Although such a result cannot

be generalized outside of an experimental setting, it affords a glimpse into the complex processes that bilinguals undergo while interacting with their environment.

#### CROSS-LANGUAGE PRIMING IN AN L1 EXPERIMENTAL SETTING

The current study also provides evidence for cross-language priming in bilinguals such that implicitly generated L2 labels facilitate recognition of auditory L1 targets despite the experiment being conducted entirely in one language. Although relatively unexplored in auditory word processing (but see Phillips et al., 2006; Pratt et al., 2013), previous studies have also found similar L2-L1 priming effects for phonologically related prime-target pairs in visual word processing (Van Wijnendaele and Brysbaert, 2002; Duyck, 2005; Zhou et al., 2010). These studies, however, used both languages in their experiments, and priming effects may, therefore, result from the artificial bilingual environment created by using both languages in the experimental setting (Grosjean, 1997). In contrast, the current study was conducted in one language. The between-language priming effect, therefore, cannot be attributed to an artificial bilingual environmental setting and the current results present the first evidence for L2-L1 priming in auditory word processing in an unbiased setting. This a useful tool for future studies to continue studying cross-language phono-lexical effects in bilinguals without presenting word stimuli auditorily or visually in both languages.

#### IMPLICATIONS FOR MODELS OF LEXICAL ACCESS DURING SPEECH PRODUCTION

The two main findings of the current study, namely that bilinguals implicitly generate the labels for pictures in both of their languages and cross-language phonological priming in participants immersed in their L1, provide an interesting addition to the debate on the kind of information that is activated during word production. It is generally accepted that speech production involves first activating the lexical node associated with the concept, followed by the corresponding phonological code (Levelt, 1989; Roelofs, 1992; Caramazza, 1997; Dell et al., 1997). Conflict abounds, however, with regard to whether the phonological information of non-selected lexical nodes is also activated. Discrete serial models of lexical access suggest that only the phonological information associated with the selected lexical node is activated (Levelt, 1989; Levelt et al., 1999). For example, when naming a picture of a *table*, semantically related lexical nodes are also activated (i.e., *couch*, *chair*). Ultimately, the lexical node for *table* is selected and the corresponding phonological information is activated, but not for the non-selected lexical nodes (i.e., *couch*, *chair*). Cascaded activation models of lexical access, in contrast, propose that the phonological information from both the selected and non-selected lexical nodes is activated (Dell, 1986; Caramazza, 1997; Dell et al., 1997).

The findings of the current study and Wu and Thierry (2011), provide useful information with regard to the processes underlying picture naming and, by extension, the extent to which conceptual and phonological levels of representations are recruited in speaking. Neither of these studies used prime-target pairs with semantic overlap (although Wu and Thierry did include a semantically related condition, prime-target pairs in the critical

rhyming conditions were not additionally semantically related). Nevertheless, both studies demonstrate that participants activate phonological information associated with non-selected lexical nodes in implicit generation of the labels for visually fixated images. In other words, participants activated phonological information associated with both L1 and L2 labels for the images. If purely semantic information associated with non-selected lexical nodes were activated, as suggested by discrete serial models (Levelt, 1989; Levelt et al., 1999), we would have expected priming effects only in the identity condition. But, the priming effect for the between-language condition in the current study provides support for cascaded activation models of lexical access (Dell, 1986; Caramazza, 1997; Dell et al., 1997), by demonstrating that phonological information from both selected (L1) and non-selected (L2) lexical nodes was activated when our participant viewed the prime pictures. Our study, furthermore, goes beyond Wu and Thierry (2011) in showing that participants activated L2 labels for the prime images despite being immersed in their L1 and tested in their L1, thereby reducing the likelihood that L2 lexical nodes would need to be retrieved.

### CONCEPTUAL ACTIVATION DURING IMPLICIT LABEL GENERATION IN BILINGUALS

The results of the current study have interesting implications for models of bilingual speech processing with regard to the activation of language representations from conceptual representations during picture viewing. The Revised Hierarchical Model is a model of bilingual word production that focuses on the connections between L1 and L2 words at the lexical and conceptual levels and how these connections develop as proficiency increases (Kroll and Stewart, 1994; Kroll and Dijkstra, 2002; Kroll et al., 2010). According to this model, connections exist between L1 and L2 translation equivalents at the lexical level. At the conceptual level, however, connections initially exist only between L1 words and their concepts. Access to conceptual representations during L2 processing must, therefore, be mediated through L1 translation equivalents. With greater L2 proficiency, access to conceptual representations during L2 processing may continue without mediation through L1 translation equivalents, with direct links between L2 words and their conceptual representations.

Models of speech production argue that upon viewing a picture, participants activate the conceptual representations associated with this image, leading to phonological activation of either one or many selected lexical nodes associated with the activated conceptual representations. Thus, evidence of the activation of the L2 label for the image might be taken to suggest that, in proficient bilinguals, conceptual representations activated are directly linked to L2 labels such that viewing the picture leads to activation of conceptual representations which in turn directly activate both L1 and L2 labels. Alternatively, it is possible that, even in proficient bilinguals, L2 words are only indirectly linked to conceptual representations such that the results of the current study are explained as follows: viewing the picture leads to activation of conceptual representations, which in turn activate the L1 label, leading to mediated activation of the L2 label from the L1 label. While our study cannot rule out this explanation entirely, we note that there were no differences in the time-course or strength of the

effects across the within-language and between-language overlap conditions, suggesting that such mediated activation through L1 labels is unlikely. To this extent, the results of the current study support the suggestion of the RHM that, in proficient bilinguals, L2 words are directly linked to their concepts.

Alternatively, we note that the results could also be explained without relying on access to the conceptual level. For example, the visual features of a picture may directly activate its label, such that, upon viewing a picture, the visual features of the picture activate the corresponding labels in both languages. Activation of word labels from picture viewing relies, in this case, not on the activation of conceptual representations but rather on the recognition of the visual features of an image. It will be interesting, therefore, for future studies to explore the extent to which conceptual representations are involved in the progression from image recognition to implicit naming/production. We also note that regardless of whether L2 labels were activated directly from conceptual representations or from the L1 labels, the results of the current study strongly suggest that bilinguals implicitly produce the labels for visually fixated images in both their languages, even when immersed in an L1 setting and when tested in their dominant language.

### CONCLUSION

The current study presented evidence that bilinguals implicitly generate the label for pictures in both of their languages. Previous studies have shown mixed results, suggesting that an L1 language environment (visual world paradigm; Weber and Cutler, 2004) or experimental task that requires participants to focus on the linguistic relationship between prime and target (rhyming judgment task; Wu and Thierry, 2011) may prevent co-activation of L2 words. By using a cross-modal ERP priming paradigm, we demonstrated not only that bilinguals implicitly generate the label for pictures in both of their languages, but also that implicitly generated L2 labels can prime related L1 words. The results provide support for cascaded activation models of lexical access, showing that phonological information associated with non-selected lexical nodes is retrieved during (implicit) picture naming.

### ACKNOWLEDGMENTS

This work was funded by the German Excellence Initiative Award to Georg-August-Universität Göttingen (Third funding line: Institutional Strategy). The authors would like to thank Nicole Altwater-Mackensen and Susan Bobb for their helpful comments on earlier version of this manuscript as well as all of the subjects who participated in the study.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.01415/abstract>

### REFERENCES

- Abelson, R. P., and Prentice, D. A. (1997). Contrast tests of interaction hypotheses. *Psychol. Methods* 2, 315–328.
- Altwater-Mackensen, N., and Mani, N. (2011). “Bilinguals activate words from both languages when listening to spoken sentences: evidence from an ERP-study,” in *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*,

- eds L. Carlson, C. Hoelscher, and T. F. Shipley (Austin, TX: Cognitive Science Society), 1382–1387.
- Alvarez, R. P., Holcomb, P. J., and Grainger, J. (2003). Accessing word meaning in two languages: an event-related brain potential study of beginning bilinguals. *Brain Lang.* 87, 290–304. doi: 10.1016/S0093-934X(03)00108-1
- Baus, C., Costa, A., and Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychol.* 142, 402–409. doi: 10.1016/j.actpsy.2013.01.010
- Blumenfeld, H. K., and Marian, V. (2005). “Covert bilingual language activation through cognate word processing: an eye-tracking study,” in *Proceedings of the XXVII Annual Conference of the Cognitive Science Society* (Stresa), 286–291.
- Blumenfeld, H. K., and Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Lang. Cogn. Proces.* 22, 633–660. doi: 10.1080/01690960601000746
- Canseco-Gonzalez, E., Brehm, L., Brick, C. A., Brown-Schmidt, S., Fischer, K., and Wagner, K. (2010). Carpet or Cárcel: the effect of age of acquisition and language mode on bilingual lexical access. *Lang. Cogn. Proces.* 25, 669–705. doi: 10.1080/01690960903474912
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208. doi: 10.1080/026432997381664
- Colomé, À. (2001). Lexical activation in bilinguals’ speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736. doi: 10.1006/jmla.2001.2793
- Colomé, À., and Miozzo, M. (2010). Which words are activated during bilingual word production? *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 96–109. doi: 10.1037/a0017677
- Costa, A., Caramazza, A., and Sebastian-Gallés, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1283–1296. doi: 10.1037/0278-7393.26.5.1283
- Costa, A., Strijkers, K., Martin, C. D., and Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21442–21446. doi: 10.1073/pnas.0908921106
- Cutler, A., Weber, A., and Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *J. Phon.* 34, 269–284. doi: 10.1016/j.wocn.2005.06.002
- Dahan, D., and Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 498–513. doi: 10.1037/0278-7393.30.2.498
- De Bruijn, E. R. A., Dijkstra, T., Chwilla, D. J., and Schriefers, H. J. (2001). Language context effects on interlingual homograph recognition: evidence from event-related potentials and response times in semantic priming. *Biling. Lan. Cogn.* 4, 155–168. doi: 10.1017/S1366728901000256
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., and Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychol. Rev.* 104, 801–838.
- Desroches, A. S., Newman, R. L., and Joanisse, M. F. (2009). Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. *J. Cogn. Neurosci.* 21, 1893–1906. doi: 10.1162/jocn.2008.21142
- Dijkstra, T., Grainger, J., and van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: the neglected role of phonology. *J. Mem. Lang.* 41, 496–518. doi: 10.1006/jmla.1999.2654
- Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: evidence for nonselective phonological activation in bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1340–1359. doi: 10.1037/0278-7393.31.6.1340
- FitzPatrick, I., and Indefrey, P. (2010). Lexical competition in nonnative speech comprehension. *J. Cogn. Neurosci.* 22, 1165–1178. doi: 10.1162/jocn.2009.21301
- FitzPatrick, I., and Indefrey, P. (2014). Head start for target language in bilingual listening. *Brain Res.* 1542, 111–130. doi: 10.1016/j.brainres.2013.10.014
- Gollan, T. H., Montoya, R., Cera, C., and Sandoval, T. (2008). More use almost always means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J. Mem. Lang.* 58, 787–814. doi: 10.1016/j.jml.2007.07.001
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., and Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Mem. Cogn.* 33, 1220–1234. doi: 10.3758/BF03193224
- Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133
- Grosjean, F. (1997). “Processing mixed language: issues, findings and models,” in *Tutorials in Bilingualism: Psycholinguistic Perspectives*, eds A. M. B. de Groot and J. F. Kroll (Mahwah, NJ: LEA), 225–254.
- Hermans, D., Bongaerts, T., De Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. Lang. Cogn.* 1, 213–229. doi: 10.1017/S1366728998000364
- Hoshino, N., and Kroll, J. F. (2008). Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106, 501–511. doi: 10.1016/j.cognition.2007.02.001
- Hoshino, N., and Thierry, G. (2011). Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain Res.* 1371, 100–109. doi: 10.1016/j.brainres.2010.11.053
- Jescheniak, J. D., Schriefers, H., Garrett, M. F., and Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *J. Cogn. Neurosci.* 14, 951–964. doi: 10.1162/089892902760191162
- Ju, M., and Luce, P. A. (2004). Falling on sensitive ears: constraints on bilingual lexical activation. *Psychol. Sci.* 15, 314–318. doi: 10.1111/j.0956-7976.2004.00675.x
- Kaushanskaya, M., and Marian, V. (2007). Bilingual language processing and interference in bilinguals: evidence from eye tracking and picture naming. *Lang. Learn.* 57, 119–163. doi: 10.1111/j.1467-9922.2007.00401.x
- Kotz, S. A., and Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *J. Neurolinguist.* 17, 215–235. doi: 10.1016/S0911-6044(03)00058-7
- Kroll, J. F., and Dijkstra, T. (2002). “The bilingual lexicon,” in *Handbook of Applied Linguistics*, ed R. Kaplan (Oxford: Oxford University Press), 301–321.
- Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008
- Kroll, J. F., van Hell, J. G., Tokowicz, N., and Green, D. W. (2010). The revised hierarchical model: a critical review and assessment. *Biling. Lang. Cogn.* 13, 373–381. doi: 10.1017/S136672891000009X
- Kutas, M., and Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307, 161–163. doi: 10.1038/307161a0
- Lagrou, E., Hartsuiker, R. J., and Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 952–965. doi: 10.1037/a0023217
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75.
- Linck, J., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x
- Mani, N., Durrant, S., and Floccia, C. (2012a). Activation of phonological and semantic codes in toddlers. *J. Mem. Lang.* 66, 612–622. doi: 10.1016/j.jml.2012.03.003
- Mani, N., Mills, D. L., and Plunkett, K. (2012b). Vowels in early words: an event-related potential study. *Dev. Sci.* 15, 2–11. doi: 10.1111/j.1467-7687.2011.01092.x
- Mani, N., and Plunkett, K. (2010). In the infant’s mind’s ear: evidence for implicit naming in 18-month-olds. *Psychol. Sci.* 21, 908–913. doi: 10.1177/0956797610373371
- Mani, N., and Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition* 121, 196–206. doi: 10.1016/j.cognition.2011.06.013
- Marian, V., Blumenfeld, H. K., and Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *J. Psycholinguist. Res.* 37, 141–170. doi: 10.1007/s10936-007-9064-9
- Marian, V., and Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Appl. Psycholinguist.* 24, 173–193. doi: 10.1017/S0142716403000092
- Marian, V., and Spivey, M. (2003b). Competing activation in bilingual language processing: within- and between-language competition. *Biling. Lang. Cogn.* 6, 97–115. doi: 10.1017/S1366728903001068

- Martin, C. D., Dering, B., Thomas, E. M., and Thierry, G. (2009). Brain potentials reveal semantic priming in both the “active” and the “non-attended” language of early bilinguals. *Neuroimage* 47, 326–333. doi: 10.1016/j.neuroimage.2009.04.025
- McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat. Neurosci.* 7, 703–704. doi: 10.1038/nn1264
- Meyer, A. S., Belke, E., Telling, A. L., and Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychon. Bull. Rev.* 14, 710–716. doi: 10.3758/BF03196826
- Meyer, A. S., and Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Mem. Cogn.* 35, 494–503. doi: 10.3758/BF03193289
- Moreno, E. M., Rodriguez-Fornells, A., and Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *J. Neurolinguist.* 21, 477–508. doi: 10.1016/j.jneuroling.2008.01.003
- Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Sec. Lang. Res.* 21, 152–174. doi: 10.1191/0267658305sr256oa
- Palmer, S. D., van Hooff, J. C., and Havelka, J. (2010). Language representation and processing in fluent bilinguals: electrophysiological evidence for asymmetric mapping in bilingual memory. *Neuropsychologia* 48, 1426–1437. doi: 10.1016/j.neuropsychologia.2010.01.010
- Phillips, N. A., Klein, D., Mercier, J., and de Boysson, C. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Res.* 1125, 116–131. doi: 10.1016/j.brainres.2006.10.002
- Phillips, N. A., Segalowitz, N., O’Brien, I., and Yamasaki, N. (2004). Semantic priming in a first and second language: evidence from reaction time variability and event-related brain potentials. *J. Neurolinguist.* 17, 237–262. doi: 10.1016/S0911-6044(03)00055-1
- Poarch, G. J., and van Hell, J. G. (2012). Cross-language activation in children’s speech production: evidence from second language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 111, 419–438. doi: 10.1016/j.jecp.2011.09.008
- Pratt, H., Abbasi, D. A.-A., Bleich, N., Mittelman, N., and Starr, A. (2013). Spatiotemporal distribution of cortical processing of first and second languages in bilinguals. II. Effects of phonologic and semantic priming. *Hum. Brain Mapp.* 34, 2882–2898. doi: 10.1002/hbm.22109
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F
- Rueschemeyer, S.-A., Nojack, A., and Limbach, M. (2008). A mouse with a roof? Effects of phonological neighbors on processing of words in sentences in a non-native language. *Brain Lang.* 104, 132–144. doi: 10.1016/j.bandl.2007.01.004
- Shook, A., and Marian, V. (2012). The bilingual language interaction network for comprehension of speech. *Biling. Lang. Cogn.* 16, 304–324. doi: 10.1017/S1366728912000466
- Sinai, A., and Pratt, H. (2002). Electrophysiological evidence for priming in response to words and pseudowords in first and second language. *Brain Lang.* 80, 240–252. doi: 10.1006/brln.2001.2597
- Sorensen, D. W., and Bailey, K. G. D. (2007). The world is too much: effects of array size on the link between language comprehension and eye movements. *Visual Cogn.* 14, 112–115. doi: 10.1080/13506280600975486
- Spalek, K., Hoshino, N., Wu, Y. J., Damian, M., and Thierry, G. (2014). Speaking two languages at once: unconscious native word form access in second language production. *Cognition* 133, 226–231. doi: 10.1016/j.cognition.2014.06.016
- Spivey, M. J., and Marian, V. (1999). Cross talk between native and second languages: partial activation of an irrelevant lexicon. *Psychol. Sci.* 10, 281–284. doi: 10.1111/1467-9280.00151
- Strijkers, K., Costa, A., and Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cereb. Cortex* 20, 912–928. doi: 10.1093/cercor/bhp153
- Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12530–12535. doi: 10.1073/pnas.0609927104
- Tokowicz, N., and MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Stud. Sec. Lang. Acquis.* 27, 173–204. doi: 10.1017/S0272263105050102
- Van Wijnendaele, I., and Brysbaert, M. (2002). Visual word recognition in bilinguals: phonological priming from the second to the first language. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 616–627. doi: 10.1037//0096-1523.28.3.616
- Von Holzen, K., and Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *J. Exp. Child Psychol.* 113, 569–586. doi: 10.1016/j.jecp.2012.08.001
- Weber, A., and Broersma, M. (2012). “Spoken word recognition in second language acquisition,” in *The Encyclopedia of Applied Linguistics*, ed C. A. Chapelle (Bognor Regis: Wiley-Blackwell), 5356–5361. doi: 10.1002/9781405198431.wbeal1104
- Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *J. Mem. Lang.* 50, 1–25. doi: 10.1016/S0749-596X(03)00105-0
- Weber, A., and Paris, G. (2004). “The origin of the linguistic gender effect in spoken-word recognition: evidence from non-native listening,” in *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, eds K. Forbus, D. Gentner, and T. Tegier (Mahwah, NJ: Erlbaum), 1446–1451.
- Wu, Y. J., and Thierry, G. (2010a). Chinese-english bilinguals reading english hear Chinese. *J. Neurosci.* 30, 7646–7651. doi: 10.1523/JNEUROSCI.1602-10.2010
- Wu, Y. J., and Thierry, G. (2010b). Investigating bilingual processing: the neglected role of language processing contexts. *Front. Psychol.* 1:178. doi: 10.3389/fpsyg.2010.00178
- Wu, Y. J., and Thierry, G. (2011). Event-related brain potential investigation of preparation for speech production in late bilinguals. *Front. Psychol.* 2:114. doi: 10.3389/fpsyg.2011.00114
- Zhou, H., Chen, B., Yang, M., and Dunlap, S. (2010). Language nonselective access to phonological representations: evidence from Chinese-English bilinguals. *Q. J. Exp. Psychol.* 63, 2051–2066. doi: 10.1080/17470211003718705

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 June 2014; accepted: 19 November 2014; published online: 09 December 2014.

Citation: Von Holzen K and Mani N (2014) Bilinguals implicitly name objects in both their languages: an ERP study. *Front. Psychol.* 5:1415. doi: 10.3389/fpsyg.2014.01415 This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Von Holzen and Mani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations

João M. Correia\*, Bernadette Jansma, Lars Hausfeld, Sanne Kikkert and Milene Bonte

Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht Brain Imaging Center (M-BIC), Maastricht University, Maastricht, Netherlands

## Edited by:

Peter Indefrey, University of Dusseldorf, Germany

## Reviewed by:

Jonas Obleser, Max Planck Institute for Human Cognitive and Brain Sciences, Germany  
Dirk Koester, Bielefeld University, Germany

## \*Correspondence:

João M. Correia, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht Brain Imaging Center (M-BIC), Maastricht University, Oxfordlaan 55, 2nd floor, Room 014, 6229 EV Maastricht, Netherlands  
e-mail: joao.correia@maastrichtuniversity.nl

Spoken word recognition and production require fast transformations between acoustic, phonological, and conceptual neural representations. Bilinguals perform these transformations in native and non-native languages, deriving unified semantic concepts from equivalent, but acoustically different words. Here we exploit this capacity of bilinguals to investigate input invariant semantic representations in the brain. We acquired EEG data while Dutch subjects, highly proficient in English listened to four monosyllabic and acoustically distinct animal words in both languages (e.g., “paard”–“horse”). Multivariate pattern analysis (MVPA) was applied to identify EEG response patterns that discriminate between individual words within one language (within-language discrimination) and generalize meaning across two languages (across-language generalization). Furthermore, employing two EEG feature selection approaches, we assessed the contribution of temporal and oscillatory EEG features to our classification results. MVPA revealed that within-language discrimination was possible in a broad time-window (~50–620 ms) after word onset probably reflecting acoustic-phonetic and semantic-conceptual differences between the words. Most interestingly, significant across-language generalization was possible around 550–600 ms, suggesting the activation of common semantic-conceptual representations from the Dutch and English nouns. Both types of classification, showed a strong contribution of oscillations below 12 Hz, indicating the importance of low frequency oscillations in the neural representation of individual words and concepts. This study demonstrates the feasibility of MVPA to decode individual spoken words from EEG responses and to assess the spectro-temporal dynamics of their language invariant semantic-conceptual representations. We discuss how this method and results could be relevant to track the neural mechanisms underlying conceptual encoding in comprehension and production.

**Keywords:** EEG decoding, EEG oscillations, speech perception, spoken word recognition, bilinguals, semantic representations, conceptual representation

## INTRODUCTION

Speech processing is a surprisingly flexible and accurate cognitive ability that allows humans to comprehend spoken language in real-time. At the individual word level, speech processing requires a continuous mapping of complex and variable auditory input signals to words and their semantic-conceptual representations. In turn, when we speak, we start from ideas and concepts and convert these into articulatory motor programs. In multilingual environments, these transformations involve the extraction of unified semantic concepts from variable acoustic/phonological word forms in native and non-native languages. When and how the bilingual brain performs these language-invariant conceptual transformations remains essentially unknown and is a focus of the present electroencephalography (EEG) study.

EEG allows studying non-invasively and with high temporal resolution the neural dynamics of speech processing. The

temporal dynamics of EEG signals are informative of temporal order effects during speech processing. ERP (event-related potential) components at early time intervals, 100–200 ms after word onset, have been associated with phonetic/phonological processing (Dumay et al., 2001; Sanders and Neville, 2003; Bonte and Blomert, 2004; Uusvuori et al., 2008). Intermediate time intervals (200–300 ms) have been suggested to reflect early aspects of lexical access (Van den Brink et al., 2001; Hagoort et al., 2004; Salmelin, 2007; Bonte et al., 2009), followed by lexical/semantic processing in the 300–600 ms window, as indicated by ERP modulations dependent on semantic attributes of words, semantic priming and semantic context (Kutas and Hillyard, 1980; Hagoort, 2008). Spatially, this temporal signature of speech processing may reflect a spread of information from primary auditory areas to anterior temporal and frontal regions, mid-inferior and posterior temporal regions (Marinkovic et al., 2003) corresponding to the

network of brain areas observed in functional magnetic resonance imaging (fMRI) studies of speech processing (Binder et al., 2000; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Complementary to ERP modulations, the oscillatory dynamics of EEG signals measured extracranially (Hagoort et al., 2004; Shahin et al., 2009; Doelling et al., 2014; Strauß et al., 2014) and intracranially (Luo and Poeppel, 2007; Giraud and Poeppel, 2012) have provided important insights regarding the function of underlying neural oscillations. Namely, an entrainment of theta band oscillations to the phoneme/syllable rate of speech signals, and the entrainment of gamma band oscillations to the phase of such theta band oscillations are suggested to reflect synchronization mechanisms that optimize the parsing of the speech signal into its relevant units (Lakatos et al., 2005; Giraud and Poeppel, 2012; Obleser et al., 2012; Peelle and Davis, 2012).

A challenge is to investigate how these temporal and oscillatory EEG dynamics encode the representation of specific speech units, such as individual words and concepts. Recently, methods based on machine learning comprising multivariate statistics (MVPA, multivariate pattern analysis, Formisano et al., 2008a; Haxby et al., 2011) have shown their potential to solve this challenge. MVPA of EEG signals extends traditional univariate methods by exploiting the interaction between multiple signal features (e.g., spectro-temporal features across multiple electrodes and/or time points) using classification algorithms (Chan et al., 2011b; Hausfeld et al., 2012; Herrmann et al., 2012; Brandmeyer et al., 2013). The higher sensitivity of MVPA to find information content within brain imaging signals has significantly contributed to our understanding of the brain's responses to speech and language. In fMRI studies, multi-voxel patterns across early and higher-order auditory cortex have been shown to successfully predict the (perceptual) identity of individual speech sounds and speaker's voices (Formisano et al., 2008b; Kilian-Hütten et al., 2011; Bonte et al., 2014). Furthermore, fMRI responses in inferior parietal areas have been shown to differentiate words across different semantic categories [e.g., tools and dwellings, Shinkareva et al. (2011)]. At a more fine-grained within-category level, MVPA was recently shown to accurately predict which spoken noun a bilingual listener was listening to in one language (e.g., "horse" in English) based on the fMRI response patterns to equivalent nouns in the other language (e.g., "paard" in Dutch; Correia et al., 2014). This generalization of the meaning of words across languages specifically relied on focal regions, including the left anterior temporal lobe (left-ATL), suggesting the existence of "hub" regions organizing semantic-conceptual knowledge in abstract form (Damasio et al., 1996; Scott et al., 2000; Patterson et al., 2007; Visser and Lambon Ralph, 2011; Correia et al., 2014). Although more challenging in terms of the robustness of single trial estimates, also spatially/temporally distributed EEG/MEG patterns have been observed to discriminate individual speech sounds (Hausfeld et al., 2012), and words from different perceptual and semantic categories (Simanova et al., 2010; Chan et al., 2011b; Sudre et al., 2012). Classification performances in EEG-MVPA studies on speech processing are typically low [e.g., below 0.55 in binary classification of spoken vowels, Hausfeld et al. (2012); or below 0.60 in binary classification of spoken words, Simanova et al. (2010)]. Besides the low signal-to-noise

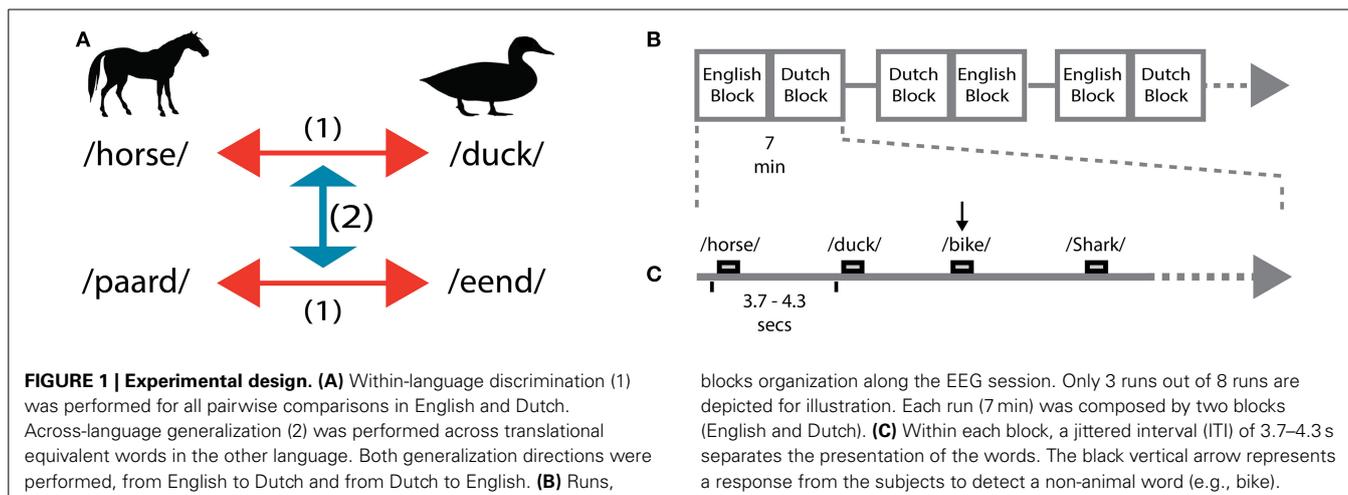
ratio of single trial EEG signals, EEG-based classification of individual words may be limited by the continuous and temporally variable processing of their phonological and semantic features (Van Petten et al., 1999). Importantly, however, multivariate approaches in EEG allow unraveling subtle differences in the neural processing of individual speech sounds that remain obscured in univariate approaches relying on average activation differences between experimental conditions.

Here, we employ MVPA to investigate spectro-temporal EEG response patterns capable of discriminating semantic-conceptual representations of words at the fine-grained level of within-category distinctions (animal nouns). To this end, we exploit the unique capacity of bilingual subjects to access semantic-conceptual information of spoken words from two languages. In separate Dutch and English blocks, we asked bilingual participants to listen to individual animal nouns (66.6% trials) and to detect non-animal target nouns (33.3% trials). The non-animal target nouns were presented as control task to ensure speech comprehension at every word presentation, but were not included in the analysis. Following supervised machine learning approaches, we trained multivariate classifiers (linear-SVM) to predict the identity of the perceived animal noun from new (untrained) samples of EEG activity (**Figure 1A**). In a first analysis we aimed to identify the EEG correlates involved in within-language word discrimination. To this end we trained classifiers to discriminate EEG responses to English (e.g., "horse" vs. "duck") and Dutch (e.g., "paard" vs. "eend") nouns. Importantly, stimuli included three exemplars of each noun, pronounced by three different female speakers, allowing for speaker-invariant word discrimination ("*within-language*"). In a second analysis we aimed to assess the EEG correlates involved in language-independent decoding of the animal nouns ("*across-language*"). Here we trained classifiers to discriminate EEG responses to words in one language (e.g., in English, "horse" vs. "duck") and tested whether this training generalizes and allows discrimination of EEG responses to the corresponding nouns in the other language (e.g., in Dutch, "paard" vs. "eend"). Importantly, all words were acoustically-phonetically distinct both within and across languages. Based on this approach, we aimed to investigate whether language-independent representations are detectable in the EEG responses to individual spoken words. In particular, this approach allowed us to extract critical time windows and frequency ranges within the EEG relevant to semantic-conceptual encoding.

## METHODS

### PARTICIPANTS

Sixteen native Dutch (L1) participants proficient in English (L2) took part in the study (8 males and 8 females, right-handed, mean age = 28.9 SD = 3.4). The participants were undergraduate or post-graduate students of Maastricht University studying or working in an English speaking environment. All participants reported normal hearing abilities and were neurologically healthy. English proficiency was assessed with the LexTALE test, a vocabulary test including 40 frequent English words and 20 non-words (Lemhöfer and Broersma, 2012). The mean test score was 89.6% correct (SD = 11.2%). This score is well above the average score (70.7%) of a large group of Dutch and Korean advanced learners



of English performing the same test (Lemhöfer and Broersma, 2012). For comparison reasons, participants also conducted the Dutch version of the vocabulary test. The mean Dutch proficiency score was 94.1% (SD = 3.3). The study was approved by the Ethical Committee of the Faculty of Psychology and Neuroscience at the University of Maastricht, The Netherlands.

## STIMULI

Stimuli consisted of Dutch and English spoken words representing four different animals (English: “Bull,” “Duck,” “Horse,” and “Shark,” and the Dutch equivalents: “Stier,” “Eend,” “Paard,” and “Haai”) and six inanimate object words (English: “Bike,” “Coat,” “Dress,” “Road,” “Suit,” and “Town”; and the Dutch equivalents: “Fiets,” “Jas,” “Jurk,” “Weg,” “Pak,” and “Stad”). All animal nouns were monosyllabic and acoustically/phonetically distinct from each other both within and across languages. Phonetic distance between word pairs was quantified using the Levenshtein distance, which gives the number of phoneme insertions, deletions and/or substitutions required to change one word into the other, divided by the number of phonemes of the longest word (Levenshtein, 1965). On a scale from 0 (no changes) to 1 (maximum number of changes), the mean (SD) Levenshtein distances corresponded to 0.83 (0.15) for Dutch word pairs, 0.93 (0.12) for English word pairs and 1.00 (0.00) for English-Dutch word pairs. Furthermore, all animal nouns had an early age of acquisition in Dutch (mean = 5.28 years SD = 0.98; De Moor et al., 2000) and a medium-high frequency of use expressed on a logarithmic scale in counts per million tokens in Dutch (mean = 1.29 SD = 0.71) and in English [mean = 1.50 SD = 0.42; Celex database, Baayen et al. (1995)]. To add acoustic variability and allow for speaker-invariant MVPA analysis, the words were spoken by three female native Dutch speakers with good English pronunciation. Stimuli were recorded in a sound proof chamber at a sampling rate of 44.1 kHz (16 bit resolution). Post-processing of the recorded stimuli was performed in PRAAT software (Boersma and Weenink, 2013) and included band-pass filtering (80–10,500 Hz), manual removal of acoustic transients (clicks), length equalization, removal of sharp onsets and offsets using 30 ms ramp envelopes, and amplitude equalization (average

RMS). Stimulus length was equated to 600 ms (original range: 560–640 ms) using PSOLA (75–400 Hz as extrema of the F0 contour). We carefully checked the stimuli for possible alterations in F0 after length equation and did not find any detectable changes. We assured that the produced stimuli were unambiguously comprehended by the participants during the stimuli familiarization phase prior to the experiment.

## EXPERIMENTAL PROCEDURES

The experimental session was organized in 8 runs, each run containing 2 blocks (one Dutch and one English). Each block included 36 nouns: 24 animal nouns and 12 (33.3%) non-animal nouns. The order of English and Dutch blocks was counter-balanced across runs: odd runs started with an English block followed by a Dutch block; even runs started with a Dutch block followed by an English block (Figure 1B). Participants were instructed to actively listen to the stimuli and to press a button (with the left index finger) whenever they heard a non-animal word. The goal of the task was to help maintaining a constant attention level throughout the experiment and to promote speech comprehension at every word presentation. All participants paid attention to the words as indicated by a mean (SD) detection accuracy of 98.3 (1.4) %. Data from non-animal trials were excluded from further analysis. The 24 animal nouns in each block corresponded to 6 repetitions of each of the 4 animal nouns. Because nouns were pronounced by 3 different speakers, each physical stimulus was repeated twice in each block. Stimulus presentation was pseudo-randomized within each block, avoiding consecutive presentations of the same words or sequences of words. Throughout the experiment, each animal noun was presented 48 times per language.

## EEG ACQUISITION AND PREPROCESSING

Data were recorded with a sampling rate of 250 Hz in an electrically shielded and sound-proof room from 62 electrode positions (Easycap, Montage Number 10, 10–20 system) relative to a left mastoid reference signal. The ground electrode was placed on the Fz electrode. Impedance levels were kept below 5 k $\Omega$ . During the EEG measurement, stimuli were presented binaurally at a comfortable intensity level. According to an event-related design

(**Figure 1C**), the averaged inter-trial-interval between two stimuli was 4 s (jittered randomly between 3.7 s and 4.3 s). Each run took 7 min, resulting in a total EEG measurement time of 56 min. A gray fixation cross against a black background was used to keep the visual stimulation constant during the whole duration of a block. Block and run transitions were marked with written instructions. Participants were instructed to minimize eye-movements during the auditory presentation and fixate on the fixation cross.

Data preprocessing was performed using EEGLab (Delorme and Makeig, 2004) and included band-pass filtering (0.1–100 Hz) followed by epoch extraction locked to the onset of the animal nouns (–1000 to 1000 ms) and baseline correction (–1000 to 0 ms).

Removal of signal artifacts was performed in two steps. First, the data were visually inspected and epochs containing non-stereotypical artifacts including high-amplitude, high-frequency muscle noise, swallowing, and electrode cable movements, were rejected (mean 4.31 trials per subject, SD 2.36). Second, stereotypical artifacts related to eye movements, eye-blinks and heart beat artifacts were corrected with extended INFOMAX ICA (Lee et al., 1999) as implemented in EEGLAB. Because data were recorded at 62 channels, runica decomposed the data in 62 component activations per subject. These component activations were categorized as EEG activity or non-brain artifacts by visual inspection of their scalp topographies, time courses, and frequency spectra. Criteria for categorizing component activations as EEG activity included (1) a scalp topography consistent with an underlying dipolar source, (2) spectral peak(s) at typical EEG frequencies, and (3) regular responses across single trials, i.e., an EEG response should not occur in only a few trials (Delorme and Makeig, 2004). Based on these criteria, component activations representing non-brain artifacts were removed, and EEG data were reconstructed from the remaining component activations representing brain activity. The resulting ICA-pruned data sets were baseline corrected (–1000 to 0 ms) and used for further analysis.

### ERP AND ERSP ANALYSIS

First, in order to validate typical EEG responses to spoken words reported in the literature, we performed univariate analyses. These were conducted in EEGLab (Delorme and Makeig, 2004) and included: (1) an ERP analysis based on the average amplitude of signal change over time with respect to baseline (–1000 to 0 ms) and (2) an ERSP (event-related spectral perturbation) analysis based on averaged power changes of all words over frequency and time with respect to baseline (–1000 to 0 ms). For the ERSP analysis we employed a Hanning taper fast fourier transform (FFT) filter from 1 to 60 Hz on a linear frequency scale with steps of 2 Hz, producing 30 filtered signals. Group statistics for the ERP and ERSP were performed at random-effects using two-sided Wilcoxon tests for each time-point vs. zero baseline and corrected for multiple comparisons using FDR (alpha = 5%).

### MULTIVARIATE CLASSIFICATION ANALYSIS

Multivariate classification was employed to investigate whether specific temporal or spectrotemporal EEG signatures enable the

discrimination of words within and across languages. To this end we used a supervised machine learning algorithm (linear support vector machines, linear-SVM; Cortes and Vapnik, 1995) as implemented by the Bioinformatics Matlab toolbox (maximum number of learning iterations = 15,000). Classifications were performed to evaluate whether patterns of EEG data pertained relevant information encoding the representations of spoken words (within-language discrimination) as well as their language invariant semantic-conceptual representations (across-language generalization). All classifications were binary (i.e., chance-level is 0.5) and involved discrimination and generalization between two words. The results of these binary predictions were then averaged across all possible pair-wise classifications. Additional methodological steps encompassing the computational strategy to validate the classification results (cross-validation) and to select the EEG features used for classification (feature selection) are described below.

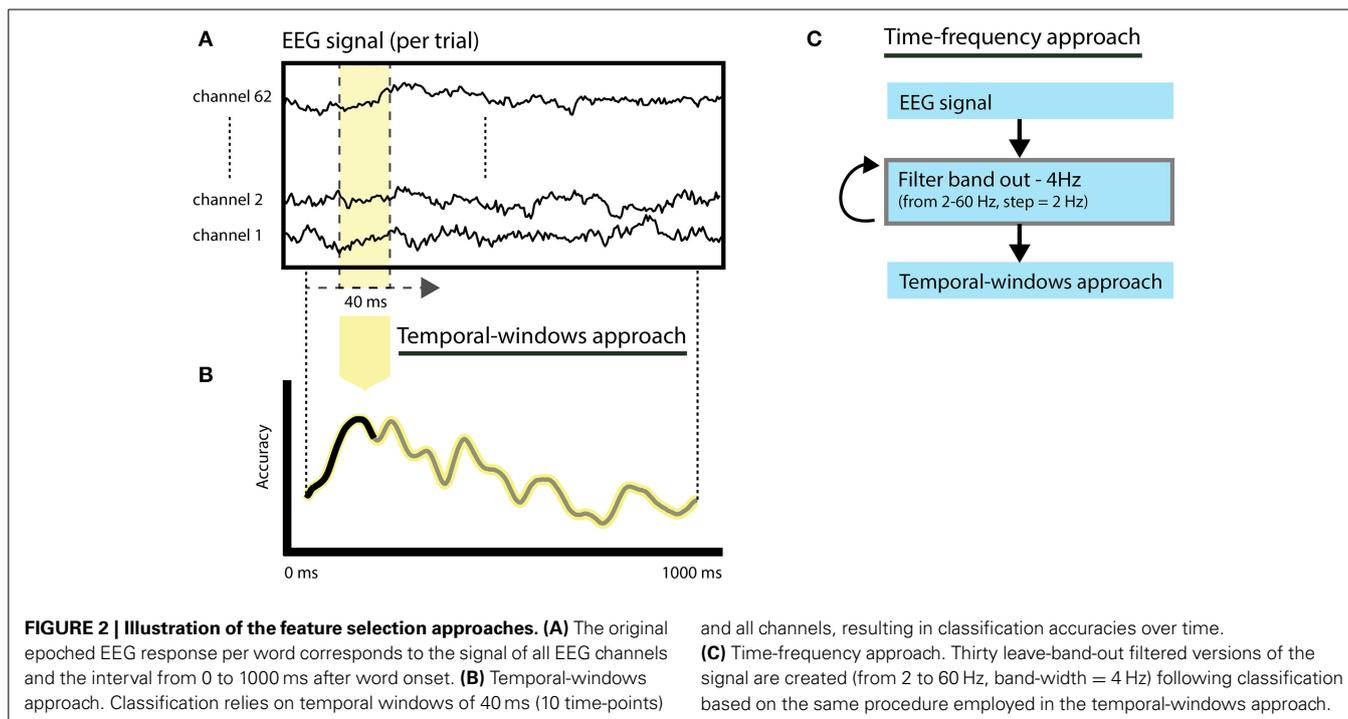
### CROSS-VALIDATION APPROACHES

Cross-validation of the multivariate classification analysis served two purposes: (1) to obtain robust estimates of the discrimination accuracies; (2) to allow generalization of classes by using distinct class groupings during the training and testing phases of classification. Cross-validation for within-language word discrimination relied on speaker identity. Here, we trained a classifier to discriminate words based on samples recorded from two out of the three speakers that pronounced the words (32 trials per word) and tested whether this training was able to generalize the left-out speaker pronouncing the same words (16 trials per word). This cross-validation procedure assured word discrimination invariant to neural activations specific to acoustic-phonetic characteristics of the speakers. Cross-validation for across-language generalization of semantic concepts relied on language independent information of the words. Here, we trained a classifier to discriminate words within one language (48 trials per word) and tested whether this training generalized to the other language (48 trials per word). Hence, in across-language generalization, we aimed to isolate semantic conceptual properties of the words that were language invariant.

### FEATURE SELECTION APPROACHES

#### **Temporal-windows approach (shifting-windows + all channels)**

To investigate the temporal evolution of spoken word decoding, we selected EEG response features (**Figure 2A**) using shifting-windows (width = 40 ms—10 time points) across all channels (**Figure 2B**). Restricting the EEG signal features to specific time windows permits the calculation of changes in classification accuracies over time informative of spoken word processing. Because the temporal-windows approach reduces the number of EEG features used for classification, it increases the temporal sensitivity of the classifiers to speaker and language invariant information of the spoken words due to a potentially better match between the training and testing patterns (Hausfeld et al., 2012). Additionally, it reduces the high dimensionality of the feature space, thus avoiding degraded classification performances (model overfitting; for a description, see Norman et al., 2006). The empirical null distribution was computed per subject using 200 label permutations.



Individual statistical significance ( $p < 0.05$ ) was calculated based on deviance from permuted accuracies. Group level statistics were calculated based on the overlap of significant subjects across time intervals using a binomial test with  $n = 16$  (number of subjects) and  $p = 0.05$  (Darlington and Hayes, 2000; Hausfeld et al., 2012) and corrected for multiple comparisons (time windows) using FDR correction ( $\alpha = 5\%$ ).

#### **Time-frequency approach (filtered-band-out + shifting-windows + all channels)**

To assess the importance of brain oscillations in specific frequency bands to the performance of the classifiers we employed a feature selection approach combining temporal shifting windows and filter-band-out (Figure 2C). The original epoched EEG responses (−1000 to 1000 ms) were filtered prior to classification using an FIR (finite impulse response) filter as implemented in EEGLab (Delorme and Makeig, 2004). The width of the filtered-out frequency band was set to 4 Hz, centered on frequencies from 2 up to 60 Hz and in frequency steps of 2 Hz, producing 30 filtered signals. For each of the filtered signal versions, we subsequently performed the *temporal-windows* approach to assess the importance of each frequency band over time. The importance of the left-out frequency band was quantified in terms of a change in classification performance with respect to the non-filtered signal. To prevent a modulation of time-frequency importance due to differences in the original classification accuracy, a normalization of the importance of each time-frequency bin with respect to the accuracy limits (0–1) was performed using “odds-ratio” normalization (Szumilas, 2010). Odds-ratio values above 1 indicate a reduction of classification accuracy after a specific frequency band is filtered out. This approach allowed us to investigate the contribution of each frequency band over time

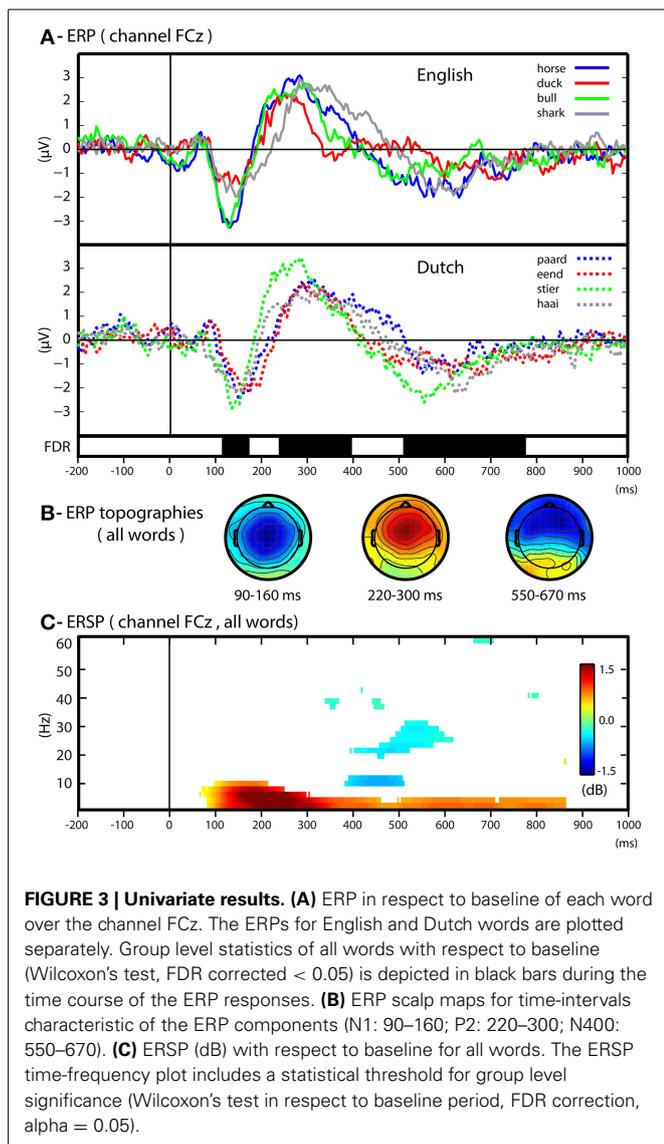
without disrupting EEG spectral interactions that may be crucial in many cognitive processes, including speech processing (Giraud and Poeppel, 2012; Henry and Obleser, 2012; Peelle and Davis, 2012). Group statistics were performed in random-effects (two-sided Wilcoxon’s test) and corrected for multiple comparisons using FDR correction ( $\alpha = 5\%$ ).

## **RESULTS**

### **ERPs AND TIME-FREQUENCY ANALYSIS**

We first conducted univariate analyses of ERP and time-frequency changes relatively to stimulus baseline in order to assess the overall spectro-temporal characteristics of EEG responses evoked by the animal words. Figure 3 illustrates the averaged ERP responses elicited by the different animal words, including the expected ERP peaks (channel Fcz, Figure 3A) and their corresponding topographies (Figure 3B), in the N1 window (120–160 ms), the P2 window (230–390 ms) and the N400 window (550–800 ms). To assess univariate differences between the ERP responses we conducted all possible word-to-word contrasts within the same language (e.g., horse vs. duck), as well as all possible concept-to-concept contrasts (e.g., horse + paard vs. duck + eend). None of the possible contrasts yielded significant differences within or across participants.

The analysis of averaged power changes in different frequency bands (Figure 3C) shows an average power increase (ERS, event-related synchronization) of slow oscillations (1–10 Hz) starting 100 ms after stimulus onset, followed by a steep reduction in alpha power (ERD, event-related desynchronization) between 400 and 500 ms. At later time intervals, the ERS of slow oscillations (1–8 Hz) was maintained. These differences did not allow the systematic discrimination of individual words nor of language-independent concepts.



### MULTIVARIATE ANALYSIS (MVPA)

The multivariate analysis consisted of assessing the ability of multivariate classifiers to discriminate words within the same language and across first and second language in bilingual subjects. To assess the contribution of specific EEG features used for classification we used two feature selection approaches: a *temporal-windows approach*, relying on restricted time intervals (40 ms) shifted over time and all EEG channels; and a *time-frequency approach*, relying on a combined selection of features using the *temporal-windows approach* and a moving filter-band-out procedure (4 Hz bands with an step of 2 Hz).

The *temporal-windows* feature selection approach enabled identifying specific time-intervals related to word decoding. Within-language discrimination (**Figure 4A**) was significantly possible throughout most of the time-course from ~50 until 620 ms after word onset. Within this broad time window, salient local maxima of accuracies were identified for the temporal windows (40 ms) around 160 ms (accuracy = 0.535), 225 ms

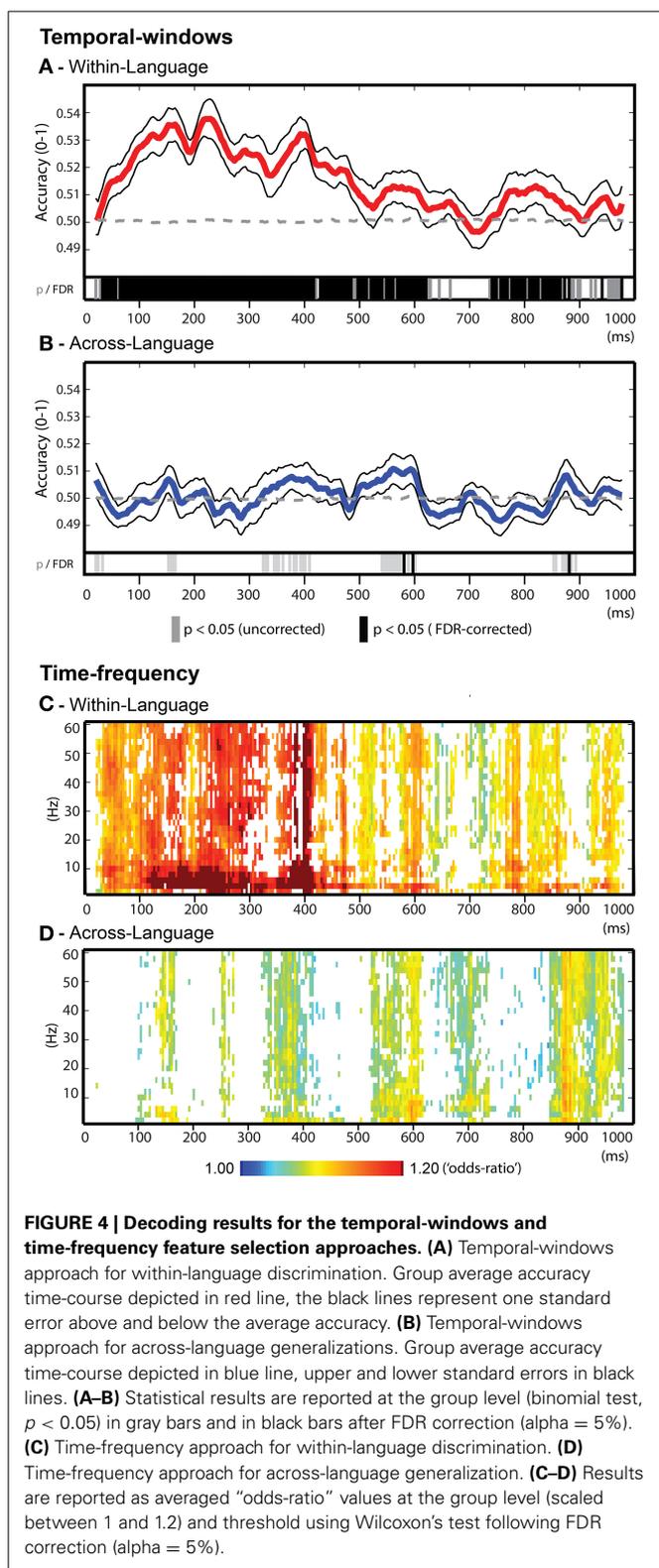
(accuracy = 0.537), 390 ms (accuracy = 0.533), 570 ms (accuracy = 0.513), and 820 ms (accuracy = 0.512). Interestingly, cross-language generalization (**Figure 4B**) led to significant classification in more restricted temporal windows with significant results between 550 and 600 ms (maximum accuracy = 0.511) and 850–900 ms (maximum accuracy = 0.508). A further time-interval showing a trend (uncorrected  $p < 0.05$ ) for across-language generalization capacity was observed around 400 ms (maximum accuracy = 0.507).

The *time-frequency* feature selection approach assessed the contribution of oscillatory activity in specific frequency bands to word decoding across the different time windows. For this purpose, “odds-ratio” values were computed, group averaged and thresholded for statistical significance (random-effects, FDR = 5%). Overall, the temporal profiles of the *time-frequency approach* match consistently with that of the *temporal-windows approach*, confirming that reductions in classification accuracy due to the omission of specific frequency bands occurred in time windows relevant for word decoding (**Figure 4C**). For within-language discrimination of words, reductions in classification accuracy especially occurred when omitting slow oscillations (below 12 Hz, delta, theta and alpha). For across-language generalization (**Figure 4D**), the period around 600 ms that showed significant generalization capacity, was characterized by accuracy reductions when filtering out frequencies up to 10 Hz (delta-theta-alpha). In other time windows a contribution of slow oscillations was also observed for this analysis, although involving slower oscillations (delta/ low theta, below 6 Hz). Visual inspection of **Figures 4C–D** further suggested that besides the sensitivities for oscillations below 12 Hz, for both types of analysis smaller classification drops occurred across gamma band (above 30 Hz) as well as across broad-band oscillation profiles.

### DISCUSSION

By combining EEG MVPA and an experimental design that exploits the unique capacities of bilingual listeners we identified specific time windows and oscillations enabling within-category discrimination of individual spoken words. We demonstrated within-language word decoding in a broad time-window from ~50 to 620 ms after word onset with a strong contribution of slow oscillations (below 12 Hz). Most importantly, we were able to isolate specific time windows, including the 550–600 ms window, in which EEG features enabled the generalization of the meaning of the words across their Dutch and English word forms. Our results demonstrate the feasibility of using MVPA to identify individual word representations based on speech evoked EEG signals. Furthermore, they indicate the advantage of feature selection approaches in assessing temporal and temporal-oscillatory EEG response features in classification.

The univariate analyses illustrate ERP and oscillatory responses typically elicited by individual spoken words (Kutas and Federmeier, 2000; Hagoort et al., 2004; Bastiaansen et al., 2008; Bonte et al., 2009; Strauß et al., 2014) indicating a progression from acoustic-phonetic to lexical-semantic processing. The ERPs to the individual words show variability as a consequence of acoustic-phonetic differences and other word-specific properties. However, these differences did not



allow the systematic discrimination of individual words nor of language-independent concepts. The prevalence of slow oscillatory activity (below 12 Hz) while subjects listened to the words indicates the crucial role of these frequencies in the

processing and comprehension of speech (Hagoort et al., 2004; Giraud and Poeppel, 2012; Strauß et al., 2014). The analysis also showed that the univariate frequency power changes were not suitable for distinguishing individual words or across-language generalization of semantic concepts.

Importantly, the multivariate analyses allowed finding neural time-course correlates of the individual words that were invariant to the acoustic-phonetic characteristics of the speakers (within-language discrimination) as well as to the language in which the meaning was presented (across-language generalization). Within-language word discrimination relied on acoustic-phonetic and semantic-conceptual differences between the nouns, but also on possible other differences reflecting their individual properties. Accordingly, within-language discrimination was possible for both approaches of feature selections employed. In the *temporal-windows* approach (**Figure 4A**), investigating the temporal evolution of classification across consecutive short time-intervals of 40 ms, classification performance was significant from  $\sim 50$  until 620 ms after word onset. In accordance with the ERP literature, decoding in this broad time window may reflect a progression from phonetic-phonological processing (100–200 ms; Dumay et al., 2001; Sanders and Neville, 2003; Bonte and Blomert, 2004; Uusvuori et al., 2008) to initial lexical access (200–300 ms; Van den Brink et al., 2001; Hagoort et al., 2004; Salmelin, 2007; Bonte et al., 2009), and lexical semantic processing (300–600 ms; Kutas and Hillyard, 1980; Hagoort, 2008). These results are also consistent with previous single trial auditory word classification (Simanova et al., 2010) that showed initial prominent classification capability centered around 240 ms followed by a second less prominent capability around 480 ms after word onset.

The second multivariate analysis - across-language generalization - relied uniquely on language invariant semantic-conceptual properties of the nouns. This analysis, and especially the temporal-window approach (**Figure 4B**), revealed language invariant EEG features coding for the animal words in much more restricted time-windows including the 550–600 ms window and the 850–900 ms window at the end of the EEG epoch. ERP research has commonly associated similar time intervals with lexical-semantic processing of words across different task and sentence contexts (Kutas and Federmeier, 2000; Hagoort, 2008). Here, we indicate the potential of EEG signals to represent semantic-conceptual information of individual words independent of their acoustic-phonetic implementation or word-form. In order to isolate these input-invariant lexical-semantic representations we used animal nouns that were acoustically-phonetically distinct both within and across languages and were presented together with non-animal nouns that served as targets. In everyday speech processing, it is more difficult to disentangle input-driven vs. input-independent processes as initial lexical-semantic access is influenced by both acoustic-phonetic word form information (McClelland and Elman, 1986; Marslen-Wilson, 1987) and semantic or task context (Bonte, 2004; Obleser et al., 2004; Çukur et al., 2013), leading to early lexical and/or semantic ERP modulations around 200–300 ms (e.g., Van den Brink et al., 2001; Bonte et al., 2006; Travis et al., 2013; Strauß et al., 2014). Our approach presents a way to disentangle these aspects of comprehension. Importantly, by using words belonging to the

same semantic category—animals—we reduced the influence of larger scale semantic category differences that can also drive the decoding of individual nouns (Simanova et al., 2010; Chan et al., 2011b; Shinkareva et al., 2011).

In later time-windows, significant classification for within-language discrimination (750–900 ms) and across-language generalization (850–900 ms) may reflect effects specific to our paradigm. That is, the slow presentation of words and/or the use of a target detection task, may have led to e.g., subvocal rehearsal in working memory (Kutas and Federmeier, 2000; Baddeley, 2003; Buchsbaum et al., 2011) and/or response monitoring toward the end of the trial (Wang et al., 2012).

In bilinguals, the active translation of written words during speech production tasks has been shown to elicit ERP differences for translation direction around 400 ms after word presentation (Christoffels et al., 2013). In the current study the effect of direct translations was minimized in several ways. First, we avoided active translations from second to native language and vice-versa by separately presenting words in Dutch and English blocks and using catch trials consisting of Dutch and English non-animal words, respectively. Furthermore, we used a selection of words with relatively early age of acquisition and of medium-high frequency of use in both languages.

To further understand the EEG temporal patterns allowing classification, we employed a time-frequency feature selection approach that assessed the relative contribution of oscillatory bands. We observed a significant contribution of slow EEG oscillations (below 12 Hz) for within-language and across-language classification, which links to the synchronization of oscillatory bands observed in the ERSP analysis. Furthermore, in the time windows during which the slower oscillations most strongly influenced classification performance, results also indicated a contribution from higher, gamma band oscillations (above 30 Hz). It would be interesting to replicate this possible co-occurrence of slower and gamma band modulations in future studies with bilinguals, and, in particular to test how they relate to suggested processing of (phonemes, syllables and semantic information (Lakatos et al., 2005; Giraud and Poeppel, 2012; Peelle and Davis, 2012; Peña and Melloni, 2012).

We may hypothesize that the neural processing underlying the EEG-based translations of animal nouns occurs in a brain network that was recently identified in an fMRI study using a comparable bilingual paradigm (Correia et al., 2014). In particular, in this previous study, language-invariant classification of animal words was found to rely on focal brain regions, including the left anterior temporal lobe (left-ATL), corroborating the existence of “hub” regions organizing semantic-conceptual knowledge in abstract form. Correspondingly, recent models of conceptual knowledge (Patterson et al., 2007), brain lesion studies (Damasio et al., 1996) and neuroimaging evidence (Visser et al., 2012; Correia et al., 2014) locate a possible semantic hub within the left-ATL, integrating distributed semantic-conceptual information throughout the cortex. Furthermore, distributed neural representations of semantic information may also connect to modality specific brain regions subserving perception and action (Martin, 2007; Meyer and Damasio, 2009). Interestingly, magnetoencephalography (MEG) studies have related time windows

starting at 400 ms after spoken word onset to semantic processing in bilateral anterior temporal areas (Marinkovic et al., 2003; Chan et al., 2011a), suggesting a putative link between the present finding of language-independent word decoding in the 550–600 ms time window and processing in these brain regions. At present, this spatial-temporal association remains speculative, but similar classification paradigms using simultaneous fMRI and EEG recordings (De Martino et al., 2011) may allow investigating the joint spatio-temporal representation of spoken words. Furthermore, earlier indications of semantic/conceptual representations of our words are observed in a spread time window between 320 and 420 ms after word onset (uncorrected  $p < 0.05$ ). These and possibly even earlier semantic activations elicited by the individual animal words may be more difficult to detect due to variability in the exact timing of these initial activations.

Overall, our results show the benefit of EEG-based MPVA to investigate the representation of semantic concepts independently of the input language and more generally of individual spoken words independently of the speaker. Although the obtained accuracies are relatively low, they demonstrate the sensitivity of multivariate classification to distinguish subtle representations extracted from single-trial EEG responses that may not be present in the averaged EEG signal across multiple trials (Makeig et al., 2002; Hausfeld et al., 2012). Furthermore, our results show the potential of feature selection approaches based on moving temporal windows to highlight time windows associated with the neural processing of specific characteristics of speech and language (e.g., language independent semantic processing, see also Simanova et al., 2010; Chan et al., 2011b; Hausfeld et al., 2012). Future studies including different sets of words, languages or feature selection approaches may help confirming the generalization of our results. Beyond decoding language-invariant semantic-concepts during listening, EEG-based MVPA may also be used to investigate whether semantic-concepts share a similar neural representation during reading and speaking (Hickok et al., 2011; Pickering and Garrod, 2013). When we speak, we start from ideas and concepts and convert these into articulatory motor programs. ERP studies on speech production (e.g., picture naming), relate early windows, 100–200 ms after stimulus onset to interactive processing of visual encoding and accessing concepts for language use (Rahman and Sommer, 2003; Redmann et al., 2014). Like in speech comprehension, this interaction between input-dependent and abstract semantic-conceptual representations in speech production, together with their strong context and task-dependency (e.g., Jescheniak et al., 2002; Aristei et al., 2011), makes it difficult to isolate abstract semantic conceptual representations using univariate analysis methods. Because our EEG-based MVPA approach may disentangle these processes, it would thus be of interest to employ this same approach in speech production studies (e.g., and Schmitt et al., 2000; Koester and Schiller, 2008). In particular, a similar bilingual paradigm involving word naming in bilingual speakers would allow investigating the timing of language-independent semantic-conceptual representations. Furthermore, the classification of spoken words across and within languages in bilingual speakers and across and within speech modality (perception and production) may allow to investigate neural representations crucial for the initiation of

speech production (Levelt, 1989; Rahman and Sommer, 2003; Indefrey and Levelt, 2004; Indefrey, 2011), as well as, for the monitoring of speech output (Hickok et al., 2011).

## REFERENCES

- Aristei, S., Melinger, A., and Rahman, A. R. (2011). Electrophysiological chronometry of semantic context effects in language production. *J. Cogn. Neurosci.* 23, 1567–1586. doi: 10.1162/jocn.2010.21474
- Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Baddeley, A. (2003). Working memory and language: an overview. *J. Commun. Disord.* 36, 189–208. doi: 10.1016/S0021-9924(03)00019-4
- Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., and Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain Lang.* 106, 15–28. doi: 10.1016/j.bandl.2007.10.006
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528. doi: 10.1093/cercor/10.5.512
- Boersma, P., and Weenink, D. (2013). Praat: doing phonetics by computer. *Glott Int.* 5, 341–345. Available online at: <http://www.praat.org/>
- Bonte, M. (2004). Developmental changes in ERP correlates of spoken word recognition during early school years: a phonological priming study. *Clin. Neurophysiol.* 115, 409–423. doi: 10.1016/S1388-2457(03)00361-4
- Bonte, M., and Blomert, L. (2004). Developmental changes in ERP correlates of spoken word recognition during early school years: a phonological priming study. *Clin. Neurophysiol.* 115, 409–423. doi: 10.1016/S1388-2457(03)00361-4
- Bonte, M., Hausfeld, L., Scharke, W., Valente, G., and Formisano, E. (2014). Task-dependent decoding of speaker and vowel identity from auditory cortical response patterns. *J. Neurosci.* 34, 4548–4557. doi: 10.1523/JNEUROSCI.4339-13.2014
- Bonte, M., Parviainen, T., Hytönen, K., and Salmelin, R. (2006). Time course of top-down and bottom-up influences on syllable processing in the auditory cortex. *Cereb. Cortex* 16, 115–123. doi: 10.1093/cercor/bhi091
- Bonte, M., Valente, G., and Formisano, E. (2009). Dynamic and task-dependent encoding of speech and voice by phase reorganization of cortical oscillations. *J. Neurosci.* 29, 1699–1706. doi: 10.1523/JNEUROSCI.3694-08.2009
- Brandmeyer, A., Farquhar, J. D. R., McQueen, J. M., and Desain, P. W. M. (2013). Decoding speech perception by native and non-native speakers using single-trial electrophysiological data. *PLoS ONE* 8:e68261. doi: 10.1371/journal.pone.0068261
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., et al. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - an aggregate analysis of lesion and fMRI data. *Brain Lang.* 119, 119–128. doi: 10.1016/j.bandl.2010.12.001
- Chan, A. M., Baker, J. M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., et al. (2011a). First-pass selectivity for semantic categories in human anteroventral temporal lobe. *J. Neurosci.* 31, 18119–18129. doi: 10.1523/JNEUROSCI.3122-11.2011
- Chan, A. M., Halgren, E., Marinkovic, K., and Cash, S. S. (2011b). Decoding word and category-specific spatiotemporal representations from MEG and EEG. *Neuroimage* 54, 3028–3039. doi: 10.1016/j.neuroimage.2010.10.073
- Christoffels, I. K., Ganushchak, L., and Koester, D. (2013). Language conflict in translation: an ERP study of translation production. *J. Cogn. Psychol.* 25, 646–664. doi: 10.1080/20445911.2013.821127
- Correia, J., Formisano, E., Valente, G., Hausfeld, L., Jansma, B., and Bonte, M. (2014). Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe. *J. Neurosci.* 34, 332–338. doi: 10.1523/JNEUROSCI.1302-13.2014
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018
- Çukur, T., Nishimoto, S., Huth, A. G., and Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nat. Neurosci.* 16, 763–770. doi: 10.1038/nn.3381
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., and Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature* 380, 499–505. doi: 10.1038/380499a0
- Darlington, R. B., and Hayes, A. F. (2000). Combining independent p values: extensions of the Stouffer and binomial methods. *Psychol. Methods* 5, 496–515. doi: 10.1037/1082-989X.5.4.496
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- De Martino, F., de Borst, A. W., Valente, G., Goebel, R., and Formisano, E. (2011). Predicting EEG single trial responses with simultaneous fMRI and relevance vector machine regression. *Neuroimage* 56, 826–836. doi: 10.1016/j.neuroimage.2010.07.068
- De Moor, W., Ghyselinck, M., and Brysbaert, M. (2000). A validation study of the age-of-acquisition norms collected by Ghyselinck, De Moor, and Brysbaert. *Psychol. Belg.* 40, 99–114.
- Doelling, K. B., Arnal, L. H., Ghitza, O., and Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage* 85 Pt 2, 761–768. doi: 10.1016/j.neuroimage.2013.06.035
- Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., and Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *J. Cogn. Neurosci.* 13, 121–143. doi: 10.1162/089892901564117
- Formisano, E., De Martino, F., Bonte, M., and Goebel, R. (2008b). “Who” is saying “what”? Brain-based decoding of human voice and speech. *Science* 322, 970–973. doi: 10.1126/science.1164318
- Formisano, E., De Martino, F., and Valente, G. (2008a). Multivariate analysis of fMRI time series: classification and regression of brain responses using machine learning. *Magn. Reson. Imaging* 26, 921–934. doi: 10.1016/j.mri.2008.01.052
- Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1055–1069. doi: 10.1098/rstb.2007.2159
- Hagoort, P., Hald, L., Bastiaansen, M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441. doi: 10.1126/science.1095455
- Hausfeld, L., De Martino, F., Bonte, M., and Formisano, E. (2012). Pattern analysis of EEG responses to speech and voice: influence of feature grouping. *Neuroimage* 59, 3641–3651. doi: 10.1016/j.neuroimage.2011.11.056
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., et al. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* 72, 404–416. doi: 10.1016/j.neuron.2011.08.026
- Henry, M. J., and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20095–20100. doi: 10.1073/pnas.1213390109
- Herrmann, B., Maess, B., Kalberlah, C., Haynes, J.-D., and Friederici, A. D. (2012). Auditory perception and syntactic cognition: brain activity-based decoding within and across subjects. *Eur. J. Neurosci.* 35, 1488–1496. doi: 10.1111/j.1460-9568.2012.08053.x
- Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422. doi: 10.1016/j.neuron.2011.01.019
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113
- Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Front. Psychol.* 2:255. doi: 10.3389/fpsyg.2011.00255
- Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001
- Jescheniak, J. D., Schriefers, H., Garrett, M. F., and Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *J. Cogn. Neurosci.* 14, 951–964. doi: 10.1162/089892902760191162
- Kilian-Hütten, N., Valente, G., Vroomen, J., and Formisano, E. (2011). Auditory cortex encodes the perceptual interpretation of ambiguous sound. *J. Neurosci.* 31, 1715–1720. doi: 10.1523/JNEUROSCI.4572-10.2011
- Koester, D., and Schiller, N. O. (2008). Morphological priming in overt language production: electrophysiological evidence from Dutch. *Neuroimage* 42, 1622–1630. doi: 10.1016/j.neuroimage.2008.06.043

- Kutas, M., and Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6
- Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911. doi: 10.1152/jn.00263.2005
- Lee, T. W., Girolami, M., and Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and supergaussian sources. *Neural Comput.* 11, 417–441. doi: 10.1162/089976699300016719
- Lemhöfer, K., and Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behav. Res. Methods* 44, 325–343. doi: 10.3758/s13428-011-0146-0
- Levelt, W. J. M. (1989). *Speaking: from Intention to Articulation*. Cambridge, MA: MIT Press.
- Levenshtein, V. I. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Dokl. Akad. Nauk SSSR* 163, 845–848.
- Luo, H., and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54, 1001–1010. doi: 10.1016/j.neuron.2007.06.004
- Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., et al. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690–694. doi: 10.1126/science.1066168
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., and Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38, 487–497. doi: 10.1016/S0896-6273(03)00197-1
- Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition* 25, 71–102. doi: 10.1016/0010-0277(87)90005-9
- Martin, A. (2007). The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45. doi: 10.1146/annurev.psych.57.102904.190143
- McClelland, J. L., and Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0
- Meyer, K., and Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends Neurosci.* 32, 376–382. doi: 10.1016/j.tins.2009.04.002
- Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430. doi: 10.1016/j.tics.2006.07.005
- Obleser, J., Elbert, T., and Eulitz, C. (2004). Attentional influences on functional mapping of speech sounds in human auditory cortex. *BMC Neurosci.* 5:24. doi: 10.1186/1471-2202-5-24
- Obleser, J., Herrmann, B., and Henry, M. J. (2012). Neural oscillations in speech: don't be enslaved by the envelope. *Front. Hum. Neurosci.* 6:250. doi: 10.3389/fnhum.2012.00250
- Patterson, K., Nestor, P. J., and Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987. doi: 10.1038/nrn2277
- Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol.* 3:320. doi: 10.3389/fpsyg.2012.00320
- Peña, M., and Melloni, L. (2012). Brain oscillations during spoken sentence processing. *J. Cogn. Neurosci.* 24, 1149–1164. doi: 10.1162/jocn\_a\_00144
- Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–347. doi: 10.1017/S0140525X12001495
- Rahman, R. A., and Sommer, W. (2003). Does phonological encoding in speech production always follow the retrieval of semantic knowledge? Electrophysiological evidence for parallel processing. *Brain Res. Cogn. Brain Res.* 16, 372–382. doi: 10.1016/S0926-6410(02)00305-1
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331
- Redman, A., Fitzpatrick, I., Hellwig, F., and Indefrey, P. (2014). The use of conceptual components in language production: an ERP study. *Front. Psychol.* 5:363. doi: 10.3389/fpsyg.2014.00363
- Salmelin, R. (2007). Clinical neurophysiology of language: the MEG approach. *Clin. Neurophysiol.* 118, 237–254. doi: 10.1016/j.clinph.2006.07.316
- Sanders, L. D., and Neville, H. J. (2003). An ERP study of continuous speech processing. I. Segmentation, semantics, and syntax in native speakers. *Brain Res. Cogn. Brain Res.* 15, 228–240. doi: 10.1016/S0926-6410(02)00195-7
- Schmitt, B. M., Münte, T. F., and Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology* 37, 473–484. doi: 10.1111/1469-8986.3740473
- Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123 Pt 12, 2400–2406. doi: 10.1093/brain/123.12.2400
- Shahin, A. J., Picton, T. W., and Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain Cogn.* 70, 259–266. doi: 10.1016/j.bandc.2009.02.008
- Shinkareva, S. V., Malave, V. L., Mason, R. A., Mitchell, T. M., and Just, M. A. (2011). Commonality of neural representations of words and pictures. *Neuroimage* 54, 2418–2425. doi: 10.1016/j.neuroimage.2010.10.042
- Simanova, I., van Gerven, M., Oostenveld, R., and Hagoort, P. (2010). Identifying object categories from event-related EEG: toward decoding of conceptual representations. *PLoS ONE* 5:e14465. doi: 10.1371/journal.pone.0014465
- Strauß, A., Kotz, S. A., Scharinger, M., and Obleser, J. (2014). Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *Neuroimage* 97C, 387–395. doi: 10.1016/j.neuroimage.2014.04.005
- Sudre, G., Pomerleau, D., Palatucci, M., Wehbe, L., Fyshe, A., Salmelin, R., et al. (2012). Tracking neural coding of perceptual and semantic features of concrete nouns. *Neuroimage* 62, 451–463. doi: 10.1016/j.neuroimage.2012.04.048
- Szumilas, M. (2010). Explaining odds ratios. *J. Can. Acad. Child Adolesc. Psychiatry* 19, 227–229.
- Travis, K. E., Leonard, M. K., Chan, A. M., Torres, C., Sizemore, M. L., Qu, Z., et al. (2013). Independence of early speech processing from word meaning. *Cereb. Cortex* 23, 2370–2379. doi: 10.1093/cercor/bhs228
- Uusvuori, J., Parviainen, T., Inkinen, M., and Salmelin, R. (2008). Spatiotemporal interaction between sound form and meaning during spoken word perception. *Cereb. Cortex* 18, 456–466. doi: 10.1093/cercor/bhm076
- Van den Brink, D., Brown, C. M., and Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *J. Cogn. Neurosci.* 13, 967–985. doi: 10.1162/089892901753165872
- Van Petten, C., Coulson, S., Rubin, S., Plante, E., and Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 394–417. doi: 10.1037/0278-7393.25.2.394
- Visser, M., Jefferies, E., Embleton, K. V., and Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *J. Cogn. Neurosci.* 24, 1766–1778. doi: 10.1162/jocn\_a\_00244
- Visser, M., and Lambon Ralph, M. A. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *J. Cogn. Neurosci.* 23, 3121–3131. doi: 10.1162/jocn\_a\_00007
- Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J.-M., Magyari, L., et al. (2012). Beta oscillations relate to the N400m during language comprehension. *Hum. Brain Mapp.* 33, 2898–2912. doi: 10.1002/hbm.21410

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 June 2014; accepted: 13 January 2015; published online: 06 February 2015.

Citation: Correia JM, Jansma B, Hausfeld L, Kikkert S and Bonte M (2015) EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations. *Front. Psychol.* 6:71. doi: 10.3389/fpsyg.2015.00071

This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2015 Correia, Jansma, Hausfeld, Kikkert and Bonte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The role of the sound of objects in object identification: evidence from picture naming

Claudio Mulatti<sup>1\*</sup>, Barbara Treccani<sup>2</sup> and Remo Job<sup>3</sup>

<sup>1</sup> Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università degli Studi di Padova, Padova, Italy

<sup>2</sup> Dipartimento di Storia, Scienze dell'Uomo e della Formazione, Università degli Studi di Sassari, Sassari, Italy

<sup>3</sup> Dipartimento di Psicologia e Scienze Cognitive, Università degli Studi di Trento, Rovereto, Italy

## Edited by:

Peter Indefrey, University of  
Dusseldorf, Germany

## Reviewed by:

Antje Meyer, Max Planck Institute  
for Psycholinguistics, Netherlands  
Pia Knoeferle, Bielefeld University,  
Germany

## \*Correspondence:

Claudio Mulatti, Dipartimento di  
Psicologia dello Sviluppo e della  
Socializzazione, Università degli  
Studi di Padova, Via Venezia, 8,  
35131 Padova, Italia  
e-mail: claudio.mulatti@unipd.it

In the present work we were concerned with the role of sound representations in object recognition. In order to address this issue we made use of a picture naming task in which target pictures might be accompanied by a white-noise burst. White-noise was thought to interfere with the representation of the sound possibly associated with the depicted object. We reasoned that if such a representation is critical for the recognition of objects strongly associated with certain sounds, white-noise interference should affect the naming of pictures representing objects with typical sounds leaving the naming of object without typical sounds unaffected. The results were congruent with the predictions and consistent with a view of the semantic representations of objects as collection of related representations, modal in nature, and mandatorily accessed.

**Keywords:** picture naming, object recognition, grounded cognition, embodied cognition, language, semantics, object processing, sound representation

## INTRODUCTION

This study deals with the role of sounds in object recognition in humans. Indeed, some objects are easily associated with a sound, i.e., some objects possess either a *typical sound* or category of sounds. This is the case, for example, of objects such as “bell” or “motorbike.” Other objects do not possess typical sounds or can be associated with particular sounds only with difficulty. This is the case, for example, of objects such as “table” or “pillow.”

Given that objects can be classified as a function of whether they possess or not a typical sound, a legitimate question is whether the typical sounds play any role in the visual recognition of the related objects. There are at least two opposed scenarios to frame this question.

In the first scenario, upon the presentation of a visual object the system first accesses an abstract representation of that object and then—depending on the task at hand—accesses the representations of information related to that object: among these representations is the representation of the typical sound. Thus, in this scenario, the access to the typical sound is post-categorical, in the sense that the object is first recognized as an instance of a particular kind (e.g., a “dog”) and then the related information is retrieved (cf. Allport, 1977; Mulatti et al., 2014). Here, the typical sound may be activated but, since its retrieval follows the identification of the object, it does not play any role in the recognition of the object.

In the second scenario, all stored representations associated to a given object are immediately and mandatorily activated upon the visual presentation of an instance of that kind of object. Here, the identification of the object does not consist in the activation of an abstract semantic representation of this object but instead corresponds to the activation of all stored representations. In other words, object identification *is* the activation of object knowledge.

For objects with a typical sound, the typical sound is part of the knowledge of that object and, therefore, the activation of the typical sound is part of the process of object identification: an object cannot be identified without its typical sound being activated. Thus, in this second scenario the access to the typical sound is pre-categorical and has a functional significance in the identification process: typical-sound activation does not only occur when it is requested by the task and it is not simply a concomitant, epiphenomenal, effect of the identification (cf. Kiefer and Barsalou, 2013).

These two scenarios can be seen as the two extreme positions of a continuum of scenarios going from post- to pre-categorical, and therefore intermediate positions are possible (Pezzulo, 2011). In this study we attempt to provide evidence in favor of one of these two extremes.

Previous studies investigating cross-modal effects in object recognition have shown that when both visual and auditory information (e.g., the picture of an object and the typical sound of an object) are presented in object recognition tasks, both types of information affect the time need to emit a response: responses are usually faster when participants are presented with cross-modal congruent stimuli (i.e., the sound refers to the object depicted in the picture) than when they are presented with incongruent stimuli (i.e., the sound is typical of another object; e.g., Laurienti et al., 2003). Based on psychophysiological and neuroimaging findings, visual and auditory inputs are thought to interact quite early (i.e., at sensory processing stages; e.g., Giard and Peronnet, 1999). Yet, according to the most accepted view, they would be integrated afterwards (e.g., Hocking and Price, 2008), at higher cognitive processing stages. Sensory information from unimodal processing channels would converge onto a modality-independent semantic system (Coltheart, 1987). Cross-modal

semantic congruency effects would arise at this processing level and, consistent with this view, they are typically interpreted within a post-categorical framework (cf., Schneider et al., 2008). Congruent visual and auditory inputs are seen as independent perceptual cues activating the same (amodal) semantic knowledge. The addition of a redundant congruent perceptual cue (e.g., the typical sound of an object when participants has to recognize a picture) can facilitate the recognition of the object by enhancing its activation level (then reducing competition) and is particularly useful when the object has many structurally and semantically similar neighbors that compete for selection (Humphreys et al., 1995). In this respect, a congruent sound does not have any facilitatory role in the recognition of an object when recognition can proceed on the basis of visual stimuli alone (e.g., Hocking and Price, 2008).

However, results of cross-modal integration studies might be equally easily interpreted by a pre-categorical account assigning to sounds a functional role in visual object recognition. Indeed, results obtained in tasks providing for the presentation of both visual and auditory stimuli related to a given object cannot help to discriminate between the two accounts: results of these studies tell us nothing about whether the typical sound of an object is activated even when only the visual form of this object is presented, nor whether the sound activation, possibly triggered by the mere presentation of visual stimuli (e.g., Nyberg et al., 2000), is simply a byproduct of object recognition processes or is critical for, and inextricable from, such processes.

The cross-modal semantic congruency paradigm does not then seem a suitable tool for the investigation of the possible functional role of typical sounds in visual object recognition. In the experiment presented below, participants are administered a visual object recognition task in which the activation of the object typical sound is neither required nor triggered by redundant auditory stimuli: we do not present the typical sound of an object or cues that can somehow evoke such a sound, but rather present stimuli that should interfere with the possible (unrequested) activation of the typical sound induced by the recognition process itself.

In this experiment, participants perform a picture naming task. Our choice of the task fell on picture naming because of two aspects that characterize it. First, picture naming requires access to the semantic system (e.g., Potter and Faulconer, 1975; Mulatti et al., 2010). Second, picture naming does not stress the processing of any particular aspect of the meaning in order to be performed, that is it does not require the retrieval of any particular feature of the meaning (Dell'Acqua et al., 2010; Mulatti and Coltheart, 2012): in the present context this means that the naming of a picture of an object possessing a typical sound does not mandatorily require the activation of sound-related representations. So, if an effect due to the typical sound were found in picture naming, we could reasonably conclude that the representation of the typical sound is mandatorily activated in object recognition because of the architecture of the semantic system and not because of the requirements of the task.

In the study, participants name pictures depicting two kinds of objects, objects possessing typical sounds and objects not possessing typical sounds. Here, possessing or not a typical sound is

an operational construct that should not be interpreted literally. An object possesses a typical sound if a sound can be easily associated to that object. An object does not possess a typical sound if no sound can be easily associated to that object.

Each picture is presented twice to each participant, once in each of two conditions. In one condition, the picture is presented along (SOA = 0) with a brief (400 ms) white-noise sound. In the other condition, the picture is presented in isolation, i.e., not accompanied by any sound. White noise should interfere with the retrieval of typical sounds. This is supported by the results of previous studies suggesting the existence of a close link between auditory perception and auditory imagery and memory (e.g., the neural structures active in auditory perception are also active in auditory imagery; see Hubbard, 2010, for a review) and showing that auditory distraction may selectively impair recall of auditory information (e.g., Vredeveldt et al., 2011).

This manipulation then allows us to investigate the possible involvement of typical sound activation in the recognition of the objects depicted in the pictures. If the access to the typical sound is post-categorical, then the concurrent presentation of white noise should not affect the naming of objects with a typical sound more than the naming of objects without a typical sound—and both should not differ from naming the same objects when presented in isolation, i.e., without white noise. This is because picture naming rests on the identification of the object stimulus, and, according to the post-categorical view, the identification of a visual object stimulus precedes—and is independent from—the activation of the representation of the typical sound. So, even if the presence of white-noise affects representation of the sound typically associated with the presented object, this would not affect object naming, regardless of whether the object possesses a typical sound or not.

Instead, if the access to the typical sound is pre-categorical, then the presence of white-noise should interfere more with the naming of objects possessing a typical sounds compared to objects not possessing typical sounds—with respect to the control condition. In the pre-categorical scenario, the activation of the typical sound representation is part of the process of object identification, for those objects that possess a typical sound. Therefore, if the presence of white-noise interferes with the activation of the representation of the typical sound, it also interferes with the identification of the object. Given that object naming rests on object identification, the presence of white-noise should interfere with object naming, but only in the case that the to-be-named object possesses a typical sound.

## EXPERIMENT

### METHODS

#### *Participants*

Thirty-two students of the Università degli Studi di Padova voluntarily participated in the experiment. They were all native Italian speakers with normal or corrected-to-normal vision, and none reported auditory impairments. Oral consent was obtained from each participant before the beginning of the experiment as required by the regulation of the ethical committee of the Università degli Studi di Padova regarding behavioral studies involving adult human participants.

## Design

A 2 Type Of Object (possessing vs. not-possessing typical sound)  $\times$  2 Presentation Condition (picture accompanied with white noise vs. alone) within-subject design was used.

## Material

128 line-drawing (black on white background) pictures of objects (half possessing a typical sound and half not possessing a typical sound) were selected as stimuli. They were taken from the databases of Bates et al. (2003), and of Dell'Acqua et al. (2000). Fourteen participants (not involved in the main study) evaluated how easily each object evokes a typical sound by means of a 7 points Likert-like scale (1 = difficult). In average, objects that were classified as possessing a typical sound received a score of 6.4 (range 5.3–7;  $SD = 0.5$ ) whereas objects that were classified as not possessing a typical sound received a score of 1.7 (range 1–2.6;  $SD = 0.5$ ). Stimuli in the two categories were balanced in terms of frequency of occurrence, name agreement, length, and phonological neighborhood size ( $ts < |1|$ ). The names of the stimuli are reported in the Appendix in Supplementary Material.

A digital hissing sound (44.1 kHz,  $-6$  dBFS) of 400 ms of duration was construed and used as the white-noise stimulus.

## Apparatus and procedure

The experiment took part in a dim-lit sound attenuated room equipped with a PC to which a 17 in. CRT monitor, a voice key, and a pair of speakers were connected. The experiment was controlled by a software developed in E-Prime 2.0. Participants were tested individually and instructed to name the picture as quickly and accurately as possible. Each trial started with the presentation of a fixation point (+) for 500 ms. At its off-set a picture was presented. Reaction times were time-locked to the onset of the picture. Pictures were presented in a single block and, as a function of the experimental condition, they presented either in isolation or accompanied ( $SOA = 0$ ) by the white-noise sound which was delivered by the speakers. The order of presentation of the stimuli for each participants was random. Apparatus and naming errors were scored manually by the experimenter. Before the picture naming experiment, participants were familiarized with the pictures and their names. The experimental session was preceded by a 20-trials practice session.

## RESULTS

### Reaction times (RTs)

Apparatus failures (2.2%) and naming errors (2.8%) were removed prior to RTs analyses. Correct RTs were submitted to the Van Selst and Jolicoeur (1994) recursive outlier trimming procedure, which excluded an additional 2.4% of the data. Mean naming latencies according to conditions are reported in **Table 1**. In the by-subjects ANOVA ( $F_1$ ), both Type Of Object (possessing vs. not-possessing typical sound) and Presentation Condition (picture accompanied with white noise vs. alone) were treated as within-subjects factors. In the by-items ANOVA ( $F_2$ ), Type Of Object was treated as a between-items factor whereas Presentation Condition was treated as a within-items factor. The analyses showed a significant main effect of Type of Object in the by-subjects analysis,  $F_{1(1, 31)} = 6.8$ ,  $MSE = 3640$ ,

**Table 1 | Mean reaction times (RTs) and percentage of errors (E%) according to conditions.**

Typical sound	White noise				Differences (RT)
	Without		With		
	RT	E%	RT	E%	
With	910	2.9	951	2.3	-41
Without	903	3.1	902	2.9	1

$p < 0.05$ , but not in the by-item analysis,  $F_{2(1, 126)} = 1.4$ ,  $MSE = 34061$ ,  $p = 0.24$ , a significant main effect of Presentation Condition,  $F_{1(1, 31)} = 4.4$ ,  $MSE = 2342$ ,  $p < 0.05$ ,  $F_{2(1, 126)} = 7.5$ ,  $MSE = 3567$ ,  $p < 0.01$ , and, crucially, a significant interaction,  $F_{1(1, 31)} = 4.9$ ,  $MSE = 2854$ ,  $p < 0.05$ ,  $F_{2(1, 126)} = 8.6$ ,  $MSE = 3567$ ,  $p < 0.005$ . Planned comparisons revealed that RTs were significantly slower when objects possessing typical sounds were presented with white-noise with respect to when presented alone,  $t$ -participants<sub>(31)} = 3.1,  $p < 0.005$ ,  $t$ -items<sub>(63)} = 3.8,  $p < 0.001$ . In contrast, RTs for the objects not possessing typical sounds were unaffected by the presence of white-noise, both  $ts < |1|$ .</sub></sub>

### Errors

Mean error percentages are reported in **Table 1**. No effects were significant in the analyses of errors,  $F_s < 1$ .

## DISCUSSION

The present study aimed at assessing the role of sound representation in object recognition. In order to address this issue we have exploited a picture naming task in which target pictures might be accompanied by a white-noise burst. White noise was thought to interfere with the representation of the sound possibly associated with the depicted object. We reasoned that if such a representation is critical for the recognition of objects strongly associated with certain sounds, white-noise interference should affect the naming of pictures representing these objects.

The results are clear cut, as a white-noise burst presented with a to-be-named picture does interfere with picture naming but only if the picture depicts an object possessing a typical sound. There are two aspects of this finding that are worth discussing.

First, in a standard picture naming task participants are only required to name the stimulus they are presented with as quickly as possible, they are *not* required to retrieve particular aspects of the meaning of the stimulus, as its typical color, smell or sound. Thus, the finding that the presentation of white noise interferes with picture naming when the stimulus depicts an object possessing a typical sound suggests that the activation of the auditory representations associated to that object is mandatory upon stimulus presentation.

Second, the fact that the naming of objects possessing a typical sound is interfered with by the concurrent presentation of a white-noise sound-stimulus suggests that the representations of sounds are activated *while* the object is being identified, that is that object-related sound are activated before

complete identification of the object had occurred. In other words, this finding is congruent with a pre-categorical view—and therefore incongruent with a post-categorical view—of the access to object-related sound representations, thus suggesting that object-related sound representations participate in object identification.

Once established that the pre-categorical scenario is more congruent with the above finding than a post-categorical scenario, a question naturally arises: why does white-noise interfere? That is, what is the mechanism that causes this interference? One possibility is to assume that auditory representations are *modal*, in the sense that acquired auditory knowledge is stored (at least partially) in the same systems that subserve auditory processing (Kiefer et al., 2008; Vermeulen et al., 2008). Thus, upon the presentation of a visual object possessing a typical sound, the corresponding modal auditory representation—residing in the auditory processing system—is activated. If the system storing auditory knowledge is also the system subserving auditory processing, then the presentation of an auditory stimulus—e.g., white-noise—will interfere with the possible concurrent activation of auditory representations—e.g., the typical sound of the object (see Connell and Lynott, 2012, for a discussion), which is what we observed.

A similar explanation has been proposed by Matheson et al. (2014) to account for the interference effects they found in a task requiring the execution of irrelevant movements while participants named picture of either animals or inanimate objects. Matheson et al. observed that the naming of manipulable artifacts was affected by concurrent motor activity, whereas no effects of motor activity were found when participants named non-manipulable animals. The authors concluded that the same neural sensorimotor networks are involved in encoding and retrieving object knowledge (cf. Barsalou, 1999, 2008) and the concurrent irrelevant motor activities interfered with the activation of motor programs that were necessary to retrieve object knowledge.

In conclusion, our finding supports a pre-categorical view of the semantic of objects and is consistent with a concept of concepts as collections of mandatorily accessed, related representations (Redmann et al., 2014) which are modal in nature.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.01139/abstract>

## REFERENCES

- Allport, D. A. (1977). "On knowing the meaning of words we are unable to report: the effects of visual masking," in *Attention and Performance VI*, ed S. Dornic (New York, NY: Academic Press), 505–534.
- Barsalou, L. W. (1999). Perceptual symbol system. *Behav. Brain Sci.* 22, 577–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychon. Bull. Rev.* 10, 344–380. doi: 10.3758/BF03196494
- Coltheart, M. (1987). "Functional architecture of the language-processing system," in *The Cognitive Neuropsychology of Language*. eds M. Coltheart, G. Sartori, and R. Job (London: Lawrence Erlbaum Associates), 1–26.
- Connell, L., and Lynott, D. (2012). When does perception facilitate or interfere with conceptual processing? The effect of attentional modulation. *Front. Psychol.* 3:474. doi: 10.3389/fpsyg.2012.00474
- Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., and Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture-word interference paradigm. *Front. Psychol.* 1:177. doi: 10.3389/fpsyg.2010.00177
- Dell'Acqua, R., Lotto, L., and 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. *Behav. Res. Methods, Instrum. Comput.* 32, 588–617. doi: 10.3758/BF03200832
- Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490. doi: 10.1162/08992999563544
- Hocking, J., and Price, C. (2008). The influence of colour and sound on neuronal activation during visual object naming. *Brain Res.* 1241, 92–102. doi: 10.1016/j.brainres.2008.08.037
- Hubbard, T. L. (2010). Auditory imagery: empirical findings. *Psychol. Bull.* 136, 302–329. doi: 10.1037/a0018436
- Humphreys, G. W., Lamote, C., and 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, 535–586. doi: 10.1080/09658219508253164
- Kiefer, M., and Barsalou, L. W. (2013). "Grounding the human conceptual system in perception, action, and internal states," in *Action Science: Foundations of an Emerging Discipline*, eds W. Prinz, M. Beisert, and A. Herwig (Cambridge, MA: MIT Press), 381–407.
- Kiefer, M., Sim, E.-J., Herrnberger, B., Grothe, J., and Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J. Neurosci.* 28, 12224–12230. doi: 10.1523/JNEUROSCI.3579-08.2008
- Laurienti, P. J., Wallace, M. T., Maldjian, J. A., Susi, C. M., Stein, B. E., and Burdette, J. H. (2003). Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Hum. Brain Mapp.* 19, 213–223. doi: 10.1002/hbm.10112
- Matheson, H. E., White, N., and McMullen, P. A. (2014). Testing the embodied account of object naming: a concurrent motor task affects naming artifacts and animals. *Acta Psychol.* 145, 33–43. doi: 10.1016/j.actpsy.2013.10.012
- Mulatti, C., Ceccherini, L., and Coltheart, M. (2014). What can we learn about visual attention to multiple words from the word-word interference task? *Mem. Cognit.* doi: 10.3758/s13421-014-0450-x. [Epub ahead of print].
- Mulatti, C., and Coltheart, M. (2012). Picture-word interference and the response-exclusion hypothesis. *Cortex* 48, 363–372. doi: 10.1016/j.cortex.2011.04.025
- Mulatti, C., Lotto, L., Peressotti, F., and Job, R. (2010). Speed of processing explains the picture - word asymmetry in conditional naming. *Psychol. Res.* 74, 71–81. doi: 10.1007/s00426-008-0182-2
- Nyberg, L., Habib, R., McIntosh, A. R., and Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11120–11124. doi: 10.1073/pnas.97.20.11120
- Pezzulo, G. (2011). Grounding procedural and declarative knowledge in sensorimotor anticipation. *Mind Lang.* 26, 78–114. doi: 10.1111/j.1468-0017.2010.01411.x
- Potter, M. C., and Faulconer, B. A. (1975). Time to understand pictures and words. *Nature* 253, 437–438. doi: 10.1038/253437a0
- Redmann, A., FitzPatrick, I., Hellwig, F., and Indefrey, P. (2014). The use of conceptual components in language production: an ERP study. *Front. Psychol.* 5:363. doi: 10.3389/fpsyg.2014.00363
- Schneider, T., Engel, A., and Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Exp. Psychol.* 55, 121–132. doi: 10.1027/1618-3169.55.2.121

- Van Selst, M., and Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Q. J. Exp. Psychol.* 47A, 631–650.
- Vermeulen, N., Corneille, O., and Niedenthal, P. M. (2008). Sensory load incurs conceptual processing costs. *Cognition* 109, 287–294. doi: 10.1016/j.cognition.2008.09.004
- Vredeveldt, A., Hitch, G. J., and Baddeley, A. D. (2011). Eye-closure helps memory by reducing cognitive load and enhancing visualisation. *Mem. Cogn.* 39, 1253–1263. doi: 10.3758/s13421-011-0098-8

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 13 June 2014; accepted: 19 September 2014; published online: 08 October 2014.*

*Citation: Mulatti C, Treccani B and Job R (2014) The role of the sound of objects in object identification: evidence from picture naming. Front. Psychol. 5:1139. doi: 10.3389/fpsyg.2014.01139*

*This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology.*

*Copyright © 2014 Mulatti, Treccani and Job. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Long-term repetition priming and semantic interference in a lexical-semantic matching task: tapping the links between object names and colors

Toby J. Lloyd-Jones<sup>1\*</sup> and Kazuyo Nakabayashi<sup>2</sup>

<sup>1</sup> Wales Institute of Cognitive Neuroscience, Department of Psychology, Swansea University, Swansea, UK

<sup>2</sup> Department of Psychology, University of Hull, Hull, UK

## Edited by:

Peter Indefrey, University of  
Dusseldorf, Germany

## Reviewed by:

Stéphanie Massol, Basque Center  
on Cognition, Brain and Language,  
Spain

Alexandra Redmann, University of  
Duesseldorf, Germany

Ines Bramao, Universidade do  
Algarve, Portugal

## \*Correspondence:

Toby J. Lloyd-Jones, Wales Institute  
of Cognitive Neuroscience,  
Department of Psychology,  
Swansea University, Singleton Park,  
Swansea SA2 8PP, Wales, UK  
e-mail: t.j.lloyd-jones@  
swansea.ac.uk

Using a novel paradigm to engage the long-term mappings between object names and the prototypical colors for objects, we investigated the retrieval of object-color knowledge as indexed by long-term priming (the benefit in performance from a prior encounter with the same or a similar stimulus); a process about which little is known. We examined priming from object naming on a lexical-semantic matching task. In the matching task participants encountered a visually presented object name (Experiment 1) or object shape (Experiment 2) paired with either a color patch or color name. The pairings could either match whereby both were consistent with a familiar object (e.g., *strawberry* and *red*) or mismatch (*strawberry* and *blue*). We used the matching task to probe knowledge about familiar objects and their colors pre-activated during object naming. In particular, we examined whether the retrieval of object-color information was modality-specific and whether this influenced priming. Priming varied with the nature of the retrieval process: object-color priming arose for object names but not object shapes and beneficial effects of priming were observed for color patches whereas inhibitory priming arose with color names. These findings have implications for understanding how object knowledge is retrieved from memory and modified by learning.

**Keywords:** color, object, name, shape, memory, repetition priming, modality-specific, semantic interference

## INTRODUCTION

Stored knowledge of object color, for instance that a strawberry or stop sign is typically red, can make an important contribution to everyday tasks such as selecting food at the supermarket or using signs when negotiating road traffic. These interactions often require retrieving color knowledge from memory along with other forms of information associated with a particular object or category of objects such as the object name or shape. To understand the properties of the different processing components mediating the memorial retrieval of object-color knowledge however, it is necessary to develop paradigms that tap components selectively. Here, we developed a novel paradigm which engaged the long-term mappings between object names and the prototypical colors for familiar objects and so allowed us to assess effects of learning within the lexical-semantic memory system for object-color knowledge. We examined the role of retrieval processes in the activation of object-color knowledge and the effect they may have on memory performance as indexed by long-term repetition priming (the benefit in performance from a prior encounter with the same or a similar stimulus). In particular, we assessed whether: (a) there were differences in the retrieval of object-color knowledge from verbal and visual modalities; and (b) retrieval modality influenced memory as indexed by priming.

The available evidence suggests that retrieval of object-color knowledge may be modality-specific. First, neuropsychological

evidence suggests a distinction between visual and verbal object-color information, for instance in knowing that a banana is yellow without consulting a visual representation (Beauvois, 1982; Beauvois and Saillant, 1985; Tanaka et al., 2001). Related to this, in studies of the development of object-color knowledge in children, younger children appear to store most of this knowledge via verbal rather than visual processing. For instance, when presented with pictures of a yellow and red banana a young child may not choose correctly, however when asked “What color are bananas?” is able to answer that bananas are yellow (Davidoff and Mitchell, 1993; Gleason et al., 2004). Note as well that, (a) visual object-color information can be accessed from a shape representation (Price and Humphreys, 1989) or via a verbal object-color representation through the use of mental imagery (Davidoff, 1991; Tanaka et al., 2001) and (b) verbal object-color information may have direct access to object and color names (Beauvois, 1982; Beauvois and Saillant, 1985; Davidoff and Mitchell, 1993; although see Tanaka et al., 2001). Second, two studies highlight differences in knowledge retrieval from verbal and visual modalities, and although the paradigms are very different from each other and from the present task they are suggestive. Naor-Raz et al. (2003) used a variation of the (Stroop, 1935) paradigm whereby participants named the colors of diagnostically colored objects (where color is a cue to identity, as in *strawberry*; Tanaka and Presnell, 1999) or object names. For objects, a Stroop-like

effect was evident with slower naming times for atypical (e.g., blue apple) than typical stimuli. In contrast, for object names this pattern was reversed and there were faster naming times for atypical stimuli. They also found that object names, but not objects, facilitated subsequent lexical decisions to associated concepts (for instance, apple primed pie when deciding whether the stimulus was a word or nonword). Naor-Raz et al. (2003) suggested that object names had more ready access to object-color information than visual objects in the color naming task. More recently, Huettig and Altmann (2010) presented the names of diagnostically colored objects within an auditory contextual sentence whilst monitoring eye movements (e.g., “The man thought about it for a while and then he looked at the frog and decided to release it back into the wild”). Object names, but not black-and-white photographs or line drawings, provided access to stored object-color information which in turn shifted overt attention to objects in the display with the same surface color. In some contexts then, it appears that object names can have more effective access to object-color information than visual objects.

In the present study, we assessed priming from object naming onto a lexical-semantic matching task. The rationale was (a) to use object naming to tap the object-color knowledge system whereby links between object and color representations may be activated at a visual (Price and Humphreys, 1989) semantic (Davidoff, 1991; Davidoff et al., 1997; Tanaka et al., 2001) or lexical level (Naor-Raz et al., 2003). Supporting this notion, there is considerable evidence that naming a familiar object is normally mediated at least by three kinds of pre-existing representation: visual input is matched to a stored visual representation of object shape; accessing this stored shape representation enables further access to a semantic representation which provides the basis for recognition; and in order to name a visually presented object a number of additional post-semantic lexical stages involved in name selection and production have also been proposed (Indefrey and Levelt, 2004). Models differ as to whether during naming information transmission at some prior stage stops or is completed before processing at a subsequent stage begins (Schriefers et al., 1990; Levelt et al., 1999) or whether it is continuously fed forward and backward between either some or all representational stages (Humphreys et al., 1995; Rapp and Goldrick, 2000). Nevertheless, in a long-term priming paradigm as used here one would expect activation from initial naming to spread to all parts of the object-color system (e.g., Lloyd-Jones and Humphreys, 1997a,b). In addition, we proposed that (b) long-term priming arises from the activation of processing components engaged across study and test tasks (for a recent review, see Cabeza and Moscovitch, 2013) and therefore from a subset of components activated during object naming and the lexical-semantic matching task. In lexical-semantic matching participants encountered a visually presented object name (Experiment 1) or object shape (Experiment 2) paired with either a color patch or color name. The pairings could either match whereby both were consistent with a familiar object (e.g., *strawberry* and *red*) or mismatch (*strawberry* and *blue*). Accordingly, we proposed that successful responses on match trials required access to lexical-semantic information about familiar objects and their prototypical colors. We used the matching task to probe knowledge about

familiar objects and their colors pre-activated during object naming.

In Experiment 1 we assessed the retrieval of object-color information from object names. We examined the priming of (a) *object name+color patch* (same object name and physical color as at study); (b) *object name+color name* (same object name and color name, where the color name corresponded to the object color at study); and (c) *object name alone* (same object name as at study but with a different color patch or color name to that encountered at study); as compared with (d) *control* (an object name and color patch or color name that had not been encountered previously). The logic was straightforward: if there was similar priming for the conditions where the same object name+color was processed across study and test as compared with the conditions where only the same name was processed across study and test, then the object-color associations activated during the naming task were not utilized by the system(s) mediating performance. In contrast, if there was greater priming for the conditions where the same object name+color was processed across study and test as compared with the conditions where only the name was processed across study and test, then the object-color associations activated during the naming task were utilized by the system(s) mediating performance (for the same logic, see Lloyd-Jones and Nakabayashi, 2009; Lloyd-Jones et al., 2012; and others). We predicted that object-color information would contribute to priming in the lexical-semantic decision task and so there would be greater priming for object name+color as compared with the name alone. We also predicted that priming would be modulated by the nature of the color retrieval cue. There is flexibility, according to processing demands, in the encoding and/or retrieval operations of the memory system for certain object properties (for a review, see Roediger and Srinivas, 1993). For instance, if the task requires a judgment about object size then size may influence priming but otherwise it may not do so (Srinivas, 1996). Similarly, object color can influence priming when it is made relevant to the task but under other circumstances it may not do so (Vernon and Lloyd-Jones, 2003). Complementary evidence also comes from a short-term priming paradigm used by Yee et al. (2012) who found color-name priming (e.g., the word *emerald* primed *cucumber*) but only when attention had been drawn to the color feature by participants previously completing a color-word Stroop task. Our main focus here was the color patch condition where the aim was to probe visual object-color memory. We proposed that the processing component engaged by the physical color during both encoding (as part of the visual object) and retrieval (by the color patch) was visual object-color information. Consistent with this idea, Price and Humphreys (1989) have shown that surface color information contributes at a visual level to object categorization and naming although we note that they found color effects only when color covered the surface of the object (not when it was used as a background color) and here color was only partially superimposed on the name (or shape in Experiment 2). Nevertheless, if we are correct, the pre-activation of visual object-color information by object naming should mediate object name+color processing in the lexical-semantic matching task and produce priming. In contrast, we expected that the color-name cue would not encourage the retrieval of visual object-color memory as effectively

because it would have to do so via the retrieval of verbal object-color information and a process of visual imagery. Rather, we proposed that the color-name cue predominantly would encourage the retrieval of verbal object-color information and as a consequence we would observe reduced long-term priming as compared with the color patch condition.

Finally, there was the intriguing possibility that interference might arise for the color-name cue because of the frequent repetition of a relatively small set of visually presented color names during retrieval. Repetition priming can have short-term negative consequences whereby retrieving a word can interfere with retrieving subsequent words from the same semantic category (for reviews, see Abdel Rahman and Melinger, 2009; Oppenheim et al., 2010). In *semantic blocking* for instance, objects are named more slowly in the context of items from the same category as compared with items from various semantic categories (Belke et al., 1985; Kroll and Stewart, 1994). These semantic context effects on language production are generally short-lived, although interference can arise across filler trials (Wheeldon and Monsell, 1994; Damian and Als, 2005; Howard et al., 2006) and in one study across experimental blocks (Vitkovitch and Humphreys, 1991). Persisting negative effects have been proposed to arise from a combination of (a) *shared semantic activation*, so that activation of one particular word or picture activates both itself and semantically-related concepts; (b) *priming*, whereby the activation/retrieval of a lexical-phonological representation facilitates the subsequent activation/retrieval of that representation, through item-specific mappings from semantics to lexical-phonology; and (c) *competition*, so that item-specific mappings from semantic to lexical-phonological representations also result in the activation of a number of lexical competitors (Howard et al., 2006; Oppenheim et al., 2010). Competition may be resolved either by lateral inhibition within the lexicon (Howard et al., 2006; but see Navarrete et al., 2010) or learning, namely small and persistent experience-driven adjustments to the mappings between semantic and lexical representations which involve strengthening the mappings for the word that is produced and at the same time weakening the mappings for semantically-related words (Oppenheim et al., 2010). Oppenheim et al. (2010) suggest that these negative effects will arise in other tasks which involve semantic-based lexical-phonological processing. Now, when color names are presented with object names here, the three conditions described previously are satisfied. On the basis of neuropsychological and developmental evidence (Beauvois, 1982; Beauvois and Saillant, 1985; Mitchell and Davidoff, 1993) there are direct links between verbal object-color (a form of semantic knowledge) and lexical-phonological color-name representations. In addition, during word recognition and reading aloud, lexical-phonological representations are always activated (although not necessarily fully specified; Frost, 1998; Coltheart et al., 2001). So, on this basis there is (a) shared semantic activation, whereby color names provide access to verbal object-color knowledge in order to make a semantic decision; (b) short-term priming, a small set of visually presented color names are presented repeatedly and item-specific priming may arise from the mappings between verbal object-color and lexical-phonological color-name representations; and (c) lexical competition, item-specific mappings

from verbal object-color knowledge to lexical-phonological representations may result in the activation of a number of lexical-phonological competitors. In a similar fashion, competition may also arise at the level of lexical-orthographic representations as there is evidence that when a word is visually presented there is activation of its orthographic/phonological competitors (McCann and Besner, 1987; Andrews, 1992). In sum, it is plausible that semantic interference will produce longer response times for color names as compared with color patches.

Furthermore, concerning long-term priming, it follows that in a system where information is transmitted continuously between representational stages, activation of visual object-color knowledge from prior object naming may exaggerate competition between subsequent verbal object-color representations and as a consequence inhibit the retrieval of color-name representations relative to the control condition. Consistent with this notion: (a) in an analogous fashion, studies have shown that visual object similarity based on shared shape features can have repercussive effects, exaggerating competition at subsequent semantic and lexical stages of the object naming system (e.g., Vitkovitch et al., 1993; Humphreys et al., 1995; Lloyd-Jones and Humphreys, 1997a,b) and (b) as Damian and Als (2005) describe, a number of studies have demonstrated that the retrieval of an object name can result in that item being a more powerful competitor on subsequent trials in which items from the same category are named (Vitkovitch and Humphreys, 1991; Wheeldon and Monsell, 1994; Vitkovitch et al., 2001). For instance, using a naming to deadline procedure where participants have to respond before they are ready resulting in various kinds of error, Vitkovitch and Humphreys (1991) found that such errors were often *perseverative*—the names of category members which were targets during an earlier block of trials. In sum, we expect inhibitory priming from the color-name retrieval cue.

## MATERIALS AND METHODS

### PARTICIPANTS

There were 189 participants in all; 21 took part in a preliminary color agreement study, 84 took part in Experiment 1 and 84 in Experiment 2. All were undergraduates at the University of Kent and participated for course credit. All had normal color vision and normal or corrected-to-normal visual acuity.

### STIMULI AND APPARATUS

The initial pool of stimuli were color photographs of 75 common objects from number of different categories. Most pictures were taken from an internet website ([www.PhotoObjects.net](http://www.PhotoObjects.net)) with a subset selected via an internet image search using the Google search engine. The objects were selected on the basis that each object had a single diagnostic color and where possible the surface color of each object was based on color agreement scores obtained by Joseph (1997) and Vernon and Lloyd-Jones (2003). We used the imaging software Adobe Photoshop CS2 to create 3 versions of each object: a correctly colored object, a grayscale object, and an incorrectly colored object. To convert correctly colored objects to grayscale all the images were converted into the Lab color mode allowing the separation of luminosity (i.e., the lightness component that can range from 0 to 100) from the

color. The lightness channel was then converted into the grayscale channel by using the grayscale mode. To convert correctly colored objects to incorrectly colored objects, we rotated the correct colors across objects ensuring that correctly and incorrectly colored objects were matched for color frequency and luminosity, with the constraint that each incorrectly colored object was not similar to the correctly colored version (e.g., we did not replace the green of a lettuce with the green of a cucumber). The incorrectly colored objects were created by selecting the surface color of an object which was pasted onto another object by using the color replacement tool. The brightness of the color-replaced object was adjusted by using the brightness contrast tool. The luminosity of grayscale images was also closely matched to that of the colored objects (i.e., there were differences only in the range of 10–15 in the Adobe lightness component).

We then examined color agreement between the surface color of each object (i.e., the color that was assigned to each object by the experimenter) and participants' knowledge of the prototypical color of each object. In a self-paced task, 21 participants wrote down the color of each object in both a perception and a memory condition. In the perception condition, each of 75 correctly colored objects was shown, in random order, one at a time on the computer screen until a response was made, and participants wrote down what they considered to be the surface color of the object. In the memory condition, participants were given the list of the names of 75 objects, and were asked to assign to each object what they thought was the object's most prototypical color from memory. The order of conditions was counterbalanced across participants. It was clear that 15 objects had strong perception-memory color disagreement (i.e., apple, aubergine, chick, chicken, elephant, giraffe, grapes, lion, onion, peach, pepper, pineapple, tank, tulip, and turtle). The surface color of 11 of these objects were then re-colored into the color which the participants thought was the most prototypical color (the surface color of 3 objects remained unchanged because participants had reported the internal rather than external color and a fourth object was excluded because name agreement was low). Finally, we selected 60 objects with the highest perception-memory agreement and prepared them for the lexical-semantic matching task: average color agreement was 80%.

For the lexical-semantic matching task we created color patches using the same correct and incorrect colors used to color the surface of the object in the preliminary study described above and in the object naming task used at study in Experiments 1 and 2. We selected the surface color of the object and pasted that color onto a box using the color replacement tool. Each color patch was partially superimposed onto either the object name or a grayscale image of the object with the color patches positioned equally to the top left, top right, bottom right and bottom left of the object or word. We did this to control as far as possible for potential differences in attention across the conditions. Participants' attention here was to a single object with name/shape+color conjoined. In contrast, had name/shape vs. color been presented spatially separately this might have encouraged subjects to attend more to either the name/shape or the color and possibly to do so to a different extent in the various conditions. The average size of the color patches was matched with

that of the color words. These sizes were also equivalent to the size of the objects and object names which were also matched with each other: 4 cm (h) × 6 cm (w). This was achieved by pairwise matching the size of each object with the size of the corresponding object name and also pairwise matching this size with the size of the corresponding color patch and color name. The font was Century Gothic in upper case 27 point. For the object-name/color-name condition we were concerned that the two components to the stimulus may not be as perceptually discriminable as the components in the other conditions and we therefore adjusted the opacity of the object and color words to 70% in order to make them more readable. A list of the stimuli are given in Supplementary material. **Figure 1** provides examples of correctly and incorrectly colored objects presented in the object naming task and **Figure 2** provides an example of each object and color format combination used in the lexical-semantic matching task for *violin*. The experiment was conducted using SuperLab Pro (Version 2.0.4) on a PC, with a microphone via a voice key system (Cedrus SV-1).

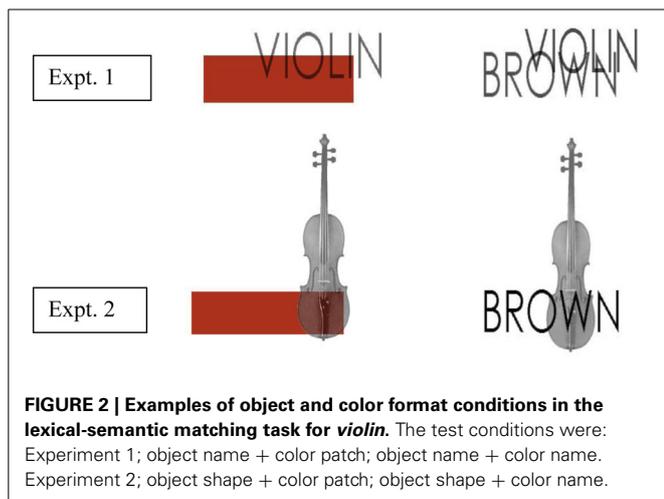
## EXPERIMENT 1

### DESIGN

The experiment comprised two phases: (a) a study phase where correctly and incorrectly colored objects were named followed by (b) a test phase in which a lexical-semantic matching task



**FIGURE 1 |** Examples of correctly and incorrectly colored objects presented in the object naming task.



was performed. In the lexical-semantic matching task participants encountered a visually presented object name paired with either a color patch or color name. The pairings could either match whereby both were consistent with a familiar object (e.g., *strawberry* and *red*) or mismatch (e.g., *strawberry* and *blue*). Accordingly, we proposed that in this experiment (but not Experiment 2) successful responses on match trials required access to lexical-semantic information about objects and their color properties. Therefore, we were most interested in the effects of *priming from naming onto match trials corresponding to correctly colored objects at test* in the following conditions: (a) *object name+color patch* (same object name and physical color as at study); (b) *object name+color name* (same object name with a color name corresponding to the object color encountered at study); and (c) *object name alone* (same object name as at study but with a different color patch or color name from that at study); as compared with (d) *control* (an object name and correct color patch or color name that had not been encountered previously).

To provide a fully balanced design there were 20 objects in the study phase with correct color and 20 with incorrect color. In the test phase there were 60 stimuli, with half in the correct color and half in an incorrect color. Within each of these two conditions at test, stimuli could be the same as at study (i.e., same correct color or same incorrect color; 20 stimuli in all), they could have changed from study (i.e., changed from correct to incorrect color or vice versa; 20 stimuli in all) or they were new stimuli (correct and incorrectly colored; 20 stimuli in all) which provided baselines against which to compare the effects of priming, where the comparison was always within correctly or incorrectly colored conditions at test. In this way, 6 lists of 10 test items were rotated through the study and test conditions, to ensure that each stimulus appeared equally often in each of the conditions across the experiment.

## PROCEDURE

The study task was to name the object, out loud, as quickly and accurately as possible. In the test phase participants were required to make speeded key press responses to indicate whether or not the color was typical of the object. Half the participants pressed

the A key for *match* responses (that the color was typical of the object) and the L key for *mismatch* responses. For the other half of participants the key mapping was reversed. Participants each received 3 practice trials. The test phase followed on from the study phase after a few minutes which was also used to brief participants as to the nature of the task: study and test phases together took approximately 15–20 min to complete. For both study and test phases each trial began with a fixation cross for 250 ms, followed by a 500 ms blank screen, and then by the stimulus which remained on screen until a response had been made.

## RESULTS AND DISCUSSION

For both this experiment and Experiment 2 we adopted the following approach. First, to ascertain whether there was a positive effect of correct color on object naming at study we directly compared correctly and incorrectly colored objects. This was important because if there was no effect of color at study then we would not expect to obtain color priming at test because color knowledge had not been contacted during the course of object naming. For priming however, we were interested only in the data for *match trials corresponding to correctly colored objects at test* because the focus of this experiment was on the retrieval of pre-existing lexical-semantic knowledge concerning real objects and their colors rather than novel representations constructed on-line during learning which is the case for mismatch trials corresponding to incorrectly colored objects at test (Musen et al., 1999; Vernon and Lloyd-Jones, 2003). Indeed, we would not expect pre-existing long-term links in semantic or lexical memory between names and colors for incorrectly colored objects (Davidoff, 1991). We therefore report the findings for mismatch trials corresponding to the priming of incorrectly colored objects at test in Supplementary material. As we expected, priming of incorrectly colored objects arose for name information alone in Experiment 1 and for shape information alone in Experiment 2: there was no priming of either name or shape in combination with color. Note also, if we analyse the response times for correctly and incorrectly colored objects at test together there is a 3-way interaction,  $F_{(2, 144)} = 6.26$ ,  $p = 0.002$ , demonstrating that the findings presented here are robust.

For the naming task at study, a trial was scored an error if: (a) participants provided an incorrect response according to the list of names in Supplementary material. Note, the average name agreement for the objects used in the study was 89%. This approach is more stringent than accepting alternative names produced by some proportion of participants nevertheless it is clear that the study had sufficient power (see summary statistics); (b) the naming latency was 2.5 standard deviations above or below the mean for that participant; or (c) a machine error occurred. In addition, responses to test trials where an error had been made to the object on the corresponding study trial were not excluded. If they were excluded it may have resulted in the removal of objects with names that were intrinsically more difficult to produce and since data from such objects would be excluded from the primed but not the unprimed conditions, this might have resulted in an illusory priming effect. Including such data is a conservative procedure (Wheeldon and Monsell, 1992). We report effect sizes, estimated using partial eta-squared ( $\eta p^2$ ) which according to

generally accepted criteria ranged from medium to large (Cohen, 1988; 0.01 = small, 0.06 = medium, 0.14 = large). For a summary of the data see **Table 1**.

### Study (object naming)

For the analysis of variance (ANOVA), the within-subjects factor was *color* (correct vs. incorrect). For response times, there was a main effect of *color*, with shorter response times for correctly colored objects (917 vs. 993 ms, respectively),  $F_{(1, 82)} = 31.61$ ,  $p = 0.000$ ,  $\eta p^2 = 0.29$ . For accuracy, there was also a main effect of *color*, with greater accuracy for correctly colored objects (87 vs. 75%),  $F_{(1, 82)} = 35.10$ ,  $p = 0.000$ ,  $\eta p^2 = 0.30$ . In sum, correct color benefited object naming performance.

### Test (lexical-semantic decision)

For the ANOVA, the within-subjects factor was *priming* of (a) object name+color (same object name and color as at study) and (b) object name alone (same object name as at study but presented with a different color) compared with (c) control (a correct object name and color that had not been encountered previously). The between-subjects factor was *color format* (color patch vs. color name). We also included the variable *stimulus list* from the rotation design (here and in Experiment 2) in order to increase power as recommended by Pollatsek and Well (1995), nevertheless if we exclude this factor the findings remain unchanged (note also, because of the counterbalanced design across both study and test, no item analyses are reported; Raaijmakers et al., 1999).

For response times, there was a main effect of *color format*,  $F_{(1, 72)} = 15.58$ ,  $p = 0.000$ ,  $\eta p^2 = 0.18$ , with longer response times to color names as compared with color patches. There was also a *priming* x *color format* interaction  $F_{(2, 144)} = 5.55$ ,  $p = 0.005$ ,  $\eta p^2 = 0.07$ . Planned comparisons (*t*-tests) revealed facilitatory priming for object name+color as compared with the control condition in the color-patch condition,  $p < 0.005$ . There was no priming for name alone compared to control. Note also, response times were shorter for object name+color as compared with the name alone condition,  $p < 0.05$ . For the color-name condition, there was inhibitory priming for object name+color as compared with the control condition,  $p < 0.005$ . There was no priming for name alone condition compared to control. Note also, response times were longer for object name+color as compared with name alone,  $p < 0.005$ .

For accuracy, there was a main effect of color format,  $F_{(1, 72)} = 49.58$ ,  $p = 0.000$ ,  $\eta p^2 = 0.41$ , with less accuracy for color patches as compared with color names. There was also a main effect

of *priming*,  $F_{(2, 144)} = 3.80$ ,  $p = 0.025$ ,  $\eta p^2 = 0.05$ , with greater accuracy for object name+color as compared with the control condition,  $p < 0.005$ . The color format x priming interaction was not significant,  $F_{(2, 144)} = 0.05$ ,  $\eta p^2 = 0.01$ ,  $p = ns$ .

Together, these findings indicate that long-term mappings between object names and prototypical colors were activated in memory. However, depending on the nature of the color cue, prior activation either helped or hindered memory retrieval. When the cue was a color patch, memory retrieval was enhanced and when the cue was a color name, memory retrieval was inhibited. There was some evidence that participants traded speed for accuracy and this contributed to the overall effect of color format, with longer response times but also greater accuracy in the color name condition. However, this cannot account for the contrasting influence of the color retrieval cue on priming. Indeed, longer baselines are normally associated with an increase in facilitative priming rather than inhibition which was observed here (Ostergaard, 1994). Moreover, there was no significant correlation between response time and accuracy for any condition: Pearson's *r* values ranged from  $-0.19$  to  $0.18$ . Rather, in combination with the effects of priming we suggest that longer lexical-semantic decision times overall for color names were driven predominantly by semantic interference.

## EXPERIMENT 2

We have argued that important differences in knowledge activation can arise according to the retrieval process. In particular, object names can have more effective access to object-color information than visual objects. To test this account further, we examined whether the findings from Experiment 1 with object names would be reproduced when object shapes provided access to object-color information. Here, we predicted that object shape, but not color, would be used by the memorial system mediating performance and so there would be equivalent priming for object shape+color as compared with the shape-alone condition. The design and procedure was the same as Experiment 1 with the exception that in the test phase decisions were made to matching or mismatching grayscale objects paired with color patches or color names. For a summary of the data see **Table 2**.

## RESULTS AND DISCUSSION

### Study (object naming)

For response times, there was a main effect of *color* with shorter response times to correctly colored objects (946 vs. 1020 ms, respectively),  $F_{(1, 82)} = 44.70$ ,  $p = 0.000$ ,  $\eta p^2 = 0.35$ .

**Table 1 | Experiment 1: Mean response times, standard error (SE), and percentage correct (%) for object name and color in the lexical/semantic matching task.**

	Object name + color			Name			Control		
	Mean	SE	%	Mean	SE	%	Mean	SE	%
<b>Format</b>									
Patch	882	34.2	84.3	922	36.3	81.2	933	41.4	79.6
Name	1160	38.2	95.1	1110	44.6	92.9	1103	43.9	90.3

**Table 2 | Experiment 2: Mean response times, standard error (SE), and percentage correct (%) for object shape and color in the lexical/semantic matching task.**

	Object shape + color			Shape			Control		
	Mean	SE	%	Mean	SE	%	Mean	SE	%
<b>Format</b>									
Patch	1006	32.1	98.6	1020	43.3	98.7	1083	40.5	98.4
Name	846	25.1	98.8	864	29.4	99.1	936	27.1	98.9

For accuracy, there was also a main effect of *color* with greater accuracy for correctly colored objects (86 vs. 77%),  $F_{(1, 82)} = 13.98$ ,  $p = 0.000$ ,  $\eta p^2 = 0.96$ . In sum, correct color benefited object naming performance.

### Test (lexical-semantic decision)

For response times there was a main effect of *priming*,  $F_{(2, 144)} = 18.13$ ,  $p = 0.000$ ,  $\eta p^2 = 0.21$ . Planned comparisons revealed facilitatory priming for object shape+color as compared with the control condition,  $p < 0.001$ , and also for shape alone as compared with the control condition,  $p < 0.001$ . (There was no difference between object shape+color and shape alone,  $p = 0.275$ .) There was also a main effect of *color format*,  $F_{(1, 72)} = 11.62$ ,  $p = 0.001$ ,  $\eta p^2 = 0.14$ , with longer response times to color patches as compared with color names. For accuracy, there was a main effect of *color format*,  $F_{(1, 72)} = 4.5$ ,  $p = 0.037$ ,  $\eta p^2 = 0.06$ , with less accuracy for color patches as compared with color names. The color format x priming interaction was not significant,  $F_{(2, 144)} = 0.22$ ,  $\eta p^2 = 0.01$ ,  $p = ns$ .

The main finding was that object shape, but not color, was used by the memorial system mediating performance. This is consistent with object shape providing a less effective retrieval cue for object-color knowledge than the object name in the lexical-semantic retrieval task. Moreover, if we directly compare priming across Experiments 1 and 2, there is a three-way interaction,  $F_{(2, 328)} = 3.84$ ,  $p = 0.022$ ,  $\eta p^2 = 0.02$ , which provides strong evidence for the claim that names are better at activating color knowledge than shapes. Note, we are not suggesting that object-color knowledge cannot be used at a visual level of analysis, as there is convincing evidence for mappings between object shape and visual object-color information (e.g., Price and Humphreys, 1989; Bramão et al., 2012; Lloyd-Jones et al., 2012). Rather, we are proposing that object shape does not provide an effective retrieval cue in the lexical-semantic matching task.

Independently, there was also an effect of color format with poorer performance for color patches (longer response times and less accuracy) as compared with color names. This contrasts with Experiment 1, where we observed longer response times for color names. Our account of semantic interference requires shared item-specific semantic activation whereby verbal object-color knowledge is activated for a particular object and also other semantically-related objects (which in turn produces lexical-phonological competition). Now, in Experiment 1 this was evident because presentation of the object name and color name specified a verbal object-color entry. Here however, we propose that the object shape initially contacted visual object-color information because object shape and visual object-color information are tightly interconnected (Price and Humphreys, 1989) and this was sufficient to make a decision. This meant that one of the conditions necessary for semantic interference was not met. Poorer performance for color patches likely reflected the fact that they share visual similarity (e.g., orange-red, blue-green) whereas color names do not, and this resulted in the activation of a greater number of visual object-color alternatives which increased competition at the visual level when making a decision. Supporting this notion, words corresponding to stimuli from the same semantic category are no more physically similar

than words corresponding to stimuli from different semantic categories (Carr et al., 1982) whereas pictures corresponding to stimuli from the same semantic category can share physical resemblance; for instance, animals, fruit and vegetables (for a recent review, see Lloyd-Jones and Nettle, 2007). For physical colors, as Braisby and Dockerell (1999) have described, particular instances can fall under different color terms. For instance, a color patch may be equally considered an instance of orange or red, just as dictionaries define olive to be yellow-green, aquamarine to be greenish-blue, and burgundy to be blackish-purple to purplish-red (Collins English Dictionary, 2014). So, for color patches their visual similarity influenced performance when the task was performed on the basis of visual information.

## GENERAL DISCUSSION

The majority of previous work on object-color knowledge has focused on object recognition and found moderate effects of color on categorization and naming (for a recent review and meta-analysis, see Bramão et al., 2011). Here, we examined the retrieval of object-color knowledge from long-term memory. We developed a novel paradigm, which we argue selectively tapped the retrieval of prototypical colors of familiar objects from object names, and used it to examine long-term priming from object naming onto lexical-semantic decisions about objects and their colors and the use of modality-specific access procedures for the retrieval of stored object-color knowledge. We found that priming varied with the nature of the retrieval process. Object-color priming arose for object names (Experiment 1) but not object shapes (Experiment 2) and beneficial effects of priming were observed for color patches whereas inhibitory priming arose with color names. The findings have implications for understanding how object knowledge is retrieved from memory and modified by learning.

The observation that object names enabled the long-term retrieval of object-color information stored in memory complements work on language comprehension showing that visual and motor representations of objects can be activated during word and sentence processing (for a review, see Zwaan, 2004; although see also Rommers et al., 2013). Such findings have often been interpreted in terms of sensorimotor theories of semantic memory whereby object knowledge is represented in a modality-specific rather than amodal fashion (Barsalou, 1999; although see Mahon and Caramazza, 2008). Moreover, our findings support (a) the claim that object names can be more effective than object shapes in retrieving stored object-color knowledge (Naor-Raz et al., 2003); and (b) the independence of object color from shape knowledge (Miceli et al., 2001). The fact that object names were particularly effective object-color retrieval cues also complements recent work by Lupyan and Thompson-Schill (2012) showing that, across short delays in picture verification tasks, semantic information is activated more effectively through the use of verbal labels (such as *cat*) as compared with non-verbal cues (such as the sound of a cat meowing) or words that do not directly refer to the object (the word *meowing*). They suggest that object names are particularly effective because they specify the concept precisely whereas other memory cues may activate a more idiosyncratic semantic representation. Here, object names

shared little physical similarity across exemplars and so activated few semantic object-color alternatives. In contrast, object shapes were visually similar (for instance, exemplars came from fruit, vegetable and animal categories) and in a system where information is continuously fed forward the co-activation of a number of competing visual representations will activate a greater number of semantic object-color alternatives (Vitkovitch et al., 1993; Lloyd-Jones and Nettlemill, 2007). So, for object shapes it is likely that access to stored object-color knowledge was more variable.

We also observed both facilitatory and inhibitory priming which was modulated by the color retrieval cue in the lexical-semantic matching task. As we shall describe, both forms of priming can be explained by learning within a lexical-semantic system comprising visual and verbal object-color knowledge and object and color names. Long-term repetition priming normally has a beneficial effect on performance and is contingent upon the overlap of perceptual, semantic, lexical and response-related processes engaged during encoding and retrieval so that priming is reduced when an item is presented in a different modality or format from study to test (Durso and Johnson, 1979; Rajaram and Roediger, 1993). In addition however, activating/retrieving a particular lexical item can have an adverse short-term effect on the retrieval of other semantically-related lexical items (Howard et al., 2006; Abdel Rahman and Melinger, 2009; Oppenheim et al., 2010). Here, we argue that when the physical color of the object was present during both encoding (as part of the object that was named) and retrieval (an object name+color patch was the memory cue) pre-existing mappings between object names and visual object-color knowledge were activated and mediated facilitatory priming. In contrast, when retrieval was cued by color names two modality-specific conditions arose which together were likely to encourage semantic interference: (a) there was less overlap in processing relative to the color patch condition because the physical color was encoded but color names were presented at retrieval. This meant that the potential benefit of long-term priming was reduced relative to the color patch condition and this allowed any effects of interference to become more apparent; and (b) color names, but not color patches, map directly onto verbal object-color knowledge (Beauvois, 1982; Beauvois and Saillant, 1985; Davidoff and de Bleser, 1993; Davidoff and Mitchell, 1993). We suggest that repeated access to verbal object-color knowledge from color names accrued categorical activation in the verbal object-color system which in turn increased competition between color names at the level of phonology and/or orthography. Long-term inhibitory priming was observed because prior object naming exaggerated effects of semantic interference by making those items particularly powerful competitors in the verbal object-color system (cf., Vitkovitch and Humphreys, 1991; Damian and Als, 2005).

Finally, in previous work we have discussed whether effects of color on object-based memory retrieval reflect either established long-term mappings between object shape and color knowledge or the creation of new temporary short-term perceptual bindings between shape and color (Vernon and Lloyd-Jones, 2003; Lloyd-Jones and Nakabayashi, 2009). For instance, in an event-related potential study Lloyd-Jones et al. (2012) observed color priming for objects in a colored-object decision task (“Is this

object correctly colored?”) from prior object naming. Priming was equivalent for correctly and incorrectly colored objects and evident early in the time course of processing (around 200 ms after stimulus onset). They suggested that the effects arose from perceptual learning which can take place after just a single study trial and has been observed for novel objects (Graf and Schacter, 1989; Wang and Bingo, 2010). Their findings contrast nicely with those presented here where we observed effects of color on memory for familiar but not novel combinations of names and colors. It is likely therefore, that color can influence memory retrieval in a number of ways. We have developed a new paradigm which combined with priming selectively engages the long-term mappings between object names and object-color knowledge and so provides a powerful tool for studying long-term object representation and retrieval.

## ACKNOWLEDGMENT

This work was supported by Leverhulme Trust award F/00236/H to Toby J. Lloyd-Jones.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00644/abstract>

## REFERENCES

- Abdel Rahman, R., and Melinger, A. (2009). Semantic context effects in language production: a swinging lexical network proposal and a review. *Lang. Cogn. Process.* 24, 713–734. doi: 10.1080/01690960802597250
- Andrews, S. (1992). Frequency and neighbourhood effects on lexical access: lexical similarity or orthographic redundancy. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 234–254. doi: 10.1037/0278-7393.18.2.234
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain. Sci.* 22, 577–609.
- Beauvois, M.-F. (1982). Optic aphasia: a process of interaction between vision and language. *Phil. Trans. Roy. Soc. Lond. B* 298, 35–47. doi: 10.1098/rstb.1982.0070
- Beauvois, M.-F. and Saillant, B. (1985). Optic aphasia for colours and colour agnosia: a distinction between visual and visuo-verbal impairments in the processing of colours. *Cognit. Neuropsychol.* 2, 1048. doi: 10.1080/02643298508252860
- Belke, E., Meyer, A. S., and Damian, M. F. (1985). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Q. J. Exp. Psychol.* 58A, 667–692.
- Braisby, N., and Dockerell, J. (1999). Why is colour naming difficult. *J. Child Lang.* 26, 23–47. doi: 10.1017/S0305000998003638
- Bramão, I., Reis, A., Petersson, K. M., and Faisca, L. (2011). The role of colour information on object recognition: a review and meta-analysis. *Acta Psychol.* 138, 244–253. doi: 10.1016/j.actpsy.2011.06.010
- Bramão, I., Faisca, L., Forkstam, C., Inácio, F., Araújo, S., Petersson, K. M., et al. (2012). The interaction between surface color and color knowledge: behavioral and electrophysiological evidence. *Brain Cogn.* 78, 28–37. doi: 10.1016/j.bandc.2011.10.004
- Cabeza, R. and Moscovitch, M. (2013). Memory systems, processing modes, and components: functional neuroimaging evidence. *Perspect. Psychol. Sci.* 8, 49–55. doi: 10.1177/1745691612469033
- Carr, T. H., McCauley, C., Sperber, R. D., and Parmalee, C. M. (1982). Words, pictures and priming: on semantic activation, conscious identification, and the automaticity of information processing. *J. Exp. Psychol. Hum. Percept. Perform.* 8, 757–777.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale NJ: Erlbaum.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi: 10.1037/0033-295X.108.1.204

- Damian, M. F., and Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1372–1384. doi: 10.1037/0278-7393.31.6.1372
- Davidoff, J. (1991). *Cognition Through Colour*. Cambridge, MA: MIT Press.
- Davidoff, J. and Mitchell, P. (1993). The colour cognition of children. *Cognition* 48, 121–137. doi: 10.1016/0010-0277(93)90027-S
- Davidoff, J., and de Bleser, R. (1993). Optic aphasia: a review of past studies and re-appraisal. *Aphasiology* 7, 135–154. doi: 10.1080/02687039308249502
- Davidoff, J., Walsh, V., and Wagemans, J. (1997). Higher-level cortical processing of colour. *Acta Psychol.* 97, 1–6. doi: 10.1016/S0001-6918(97)00020-6
- Durso, F. T., and Johnson, M. K. (1979). Facilitation in naming and categorizing repeated pictures and words. *J. Exp. Psychol. Hum. Learn. Mem.* 5, 449–459. doi: 10.1037/0278-7393.5.5.449
- Frost, R. (1998). Towards a strong phonological theory of visual word recognition: true issues and false trails. *Psychol. Bull.* 123, 71–99. doi: 10.1037/0033-2909.123.1.71
- Gleason, T. R., Fiske, K. E., and Chan, R. K. (2004). The verbal nature of representations of the canonical colours of objects. *Cognit. Dev.* 19, 1–14. doi: 10.1016/S0885-2014(03)00044-3
- Graf, P., and Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 930–940. doi: 10.1037/0278-7393.15.5.930
- Howard, D., Nickels, L., Coltheart, M., and Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition* 100, 464–482. doi: 10.1016/j.cognition.2005.02.006
- Huetting, F., and Altmann, G. T. M. (2010). Looking at anything green when hearing ‘frog’: how object surface colour and stored object colour knowledge influence language-mediated overt attention. *Q. J. Exp. Psychol.* 64, 122–145. doi: 10.1080/17470218.2010.481474
- Humphreys, G. W., Lamote, C., and 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, 535–586. doi: 10.1080/09658219508253164
- Indefrey, P. and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001
- Joseph, J. E. (1997). Colour processing in object verification. *Acta Psychol.* 97, 95–127. doi: 10.1016/S0001-6918(97)00026-7
- Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008
- Levelt, W., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75. doi: 10.1017/S0140525X99001776
- Lloyd-Jones, T. J., and Humphreys, G. W. (1997a). Perceptual differentiation as a source of category effects in object processing: evidence from naming and object decision. *Mem. Cogn.* 25, 18–35. doi: 10.3758/BF03197282
- Lloyd-Jones, T. J., and Humphreys, G. W. (1997b). Categorizing chairs and naming pears: category differences in object processing as a function of task and priming. *Mem. Cogn.* 25, 606–624. doi: 10.3758/BF03211303
- Lloyd-Jones, T. J., and Nakabayashi, K. (2009). Independent effects of colour on object identification and memory. *Q. J. Exp. Psychol.* 62, 310–322. doi: 10.1080/17470210801954827
- Lloyd-Jones, T. J. and Nettlemill, M. (2007). Sources of error in pictures naming under time pressure. *Mem. Cogn.* 35, 816–836. doi: 10.3758/BF03193317
- Lloyd-Jones, T. J., Roberts, M. V., Leek, E. C., Fouquet, N. C., and Truchanowicz, E. G. (2012). The time course of activation of object shape and shape+colour representations during memory retrieval. *PLoS ONE* 7:e48550. doi: 10.1371/journal.pone.0048550
- Lupyan, G., and Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *J. Exp. Psychol. Gen.* 141, 170–186. doi: 10.1037/a0024904
- Mahon, B. Z. and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004
- McCann, R. S., and Besner, D. (1987). Reading pseudohomophones: implications for models of pronunciation assembly and the locus of word frequency effects in naming. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 14–24. doi: 10.1037/0096-1523.13.1.14
- Miceli, G., Fouch, E., Capasso, R., Shelton, J. R., Tomaiuolo, F., and Caramazza, A. (2001). The dissociation of colour from form and function knowledge. *Nat. Neurosci.* 4, 662–667. doi: 10.1038/88497
- Mitchell, P., and Davidoff, J. (1993). The colour cognition of children. *Cognition* 48, 121–137. doi: 10.1016/0010-0277(93)90027-S
- Musen, G., Szerlip, J. S., and Szerlip, N. J. (1999). Attentional requirements for object-location priming. *Psychon. Bull. Rev.* 7, 646–653. doi: 10.3758/BF03213002
- Naor-Raz, G., Tarr, M. J., and Kersten, D. (2003). Is colour an intrinsic property of object representation? *Perception* 32, 667–680. doi: 10.1068/p5050
- Navarrete, E., Mahon, B. Z., and Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychol.* 134, 279–289. doi: 10.1016/j.actpsy.2010.02.009
- Ostergaard, A. L. (1994). Dissociations between word priming effects in normal subjects and patients with memory disorders: multiple memory systems, or retrieval?. *Q. J. Exp. Psychol.* 47, 331–364. doi: 10.1080/14640749408401115
- Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2010). The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. *Cognition* 114, 227–252. doi: 10.1016/j.cognition.2009.09.007
- Pollatsek, A., and Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 785–794. doi: 10.1037/0278-7393.21.3.785
- Price, C. J., and Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *Q. J. Exp. Psychol.* 41, 797–828. doi: 10.1080/14640748908402394
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., and Gremmen, F. (1999). How to deal with ‘The language-as-fixed-effect fallacy’: common misconceptions and alternative solutions. *J. Mem. Lang.* 41, 416–426. doi: 10.1006/jmla.1999.2650
- Rajaram, S., and Roediger, H. L. 3rd. (1993). Direct comparison of four implicit memory tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 765–777. doi: 10.1037/0278-7393.19.4.765
- Rapp, B., and Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychol. Rev.* 107, 460–499. doi: 10.1037/0033-295X.107.3.460
- Roediger, H. L., and Srinivas, K. (1993). “Specificity of operations in perceptual priming,” in *Implicit Memory: New Directions in Cognition, Neuropsychology and Development*, eds P. Graf and M. E. J. Masson (Hillsdale NJ: Erlbaum), 17–48.
- Rommers, J., Meyer, A. S. and Huetting, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychol. Sci.* 24, 2218–2225. doi: 10.1177/0956797613490746
- Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *J. Mem. Lang.* 29, 86–102. doi: 10.1016/0749-596X(90)90011-N
- Srinivas, K. (1996). Size and reflection effects in priming: a test of transfer appropriate processing. *Mem. Cogn.* 24, 441–452. doi: 10.3758/BF03200933
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Tanaka, J. W., and Presnell, L. M. (1999). Colour diagnosticity in object recognition. *Percept. Psychophys.* 61, 1140–1153. doi: 10.3758/BF03207619
- Tanaka, J. W., Weiskopf, D., and Williams, P. (2001). The role of colour in high-level vision. *Trends Cogn. Sci.* 5, 216–222. doi: 10.1016/S1364-6613(00)01626-0
- Vernon, D., and Lloyd-Jones, T. J. (2003). The role of colour in implicit and explicit memory performance. *Q. J. Exp. Psychol.* 56, 779–802. doi: 10.1080/02724980244000684
- Vitkovitch, M., and Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: it’s in the links. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 664–680. doi: 10.1037/0278-7393.17.4.664
- Vitkovitch, M., Humphreys, G. W., and Lloyd-Jones, T. J. (1993). On naming a giraffe a zebra: picture naming errors across different object categories. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 243–259. doi: 10.1037/0278-7393.19.2.243
- Vitkovitch, M., Rutter, C., and Read, A. (2001). Inhibitory effects during object name retrieval: The effect of interval between prime and target on picture

- naming responses. *Br. J. Psychol.* 92, 483–506. doi: 10.1348/000712601162301
- Wang, G., and Bingo, J. (2010). Three-dimensional object recognition learning alters an early ERP component of N1. *Clin. Neurophysiol.* 121, 1473–1480. doi: 10.1016/j.clinph.2010.03.032
- Wheeldon, L. R., and Monsell, S. (1992). The locus of repetition priming of spoken word production. *Q. J. Exp. Psy.* 44, 723–761. doi: 10.1080/14640749208401307
- Wheeldon, L. R. and Monsell, S. (1994). Inhibition of spoken word production in priming a semantic competitor. *J. Mem. Lang.* 33, 332–356. doi: 10.1006/jmla.1994.1016
- Yee, E., Ahmed, S. Z., and Thompson-Schill, S. L. (2012). Colourless green ideas (can) prime furiously. *Psychol. Sci.* 23, 364–369. doi: 10.1177/0956797611430691
- Zwaan, R. A. (2004). “The immersed experiencer: toward an embodied theory of language comprehension,” in *The Psychology of Learning and Motivation*, ed B. H. Ross (New York, NY: Academic Press), 35–62.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 13 March 2014; accepted: 06 June 2014; published online: 24 June 2014.  
Citation: Lloyd-Jones TJ and Nakabayashi K (2014) Long-term repetition priming and semantic interference in a lexical-semantic matching task: tapping the links between object names and colors. *Front. Psychol.* 5:644. doi: 10.3389/fpsyg.2014.00644  
This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.  
Copyright © 2014 Lloyd-Jones and Nakabayashi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

