Published in final edited form as:

Lang Cogn Neurosci. 2019 ; 34(1): 69–86. doi:10.1080/23273798.2018.1501500.

Semantic interference in speech error production in a randomized continuous naming task: Evidence from aphasia

Denise Y. Harvey^{a,b,*}, Hilary J. Traut^{a,†}, and Erica L. Middleton^a

^aMoss Rehabilitation Research Institute, Elkins Park, Pennsylvania, USA

^bDepartment of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[†]Now at the Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, Colorado, USA

Abstract

Naming pictures from the same semantic category hinders subsequent naming from that category (i.e., semantic interference), irrespective of the number of intervening different-category exemplars named. Persistent semantic interference has been well documented in chronometric studies, and has been attributed to experience-driven adjustments in the strength of connections between semantic and lexical representations. However, whether parallel effects exist in speech error data remains unclear. In the current study, people with aphasia, a speaker population prone to naming errors, provided naming responses to a large picture corpus presented in random order that comprised multiple exemplars drawn from several different categories. We found persistent semantic interference in the task in semantic error rates specifically, and that semantic similarity between consecutive related exemplars modulated the effect. The results provide further evidence for the presumed lexical-semantic locus and mechanism(s) underlying semantic interference.

Keywords

semantic interference; lexical retrieval; aphasia; incremental learning

Introduction

A substantial and growing body of evidence shows that oral naming of an object is adversely affected by previous experience naming other items from the same category (e.g., Brown, 1981; Kroll & Stewart, 1994; Vitkovitch & Humphreys, 1991; Wheeldon & Monsell, 1994). Such *semantic interference* occurs even when two same-category items are separated by multiple trials of items from other, unrelated categories (e.g., Brown, 1981; Damian & Als, 2005; Howard, Nickels, Coltheart, & Cole-Virtue, 2006). This feature of semantic interference has motivated contemporary theoretical accounts of lexical access for speech

*Corresponding Author: Denise Y. Harvey, dyharvey@gmail.com; Research Department, Moss Rehabilitation Research Institute, 50 Township Line Road, Elkins Park, PA 19027.

A portion of this work was presented at the 7th annual meeting of the Society for the Neurobiology of Language (Chicago, IL).

Disclosure Statement

The authors report no potential conflicts of interest.

production where each act of word retrieval exerts lasting change in the strength of connections between semantic and lexical representations, persistently influencing future retrievability (i.e., *incremental learning accounts of language use*; Howard et al., 2006; Oppenheim, Dell, & Schwartz, 2010; see also Damian & Als, 2005; Vitkovitch & Humphreys, 1991). Researchers have extensively explored the factors that impact semantic interference using chronometric measures (Belke, 2008, 2013; Belke, Meyer, & Damian, 2005b; Biegler, Crowther, & Martin, 2008; Damian & Als, 2005; Damian, Vigliocco, & Levelt, 2001; Harvey & Schnur, 2016; Howard et al., 2006; Navarrete, Del Prato, & Mahon, 2012; Riès, Karzmark, Navarrete, Knight, & Dronkers, 2015; Rose & Abdel Rahman, 2016, 2017; Schnur, 2014; Vigliocco, Vinson, Damian, & Levelt, 2002). Relatively fewer studies have examined whether similar semantic interference phenomena manifest in naming error rates (Harvey & Schnur, 2015; McCarthy & Kartsounis, 2000; Navarrete, Mahon, & Caramazza, 2010; Schnur, Schwartz, Brecher, & Hodgson, 2006; Wilshire & McCarthy, 2002). This state of affairs is attributable to the literature's focus on neurotypical college-aged adults, a speaker population that rarely makes enough errors to warrant empirical evaluation (cf. Navarrete et al., 2010). Following the assumption that the same mechanism(s) that slow word retrieval also cause the production system to err (Oppenheim et al., 2010), semantic interference and the factors that mediate it as established in naming latencies should also impact naming error production. In the present study, people with aphasia (PWA)—a speaker population in which naming errors are commonplace—provided naming responses to a corpus of pictured objects comprising a large number of exemplars from multiple semantic categories. By analyzing different naming error types, this work goes a step beyond chronometric studies of semantic interference by more directly assessing the stage(s) of word production at which semantic interference and its mediating factors exert their effects.

Characteristics of semantic interference

The *blocked-cyclic naming task* and *continuous naming task* are popular paradigms for studying semantic interference effects in lexical access for speech production. In blocked-cyclic naming, pictures of objects are presented serially for naming in a set that draws from the same category (hereafter, *homogenous condition*) or different categories (hereafter, *mixed condition*). Each set of items is presented for naming in multiple (e.g., 6) contiguous cycles, which together comprise a block (e.g., Damian et al., 2001). There are two indices of semantic interference in this paradigm—(1) greater naming difficulty for items in the homogenous versus mixed condition, collapsed across cycles (e.g., Brown, 1981; Kroll & Stewart, 1994), and (2) *cumulative* semantic interference in the form of a growing disadvantage across cycles in the homogenous versus mixed condition (e.g., Schnur et al., 2006; but see Belke, 2008; Belke & Stielow, 2013). The continuous naming task uses a subtler manipulation of semantic context (e.g., Gordon & Cheimariou, 2013; see also Belke, 2013; Belke & Stielow, 2013; Howard et al., 2006). In that task, several exemplars from each of multiple categories are presented serially for naming with each exemplar presented only once, and the number of other-category trials (hereafter, *lag*) between members of a category varies (i.e., typically anywhere from 2–8; but see Gordon & Cheimariou, 2013; Schnur, 2014). The index of semantic interference in this task corresponds to a linear increase in naming difficulty with each additional item from a category (hereafter, *ordinal position*) in a

manner that is insensitive to lag (Belke, 2013; Howard et al., 2006; Navarrete et al., 2010; but see Schnur, 2014 for a nuanced account of lag effects). Thus, word retrieval becomes increasingly difficult either when categorically related objects are named consecutively and repeatedly, or when each object is named once only separated by trials of unrelated objects.

By most accounts, semantic interference in blocked-cyclic and continuous naming paradigms reflects a fundamental property of semantically-driven lexical access, namely that retrieval of a word (i.e., *target*) activates other words that share its semantic features (i.e., *non-targets*), rendering word retrieval a “competitive” process (e.g., Howard et al., 2006; Levelt, Roelofs, & Meyer, 1999; Oppenheim et al., 2010). Often described to be a function of the degree of semantic feature overlap with a target word, the number and/or strength of activated non-targets influences how difficult it is to distinguish the target from its competitors (e.g., Vigliocco, Vinson, Lewis, & Garrett, 2004; but see Mahon, Costa, Peterson, Vargas, & Caramazza, 2007). For instance, naming is more difficult when a word has many closely-versus distantly-related words in the lexicon (e.g., Mirman, 2011; see also Mirman & Graziano, 2013; Rabovsky, Schad, & Abdel Rahman, 2016). Such *semantic neighborhood density effects* reflect item-inherent attributes (i.e., the organization of the semantic space a word inhabits with respect to other same-category items) that impact the ease of a word’s retrieval. Likewise, studies exploring *semantic distance effects* in paradigms eliciting semantic interference have shown that the competitiveness of a word’s co-activated non-targets can be altered experimentally by manipulating whether a word is named amidst words that are more or less semantically similar. To illustrate, in the blocked-cyclic naming task, semantic interference is greater for sets composed of (1) more semantically similar (e.g., *zebra* and *horse*) versus less semantically similar (e.g., *zebra* and *dog*) category members (Navarrete et al., 2012); and (2) more similar (e.g., clothing and body parts) versus less similar (e.g., clothing and vehicles) categories (Vigliocco et al., 2002). In the continuous naming task, cumulative semantic interference has also been shown to depend on the semantic similarity between same-category exemplars (Rose & Abdel Rahman, 2017; see Alario & Moscoso del Prado Martín, 2010 for discussion). Thus, the ease with which words are retrieved is affected by both the presence or absence of competitors, and the distance of those competitors from the target – whether these reflect item-inherent semantic properties of the target or experimentally imposed competition from item sets.

It is generally accepted that semantic interference phenomena are not reducible to interference from residual activation of related words named in preceding trials. Subsequent to input from semantics, activation in the lexical system dissipates quickly (e.g., unlikely to last beyond a second or two; Dell, 1988) whereas semantic interference in naming is a relatively long-lasting effect (e.g., Damian & Als, 2005; Howard et al., 2006; see also Vitkovitch & Humphreys, 1991). For example, studies investigating the time course of semantic interference reveal that in both the blocked-cyclic and continuous naming tasks semantic interference is unaffected by (1) different response stimulus intervals (RSIs) between trials (e.g., 1 versus 5 seconds; Biegler et al., 2008; Schnur et al., 2006; Schnur, 2014) and (2) naming several unrelated objects between two same-category objects (e.g., Belke, 2013; Damian & Als, 2005; Navarrete et al., 2012; cf. Schnur, 2014). That semantic interference persists across time and accrues in an incremental fashion as a function of prior relevant experience (i.e., naming from the category) has motivated the view that each act of

word retrieval is a learning experience in the form of adjustments to the strength of connection weights between semantic and lexical representations (Damian & Als, 2005; Howard et al., 2006; Oppenheim et al., 2010; Vitkovitch & Humphreys, 1991). To summarize, in addition to enhanced semantic interference with increasing semantic similarity between items, another central feature of the effect is its relative insensitivity to the passage of time and intervening irrelevant experience (i.e., naming from different categories).

Computational accounts of semantic interference

Howard et al. (2006) modeled semantic interference in latencies in the continuous naming task by implementing shared activation (i.e., co-activation of a target and related non-targets), priming, and competition. According to Howard et al., subsequent to target selection (e.g., *horse*), priming strengthens the target's lexical-semantic connection weights (hereafter, *weights*). Prior naming makes *horse* more available for selection and thus, a stronger competitor when subsequently selecting a category exemplar not yet named (e.g., *zebra*). Howard et al. implemented competition via lateral inhibitory connections between lexical candidates, where previously strengthened items exert greater inhibition on a future same-category naming trial. Thus, competition accrues with each named same-category exemplar, resulting in cumulative semantic interference across ordinal positions within a category.

Also implementing the assumed properties of shared activation, priming, and competition, Oppenheim et al. (2010) conducted a series of computational investigations to capture semantic interference effects in both the continuous naming and blocked-cyclic naming paradigms. In contrast to Howard et al., in the Oppenheim et al. model, the learning algorithm legislates two kinds of changes wherein producing a word strengthens its lexical-semantic weights *and* weakens weights of co-activated non-target representations. In the continuous naming task, cumulative semantic interference results from weight weakening that incrementally slows latencies with each named category exemplar. In the blocked-cyclic naming task, weight strengthening in both conditions confers faster retrieval with repeated presentations of words (i.e., repetition priming). However, this facilitation is increasingly attenuated across cycles in the homogenous condition due to weight weakening, resulting in cumulative semantic interference and an overall increase in naming latency (i.e., collapsing across cycles) in the homogeneous versus mixed condition (cf. Navarrete, Del Prato, Peressotti, & Mahon, 2014). Thus, Oppenheim et al. (2010) provide a computational framework in which the same learning mechanism accounts for semantic interference phenomena in latencies in both the continuous naming and blocked-cyclic naming paradigms.

Oppenheim et al.'s computational architecture also captured semantic interference effects in errors with minimal modification. By adding noise to how individual word units are activated from semantic features while leaving other model parameters and processes unchanged, the model simulated both error rates and types of errors documented in studies of PWA performing blocked cyclic naming (i.e., Schnur et al., 2006, Experiment 2; Hsiao, Schwartz, Schnur, & Dell, 2009). Specifically, the noise-added version of the model

produced increased semantic errors (e.g., saying “cat” for the picture DOG) and omission errors (i.e., failure to produce a response) in homogenous versus mixed blocks; and, this effect increased across cycles, consistent with Schnur et al. (2006, Experiment 2). The model also successfully captured an additional aspect of Schnur et al.’s error data reported in Hsiao et al. (2009) in which the observed increase in semantic error rates represented erroneous retrieval of intra-versus extra-experimental semantic competitors. To summarize, in the Oppenheim et al. model, the same set of processes underlie semantic interference as it manifests in latencies and error rates in either blocked-cyclic or the continuous naming task.

However, there are reasons to question whether the same set of processes are involved in the manifestation of semantic interference across the two paradigms and dependent measures. To start, a host of studies point to important processing differences between blocked-cyclic and continuous naming. Blocked-cyclic naming entails repeatedly naming a small set of exemplars, which evidence suggests places greater demands on control mechanisms and/or invites subject strategies that are less likely to be involved in continuous naming (e.g., Crowther & Martin, 2014; see Belke & Stielow, 2013 for review). For instance, semantic interference increases with the addition of a working memory load (induced by a digit-span task) in blocked-cyclic but not continuous naming (Belke, 2008; Belke & Stielow, 2013). Furthermore, interference effects within the same individuals are not correlated across the two tasks (Hughes & Schnur, 2017). Similar distinctions between interference effects in the two tasks have likewise been found in studies of PWA. In Schnur et al. (2006) the increase in errors in the homogeneous versus mixed conditions was reliable across the group, but only significant in—and thus carried by—the Broca’s subset of PWA. Broca’s aphasia is typically associated with damage to the left inferior frontal gyrus (LIFG), a region within the prefrontal cortex (PFC) important for cognitive control processes that serve to select a target from among competing representations (reviewed in Novick, Trueswell, & Thompson-Schill, 2010). This, along with the finding that percent damage to the LIFG predicts the magnitude of cumulative semantic interference in errors in blocked-cyclic naming (Schnur et al., 2009), suggests a role for dysfunctional cognitive control and/or controlled selection in how semantic interference manifests in this task (see also Biegler et al., 2008; Harvey & Schnur, 2015; Wilshire & McCarthy, 2002). By contrast, PFC damage had no effect on the magnitude of cumulative semantic interference in the continuous naming task in a study comparing PWA with age- and education-matched controls (Riès et al., 2015).

Regarding dependent measures, because of the likely greater involvement of cognitive factors in the manifestation of semantic interference in blocked-cyclic naming, continuous naming may be the more preferred paradigm for testing whether the same set of processes govern semantic interference effects in latencies and naming errors. Furthermore, continuous naming is a more naturalistic task that more closely resembles the experience of everyday naming. However, among the few continuous naming studies that have analyzed error rates in neurotypical speakers (Belke, 2013; Howard et al., 2006; Navarrete et al., 2010; Rose & Abdel Rahman, 2017; Runnqvist, Strijkers, Alario, & Costa, 2012; Schnur, 2014), only two prior studies have shown that errors accumulate across ordinal position (Navarrete et al., 2010; Rose & Abdel Rahman, 2017). Likewise, the one study to investigate PWA, a speaker population where variability in naming accuracy may be expected to provide power to detect cumulative semantic interference, found a cumulative effect only in latencies, not errors

(Riès et al., 2015). In addition to the lack of relevant evidence, another motivation for explicating learning-based accounts of semantic interference in terms of naming errors is the potential relevance to clinical or applied domains, as naming errors are more disruptive to effective communication than delayed, accurate responding. Thus, in the present study we employ a version of the continuous naming task in PWA. Along with other indices of semantic interference described below, of key interest is whether semantic interference in error rates accrues as a function of ordinal position in a fashion that is independent of lag. Such a finding would support the notion of a qualitatively similar learning mechanism underlying the manifestation of semantic interference in latencies and speech errors.

Another aim of the current research was to explore whether semantic similarity of an exemplar to its category (hereafter, *item similarity*) influences subsequent naming from that category. As mentioned above, previous research shows that semantic interference effects are enhanced when related items are more versus less semantically similar (Navarrete et al., 2012; Rose & Abdel Rahman, 2017; Vigliocco et al., 2002). However, to our knowledge, only one study using continuous naming has investigated semantic distance effects (Rose & Abdel Rahman, 2017). That study found cumulative semantic interference for very closely related items (e.g., primates), whereas more distantly related same-category items (e.g., animals) showed no such effect. These findings could be consistent with a framework where semantic similarity must reach a threshold to trigger incremental learning. Alternatively, it is possible that their study design was underpowered to detect existing but weak interference effects in the distantly related same-category items. In the present study, we re-examine an effect of semantic similarity on semantic interference manifesting in error rates, but use a continuous measure of semantic similarity. To more directly establish a causal effect of semantic similarity on incremental learning than in prior work, we focus on the impact of an exemplar's similarity to its category on naming error incidence for the upcoming exemplar in that same category.

Notably, in going a step beyond previous research examining errors in the continuous naming task, we measure semantic interference phenomena as they may manifest in specific error types. The relationship between a naming error and target response provides evidence of the error's stage of origin during lexical access. Except in cases of multi-modal semantic deficits from stroke (for discussion, see Jeffries & Lambon Ralph, 2006), it is generally understood that a preponderance of semantic naming errors in aphasia likely reflect problems in the mapping between semantics and an intermediary lexical representation (hereafter, *word*) prior to phonological processing (e.g., Caramazza, 1997; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Rapp & Goldrick, 2000; Schwartz, Dell, Martin, Gahl, & Sobel, 2006; cf., Caramazza & Hillis, 1990). As a point of contrast, phonological errors – word or nonword errors that phonologically resemble a target (e.g., “grain” or “prain” for train, respectively) – are generally understood to reflect difficulty in phonological retrieval and/or encoding after successful lexical-semantic encoding (e.g., Caramazza, 1997; Dell et al., 1997; Goldrick & Rapp, 2007). Thus, we evaluate whether semantic interference phenomena involve adjustments in lexical-semantic weights specifically (e.g., Oppenheim et al., 2010; see also Damian & Als, 2005) by measuring such phenomena in two binary dependent variables, one corresponding to semantic errors versus all other responses, and the other corresponding to all other errors (hereafter *other errors*, which includes phonological

errors) versus all other responses. Because latencies are sensitive to difficulty arising from multiple stages of processing (e.g., Belke, 2008; Belke & Stielow, 2013), an advantage of our approach is it more directly taps semantic interference as it impacts the lexical-semantic stage of word retrieval. Furthermore, by studying semantic errors, we are in a position to confirm or qualify conclusions of chronometric studies that have examined the locus of semantic interference effects (e.g., Belke, Brysbaert, Meyer, & Ghyselinck, 2005a; Belke, 2013; Damian et al., 2001; Harvey & Schnur, 2016; Kroll & Stewart, 1994; Riley, McMahon, & de Zubicaray, 2015; Vitkovitch & Humphreys, 1991).

It is worth noting that the other error dependent variable includes circumlocutions and omissions, and these error types may have a lexical-semantic origin, as damage to a brain region associated with semantic error production – the mid-anterior temporal lobe (e.g., Schwartz et al., 2009) – has also been associated with circumlocutions that describe rather than name the target (Schwartz et al., 2011; Walker et al., 2011) and omissions (Chen, Middleton, & Mirman, 2018). However, we decided to restrict the main dependent variable of interest to semantic errors because, compared to circumlocutions and omissions, semantic errors are most transparently reflective of semantic-to-lexical mapping failure. Specifically, producing a semantically-related word instead of the target is suggestive of co-activation of semantically related words and errant selection of a competitor. In the case of circumlocutions or omissions, as no naming attempt is provided, the inference relating the naming failure to faulty semantic-to-lexical mapping is necessarily less direct.

The current study

The present study administered a form of the continuous naming task to a group of 15 PWA with word retrieval impairments. The task involved collecting naming responses to a large (N=615) picture corpus of common, everyday objects. The items were presented in random order at each of two complete administrations of the corpus per participant. The corpus comprised 19 categories with a range of exemplars from each category (hereafter, referred to as “related” items/categories). The remaining unclassifiable items formed an “unrelated” category that was included as a control condition in Analysis I (see below). Due to random presentation, large corpus size, and variable number of exemplars per category, the procedure eliminated as much as possible recognition of the relationship of category exemplars across trials. A normative study on the corpus items provided the item similarity values, a by-item measure of semantic distance that captured the degree to which a particular item was similar to other experimental items in its category.

With this experimental design, we conducted three analyses to investigate whether semantic interference phenomena established in naming latencies are similarly exhibited in naming errors. Briefly, the goal of *Analysis I* was to establish an overall semantic interference effect by comparing error rates when naming related versus unrelated items presented in a random (unstructured) fashion (following Gordon & Cheimariou, 2013). *Analysis II* provided a test of incremental learning as a driver of semantic interference in our task by inspecting whether error rates accrue as a function of ordinal position in the related categories and whether such accrual is independent of lag. *Analysis III* sought to clarify whether semantic distance impacts incremental learning in errors by investigating whether error rates for a given target

increased as a function of the item similarity of the preceding same-category exemplar. Across all analyses, we investigated a possible lexical-semantic locus of these semantic interference phenomena by measuring changes in semantic error rates specifically. Parallel analyses on other errors were carried out to (1) identify whether the factors under investigation uniquely impact semantics-to-word mapping specifically; and (2) provide control comparisons for other purposes, as described in more detail below.

Method

Participants

Participants gave informed consent under a protocol approved by the Institutional Review Board of Einstein Healthcare Network, and were reimbursed \$15 per hour of participation. The participants were selected from a large pool (>150) of well-characterized and potentially available people with aphasia following stroke for enrollment into treatment studies. Enrollment continued into those treatment studies to reach the desired N for each study. The current dataset is composed of the naming baselines for participants enrolled into those studies (Middleton, Schwartz, Rawson, & Garvey, 2015; Middleton, Traut, & Stabile, 2018). The goal of those studies was to evaluate the impact of different kinds of naming treatment on naming impairment attributable to lexical access deficit. Thus, the sample is not an unselected sample of PWA, but is restricted to individuals that exhibit a lexical access disorder. This is important for the goal of delineating the action of the learning mechanisms under study as impacting lexical processes, rather than processes that are peripheral to word retrieval.

The 15 participants (four female) were all right-handed native English speakers with chronic aphasia (> 6 months post-stroke onset) secondary to left hemisphere stroke. Mean (*M*) age was 60 years (standard deviation (*SD*) = 7.9; *range* = 49–77), mean education level was 15 years (*SD* = 2.9; *range* = 10–21), and mean months post onset was 56.1 (*SD* = 44.2; *range* = 8–145). Aphasia severity as determined by the Western Aphasia Battery (WAB) Aphasia Quotient (AQ) (Kertesz, 1982) varied considerably across the group (*M* = 74.9; *SD* = 13.6; *range* = 48.4–92.9). Six participants were diagnosed with fluent aphasia (conduction and anomic subtypes), and nine with nonfluent aphasia (Broca's, transcortical motor; see Table 1).

Furthermore, for considerations of experimental sensitivity, it was important that semantic naming errors committed in the main task were generally attributable to faulty semantic-to-lexical mapping rather than dysregulated semantics or degraded semantic/lexical representations. Thus, the participants selected for inclusion in the study demonstrated generally good nonverbal semantic comprehension (mean accuracy: 91% on the Pyramids and Palm Trees (PPT) test; Howard & Patterson, 1992) and word comprehension (mean accuracy: 93% on a spoken word-to-picture verification task; Roach, Schwartz, Martin, Grewal, & Brecher, 1996). See Table 1 for participant scores and group averages on relevant neuropsychological assessments.

Materials

Stimuli consisted of 615 pictures of common objects selected from published picture naming corpora (Szekely et al., 2004; Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010) and various Internet sources. Visual complexity and name agreement values were taken from published corpora when available. When these values were unavailable, they were obtained via normative studies in which at least 40 responses were collected per item to determine visual complexity and name agreement. Properties of the pictures and target names are summarized in Table 2. Log frequency values for picture names were collected from SUBTLEX_{US} (Brysbaert & New, 2009). Picture names in the 615-item corpus that were not listed in the SUBTLEX_{US} database were assigned a log frequency of zero (i.e., 33 items).

For the purposes of the present study, the 615-item corpus was divided into the 19 related categories based on category production norms (Van Overschelde, Rawson, & Dunlosky, 2004) and experimenter intuition (see Table 3). The related categories were formed to maximize within-category similarity and were composed of between 16 to 37 exemplars ($M = 28$, $SD = 6$). The unrelated category comprised 78 items that were unclassifiable into any of the related categories.

The related categories were organized around phylogenetic relationships (i.e., natural taxonomies, e.g., mammals) or functional similarities (e.g., vehicles; kitchen items; furnishings). Item similarity ratings were collected using Amazon's Mechanical Turk. In this procedure, participants provided two types of responses--(1) naming, followed by (2) a similarity rating--to each depicted exemplar in a category. First, the exemplars of the category (e.g., the 25 body parts) were presented one at a time, and the participant typed in the name for each exemplar (e.g., finger, chest, lung) to ensure the participant was accessing the semantics of each image. Immediately following the naming of the last exemplar (e.g., nose) in the category, the items in that category were again presented one at a time but in a new random order for the similarity rating. For that rating, the participant was instructed to consider all items in the set they had just named, and to base their ratings off a word or brief phrase they felt described the set of images. Participants were instructed to rate each item for how well it 'fit' with the other members in the set on a 1 – 5 scale (1 = fits very well; 5 = does not fit well). As such, these ratings served as a measure of the semantic similarity between a given exemplar and other exemplars in its category, with lower values corresponding to higher similarity. All 19 related categories underwent this norming with a minimum of 26 responses per item per category. Table 3 lists the categories with token exemplars and average item similarity rating per category.

Procedure

Each participant provided naming responses to the 615-item corpus twice without feedback, providing baseline observations for later naming treatment (Middleton et al., 2015, 2018). Data from both administrations of the 615-item corpus were considered for analysis in the present study (see *Analyses* section for detailed information regarding data filtering). Each administration of the corpus was completed in different weeks, requiring 1–2 sessions per administration. The pictures were presented in random order on a desktop or laptop computer for confrontation naming. On each naming trial, the picture was shown and the

participant was provided up to 20 seconds to produce the name. Participants were asked to name the picture as best they could. The experimenter advanced the trial when the participant indicated they were finished attempting to name the picture. This procedure was instituted in order to avoid experimenter-provided feedback of any kind. If the participant did not indicate they had given their final answer within 20 seconds, the trial ended automatically. An intervening “Ready?” screen separated trials, and the experimenter initiated each trial when the participant was ready.

Response coding

Participants’ verbal responses were digitally recorded, transcribed, and checked by research assistants trained in the International Phonetic Alphabet. Expert coders applied a coding scheme to the first complete (i.e., nonfragment) response on each trial adapted from the PNT (Roach et al., 1996; see Schwartz et al., 2006 for more details). A response was coded as accurate if all target phonemes were produced in the correct order with no erroneously inserted phonemes. A response was also counted as accurate if the singular form of the target was produced in place of the plural or vice versa (e.g., socks for the target sock). The primary error category of interest was *semantic error*, which included a substituted noun that was a synonym of the target, a category coordinate, superordinate/subordinate, or strong associate (e.g., “fire” for the target grill). Noun responses that were semantically *and* phonologically related to a target (i.e., mixed errors) were classified as semantic errors because the lexical status of the response error and its semantic relationship to the target suggests origination from retrieval of the wrong word from semantics rather than faulty phonological mapping coincidentally creating a semantically-related error response (see Dell et al., 1997). All inaccurate responses that were not classified as semantic errors were coded as “other errors.” Omissions and circumlocutions were also coded as other errors.

Analyses

The procedure produced 18,450 trials (i.e., 2 administrations $\times$ 615 items $\times$ 15 participants). Seventy-two trials were dropped from the dataset due to experimenter error. Of these, 67 were eliminated from one administration of the corpus for one participant because the experimenter erroneously presented the objects in alphabetic rather than random order. We further filtered the dataset to include only the first six items per related category presented within a single session for a given participant (recall that for all participants, a single administration of the corpus required more than one session). This was done to maximize power to detect an effect of ordinal position within a category, as semantic interference is usually assessed within a single session rather than across multiple sessions held on different days. Moreover, due to the random presentation order of the 615-item corpus across multiple sessions, the number of items per related category presented within a session varied considerably across participants. Restricting the dataset to the first six items per related category in a session maximized the number of observations per cell of the design—98% (i.e., 1080) of the 1102 total possible 6-item sets contained six items per category (Note: 1102 sets is equal to 4 sessions $\times$ 19 categories $\times$ 13 participants *plus* 3 sessions $\times$ 19 categories $\times$ 2 participants). The 22 incomplete sets (i.e., 5 or fewer items per category) were included because the missing observations ($n=46$) constituted a miniscule percentage of the related data set (0.7%), and excluding the incomplete sets from the analyses only changed

the estimates trivially. For the unrelated category, all valid unrelated item trials were included. In sum, the filtered dataset included 6566 trials of related items and 2329 trials of unrelated items, for a total of 8895 trials.

Analyses were conducted using logistic mixed effects regression implemented in the lme4 package of R version 3.2.1 (R Core Team, 2015; see also Baayen, Davidson, & Bates, 2008). The dependent measure of primary interest was a binary variable corresponding to semantic errors versus all other responses (i.e., correct responses and other errors). Analogous, but separate, models were applied to a secondary (binary) dependent variable corresponding to other errors versus all other responses (i.e., correct responses and semantic errors). Semantic and other error models reported in each of the three analyses contain the same trial-level data, with the only difference being the outcome measure.

We adopted a model comparison approach to assess whether the fixed effects of theoretical relevance impacted performance above and beyond contributions from other factors that may affect naming. For each of the analyses, we first fit a base model to account for the potential influence of variables not of theoretical influence (hereafter, *covariates*) on error production. To identify covariates that significantly contribute to the outcome measure, and in turn need to be controlled for when testing the impact of variables that are of theoretical interest (hereafter, *fixed effects of interest*), we used an exploratory, data-driven approach as implemented in the drop1 (chi-squared test) package of R (discussed in Mirman, 2014). In all base models, the following covariates were considered for inclusion: log-transformed session trial number (i.e., to account for fatigue across a given session; see Alario & Moscoso del Prado Martin, 2010; Gordon & Cheimariou, 2013), and item-specific factors including log-transformed word frequency from the SUBTLEX_{US} project (Brysbaert & New, 2009), name agreement, number of phonemes (hereafter, *length*), and visual complexity (see Materials). Covariates were removed from the base model if they did not significantly predict naming performance prior to adding the fixed effects of interest (see Results for covariates included in each of the base models and the fixed effects of interest assessed in each analysis). To the base model, we sequentially added fixed effects of interest to assess if their inclusion significantly improved model fit by a chi-squared log-likelihood test.

Lastly, all models described below included a by-participants random intercept to account for individual differences in the outcome measure in the context of models assessing group-level effects. By-item random effects were not included because less than 13% of items submitted for analyses were repeated across all participants. For all models, maximal random effect structures (Barr, Levy, Scheepers, & Tily, 2013) were implemented when possible (i.e., by-participant random slopes for the fixed effects of theoretical relevance) and simplified if required for model convergence. We included by-participant random slopes in the base model in cases where we were able to fit by-participants random slopes without triggering convergence warnings (indicative of model overfitting), following the recommendation of Barr et al. (2013) (see p. 277 for discussion).

Results

In the following sections, we describe covariates included in each base model and the fixed-effects of theoretical interest for each of the three analyses. We focus on results pertaining to the fixed-effects of theoretical interest, but full model results including estimates and associated significance values for covariates are provided in Supplementary Material.

Analysis I: Effect of relatedness

If semantic interference manifests in error rates, *and* the effect arises at a lexical-semantic level, we expect to observe increased semantic error rates when naming related versus unrelated items presented in a random (unstructured) fashion (following Gordon & Cheimariou, 2013). To assess whether the locus of the effect was specific to the lexical-semantic level, parallel analyses were conducted on the other errors dependent variable. Analysis of other errors also served as an additional control comparison to establish that general item difficulty did not differ for the related versus unrelated categories. As such, the fixed effect of interest was category type (related versus unrelated items).

Data included in Analysis I consisted of 8,895 responses, with 1,093 classified as semantic errors, 2,732 classified as other errors, and 5,065 classified as accurate. Table 4 summarizes the count and percentages of semantic errors, other errors, and accurate responses as a function of category type (related versus unrelated items). Fixed/random effects structures and model comparison results are reported in Table 5 (see Supplementary Material for model output). Covariates that significantly contributed to the base model fit of semantic errors included log-transformed word frequency, name agreement, and visual complexity. The base model for other errors additionally included length as a covariate. Both semantic and other error models included by-participant random intercepts and by-participant random slopes for the fixed effect of category type.

For the Analysis I models, adding the two-level factor of category type (related versus unrelated) to the base model significantly improved model fit for semantic errors, $\chi^2(1) = 8.36$, $p = .004$, and other errors, $\chi^2(1) = 15.45$, $p < .001$. However, the effects went in opposite directions. Participants made more semantic errors for related versus unrelated items [*Estimate* = $-.44$, *SE* = $.14$, $p = .001$], but *fewer* other errors for related versus unrelated items [*Estimate* = $.36$, *SE* = $.07$, $p < .001$]. Note that in both models the related items serve as the reference condition, resulting in a negative coefficient in the semantic error model and a positive coefficient in the other error model.

The results from Analysis I support the prediction that semantic interference impacts error production in an unstructured version of the continuous naming task as demonstrated by increased semantic error rates for related versus unrelated items. Importantly, that the effect manifested in semantic (and not other) errors points to a lexical-semantic locus of the semantic interference effect, and that the related items were not simply more error-prone overall. Thus, the findings are consistent with the conclusion that the related items were more susceptible to semantic error production because of interference from previously named category exemplars (i.e., semantic interference was contextually imposed).

However, we cannot rule out the possibility that related items differ intrinsically from unrelated items (e.g., higher semantic neighborhood density), resulting in greater semantic error rates (e.g., Mirman, 2011; Mirman & Graziano, 2013). Specifically, by virtue of being amenable to categorization, the related items likely have more and/or stronger semantic competitors in the lexicon compared to unrelated items, increasing the likelihood of semantic error production. Thus, in Analysis II, we focused on the related items only, and analyzed error rates as a function of ordinal position to more directly assess a role for incremental learning in semantic error production in our task. Importantly, we first assessed the alternative explanation of the findings of Analysis I, namely that intrinsic semantic properties of items in the related categories may have contributed, at least in part, to the finding of increased semantic error rates for related versus unrelated items. Here, we used our measure of item similarity (described in the Materials section) as a proxy for intrinsic semantic properties that may affect semantic error rates on the current trial. We then tested whether semantic error rates increase with ordinal position after accounting for the impact of an item's similarity on current trial responding. We also explored a predicted property of such a learning-based mechanism, that of insensitivity to lag.

Analysis II: Test of incremental learning

If incremental learning affects error production in our continuous naming task, and if such learning impacts lexical-semantic retrieval specifically, then semantic but not other errors are expected to accrue across ordinal position. Indicative of learning, and not short-term influences from residual activation of prior category members, we expect rate of accrual to be unaffected by lag. To assess whether the item-inherent factor of semantic similarity influences performance independent of prior naming from the category, we first added to the base model our measure of item similarity. This also allowed us to disentangle potential item-level (current trial) effects from those that are indicative of experience-dependent (incremental learning) semantic interference, i.e. the effect of ordinal position within a category. Thus, in Analysis II, the fixed effects of interest were item similarity, ordinal position, lag, and the interaction between ordinal position and lag.

Of the 6,566 responses to related items, 866 were classified as semantic errors, 1,903 were classified as other errors, and 3,797 were classified as correct. Table 6 reports the count and percentages of semantic and other errors as a function of ordinal position along with average item similarity at each ordinal position and lag between ordinal positions. Covariates that significantly contributed to the base model fit of semantic errors included log-transformed word frequency, name agreement, and visual complexity. The base model for other errors additionally included length and log-transformed session trial number. Both semantic and other error models included by-participant random intercepts. Fixed/random effects structures and model comparison results are reported in Table 7 (see Supplementary Material for model output).

To assess the impact of item similarity on naming performance, we transformed each item's similarity value into a z-score derived from the distribution of its category's similarity ratings. Adding the z-transformed item similarity measure to the base model significantly improved model fit for semantic errors, $\chi^2(1) = 7.70$, $p = .006$, with the direction of the

coefficient indicating increasing semantic errors with increasing item similarity [$Estimate = -.11$, $SE = .04$, $p = .007$; i.e., recall that in the similarity rating task, a lower item score corresponds to higher similarity]. The model fit for other errors also significantly improved with the addition of item similarity z-score, $\chi^2(1) = 24.35$, $p < .001$. However, in contrast to semantic errors, other errors *decreased* with increasing item similarity [$Estimate = .15$, $SE = .03$, $p < .001$]. We then added the fixed effect of ordinal position within a category (1–6) to determine whether incremental learning impacts semantic error production above and beyond item similarity. Ordinal position significantly improved model fit of semantic errors $\chi^2(1) = 6.64$, $p = .01$, where the positive coefficient [$Estimate = .06$, $SE = .02$, $p = .01$] indicates that semantic errors increased as a function of ordinal position within category. Ordinal position also improved model fit of other errors, $\chi^2(1) = 8.11$, $p = .004$, but here too the effect was in the opposite direction: other errors *decreased* across ordinal positions [$Estimate = -.07$, $SE = .02$, $p = .004$] (see Figure 1).

To evaluate whether lag interacts with rate of cumulative interference, lag was log-transformed (i.e., *log-lag*) due to a non-normal distribution. Furthermore, the dataset was constrained to ordinal positions 2 through 6 because lag is not relevant to the first item that appears from a related category within a session. As in the previous model comparison with all 6 ordinal positions, we added item similarity to the base model, and found that it improved model fit in the analysis on semantic errors, $\chi^2(1) = 5.65$, $p = .017$, and other errors, $\chi^2(1) = 20.59$, $p < .001$, in a manner consistent with the previous results (i.e., semantic errors *increased* with increasing item similarity [$Estimate = -.10$, $SE = .04$, $p = .02$] whereas other errors *decreased* with increasing item similarity [$Estimate = .15$, $SE = .03$, $p < .001$]). Also consistent with the previous results, adding the fixed effect of ordinal position (2–6) significantly improved model fit of semantic errors, $\chi^2(1) = 4.40$, $p = .036$, with increasing semantic errors [$Estimate = .06$, $SE = .03$, $p = .036$] across ordinal positions within a category. However, neither log-lag nor the interaction between log-lag and ordinal position improved model fit in the analysis of semantic errors, p 's $> .54$. In the analysis of other errors, adding the fixed effect of ordinal position significantly improved model fit, $\chi^2(1) = 5.27$, $p = .022$; however, the model produced a convergence error, indicating that the results cannot be reliably interpreted. Moreover, adding the fixed effects of lag and its interaction with ordinal position did not improve model fit, p 's $> .11$.

To summarize, in Analysis II, semantic errors accrued across ordinal positions within a category in a fashion that was insensitive to lag – even after accounting for the influence of semantic similarity on semantic error production on a given trial. That semantic interference in the current study was insensitive to lag parallels prior chronometric evidence (e.g., Gordon & Cheimariou, 2013; Howard et al., 2006; cf. Schnur, 2014) and aligns with two prior studies demonstrating a cumulative effect in neurotypical speakers' error rates (Navarrete et al., 2010; Rose & Abdel Rahman, 2017). Going beyond that work, the present documented effect of ordinal position in semantic errors constitutes direct evidence of incremental learning impacting the lexical-semantic stage of word retrieval. In this case, that other errors did not increase with ordinal position suggests the increase in semantic errors with ordinal position was not likely due to increasing fatigue, and thus naming failure, across the session.

We also found that semantic error rates increased as a function of increasing item similarity. In the Analysis II models, item similarity values corresponded to the current item being named. Thus, the straightforward interpretation of this result is that item similarity captures item-inherent differences in factor(s) that influence semantic error production (e.g., items with higher item similarity values may inhabit denser regions of semantic space and in turn elicit more semantic errors; e.g., Mirman, 2011; Vigliocco et al., 2004). This, along with the accrual of semantic errors across ordinal position, suggests that semantic error rates are influenced by both intrinsic semantic properties (i.e., current target's item similarity) *and* experimentally imposed competition due to having named from the category previously. However, whether item similarity values of preceding exemplars impact current trial naming performance remains to be clarified (cf. Rose & Abdel Rahman, 2017). To isolate potential semantic distance effects on incremental learning in the present design, it is necessary to focus on the first two category exemplars presented within a session. This is because similarity values of successive members of a category (i.e., within the 6-item sets) was uncontrolled, and it is not obvious in what way the item similarity of multiple prior exemplars may interact to affect performance in each successive ordinal position in a category. Thus, in Analysis III, we assessed whether the item similarity of the *first* category exemplar presented within a session (hereafter, *position 1*) impacts naming of the subsequent exemplar (hereafter, *position 2*). If so, this would suggest semantic distance modulates the learning-based mechanism underlying persistent semantic interference.

Analysis III: Follow-up test of semantic distance effect

In Analysis III, we restricted the data set to naming responses at ordinal position 2, with position 1 item similarity and position 2 item similarity reserved as fixed effects of theoretical interest. Position 2 item similarity was included as a predictor to account for possible (e.g., intrinsic) similarity influences other than item similarity of position 1. More importantly, if position 1 item similarity heightens semantic error rates at position 2, this suggests semantic distance interfaces with a learning mechanism to affect future exemplar retrieval.

Data included in Analysis III consisted of 1,102 responses (i.e., responses to related items at position 2), of which 135 were classified as semantic errors, 328 were classified as other errors, and 639 were classified as accurate (see Table 6). Covariates considered when fitting the base models reflected the values of position 2 items. Only position 2 name agreement significantly contributed to the base model of semantic errors. The base model of other errors included position 2 log-transformed word frequency, length, and visual complexity. To the base model, the fixed effects of position 2 item similarity and position 1 item similarity were sequentially entered. Both semantic and other error models included by-participant random intercepts. Fixed/random effects structures and model comparison results are reported in Table 8 (see Supplementary Material for model output).

Adding position 2 similarity to the semantic and other error base models did *not* significantly improve model fit, p 's > .16. For semantic errors, adding position 1 similarity to the model controlling for position 2 similarity did significantly improve model fit, $\chi^2(1) = 4.05$, $p = .044$, with increasing position 1 item similarity associated with increased

semantic error rates [$Estimate = -.21$, $SE = .11$, $p = .048$]. However, in contrast to semantic errors, position 1 item similarity did not significantly improve model fit of other errors, $p = .219$.

The findings from Analysis III provide support for the assumption that semantic similarity has an effect on future same-category naming, and thus incremental learning, as the item similarity of the first exemplar impacted semantic error rates on the second exemplar. Furthermore, aligning with the findings above, this effect was found only in semantic, and not other, errors, suggesting incremental learning and its modulation by semantic distance has a lexical-semantic locus. We discuss the theoretical implications of this and our other findings in the General Discussion.

General Discussion

The goal of this research was to investigate properties of semantic interference manifesting in specific error types to explore possible empirical parallels derived from chronometric semantic interference studies. In the present study, a group of PWA provided naming responses to a large corpus of pictures. The corpus comprised multiple categories with a range of exemplars, and several unclassifiable objects that formed an “unrelated” category. By analyzing rates of different error types, we assessed whether incremental learning (1) characterizes the way in which the production system errs in response to semantic context manipulations, and if so, whether such semantic interference (2) exerts its effects at a lexical-semantic level of processing. In an initial comparison of error rates for related versus unrelated items, Analysis I established an overall effect of semantic context in our unstructured version of the continuous naming task (see also Gordon & Cheimariou, 2013), with enhanced semantic (but not other) errors for related versus unrelated items. Analysis II revealed that within the related categories, semantic (and not other) error rates increased across ordinal position independent of lag. The accrual of semantic errors was significant after controlling for the similarity between an item and its category, a factor that may reflect intrinsic semantic properties shown to affect semantic error production on current trial responding, such as semantic neighborhood density. In Analysis III, we further explored semantic similarity effects as they may influence the degree to which incremental learning affects future same-category naming. Analysis III showed that greater item similarity of the first (and not the second) category exemplar presented within a session was associated with increased semantic error production when naming the second category exemplar. Together, these findings provide evidence that incremental learning operating at a lexical-semantic level underlies cumulative semantic interference manifesting in speech errors, and offer additional insight into the role of semantic distance in learning-based interference effects.

According to the computational framework put forth by Oppenheim et al. (2010), semantic interference results from experience-dependent changes in the semantic-to-lexical connection weights, which affects both the speed *and* accuracy of future naming from the category. In line with this account, we found that cumulative semantic interference manifests in speech error data, exhibiting characteristics parallel to those established in chronometric studies using the continuous naming task (e.g., Belke, 2013; Howard et al., 2006; Navarrete et al., 2010). Specifically, we find that error rates increased across ordinal position in a

manner that was insensitive to lag, consistent with chronometric evidence demonstrating persistent interference effects (e.g., Howard et al., 2006; cf. Schnur, 2014). Critically, our finding that the effect was specific to semantic, and not other, errors indicates that semantic interference arises when mapping from semantic to lexical representations. This aligns with a number of chronometric studies that have provided evidence indicative of a lexical-semantic locus of interference by manipulating factors that tap different levels of the production system (e.g., Navarrete et al., 2010; cf. Damian et al., 2001). For instance, semantic interference generally does *not* occur in tasks that tap the semantic level without lexical access (i.e., manually categorizing pictures; e.g., Damian et al., 2001; Belke, 2013; Riley et al., 2015; Vitkovitch & Humphreys, 1991, Experiment 2; but see Wei & Schnur, 2016 for evidence in associative picture matching) or those that tap the lexical level but retrieval is not semantically-driven (i.e., written word naming; Belke et al., 2005a; Kroll & Stewart, 1994; Vitkovitch & Humphreys, 1991, Experiment 3). These findings have motivated contemporary theoretical accounts of semantic interference that localize experience-dependent changes in lexical access to the lexical-semantic stage of retrieval (e.g., Howard et al., 2006; Oppenheim et al., 2010). However, focusing on the specific types of errors that pattern with the expected characteristics of semantic interference (i.e., persistent accrual of errors with successive naming from the category) represents a more direct measure of the stage at which interference exerts its effects in light of prominent accounts of aphasic error production (e.g., Dell et al., 1997; Schwartz et al., 2006). In turn, our results extend those of previous studies investigating semantic interference effects in latency data by providing more direct evidence that incremental learning within the lexical-semantic connection weights gives rise to semantic interference in naming.

Our interpretation that the incremental learning effects localized to the mapping between semantics and words hinges on claims about the origin of semantic naming errors in aphasia. However, as reported in two cases of deep dyslexia, semantic errors in oral naming in aphasia can arise from disrupted access to phonological forms for output (e.g., Caramazza & Hillis, 1990). Nevertheless, an array of evidence converges to a predominantly pre-phonological origin of semantic errors in naming in aphasia. First, semantic errors are attributed solely to disrupted semantics-to-words mapping in the two-step interactive model of lexical access (e.g., Dell et al., 1997; Schwartz et al., 2006), an influential, computationally explicit framework that has been successful in modeling individual PWA's naming accuracy and error response patterns in large case series investigations with diverse samples. Second, semantic and phonological naming errors in stroke aphasia have different neuroanatomical substrates, implying different functional loci. Semantic errors in naming are strongly associated with damage to the mid-to-anterior temporal lobe (Mirman, Zhang, Wang, Coslett, & Schwartz, 2015b; Schwartz et al., 2009; Walker et al., 2011). In contrast, phonological errors in word production are associated with damage to posterior superior temporal and frontoparietal regions (Buchsbaum et al., 2011; Fridriksson et al., 2016; Mirman et al., 2015a, 2015b; Schwartz, Faseyitan, Kim, & Coslett, 2012).

Arguments have been made that semantic naming errors in aphasia can arise from dysregulated or disrupted semantics prior to lexical retrieval (Hillis, Rapp, Romani, & Caramazza, 1990; Jefferies & Lambon Ralph, 2006). However, studies using voxel-based lesion-symptom mapping have shown that semantic errors in naming are associated with

damage to regions implicated in semantic cognition (i.e., mid-to-anterior temporal lobe) – even after accounting for verbal and non-verbal conceptual processing deficits (Schwartz et al., 2009, 2011; Walker et al., 2011). More importantly, our sample evinced generally intact core semantics and word comprehension abilities, arguing against a purely semantic locus of semantic errors in naming in our task.

That cumulative semantic interference manifested in semantic (and not other) errors may also provide an explanation for why the preponderance of continuous naming studies have failed to find an effect in speech errors (Belke, 2013; Howard et al., 2006; Riès et al., 2015; Runnqvist et al., 2012; Schnur, 2014; cf. Navarrete et al., 2010; Rose & Abdel Rahman, 2017). In the study of continuous naming in PWA, Riès et al. (2015) did not analyze specific error types, which may have obscured detecting a cumulative effect. Indeed, in the current study, the rate of errors overall (i.e., collapsed across error type) remained stable with successive same-category naming. This is because semantic errors increased across ordinal position, whereas other errors decreased, indicating that accuracy alone may not be a sensitive enough measure of semantic interference effects in naming. The specificity of semantic, and not other, errors in persistent interference effects has also been shown in neurotypical speakers (Vitkovitch & Humphreys, 1991). Specifically, Vitkovitch and Humphreys demonstrated that semantic errors in picture naming occurred at a greater than chance level under speeded deadline conditions (Experiment 1), and the rate with which semantic errors occurred increased if a semantically related versus unrelated picture had been named in a previous block of trials (Experiment 2). This, along with the finding that semantic error rates in PWA are not conflated with overall aphasia severity, suggests that the semantic error production in both populations reflect points along a continuum corresponding to the fidelity of mapping from semantics to lexical representations (Dell et al., 1997; Schwartz et al., 2006). Thus, parallel findings of long-lasting semantic interference manifesting specifically in semantic errors as observed here in PWA and prior studies of neurotypical speakers (see also Vigliocco et al., 2004) indicates that semantic interference effects in the unimpaired system are not qualitatively different from those observed in PWA – at least in the case of continuous naming (which may not be the case in blocked-cyclic naming; e.g., Belke & Stielow, 2013).

Another important finding in the current study concerns the impact of semantic similarity on subsequent naming from the category. To our knowledge, the present research and that of Rose and Abdel Rahman (2017) are the only studies to explore whether (and how) semantic similarity affects the accumulation of semantic interference in the continuous naming task. Rose and Abdel Rahman manipulated semantic distance in a binary fashion (i.e., basic-level versus superordinate-level categories), and found cumulative semantic interference (in latencies and errors) *only* for items from within a basic level category (e.g., primates: gorilla, orangutan, chimpanzee) and not those from different basic levels under the superordinate category (e.g., animals: shark, donkey, cobra). Based on these findings, we might conclude that the strength and/or number of competitors must reach some “threshold” to initiate the incremental learning mechanism that underlies persistent semantic interference effects. However, in the current study, the degree to which the first named exemplar from a given category was similar to other category exemplars predicted semantic (and not other) error rates for the next same-category exemplar presented for naming. This is consistent with a

view where semantic similarity exerts graded adjustments within the system, rather than having an “all or none” effect on semantic interference (for similar evidence in blocked-cyclic naming, see Navarrete et al., 2012; Vigliocco et al., 2002). While contemporary models of semantic interference in naming (Howard et al., 2006; Oppenheim et al., 2010) do not explicitly simulate semantic distance effects, the assumed property of shared activation by extension indicates that semantic similarity may represent the “relevance” of prior naming experience, and in turn directly influence the magnitude of experience-driven changes in lexical-semantic connections weights following a naming event (with more similar items inducing greater changes).

While the findings of this research support a learning-based account of semantic interference in error production, there are additional aspects of the results and limitations of the study that warrant discussion. First, in Analysis II, visual inspection of the increase in semantic error rates across ordinal position resembles a non-linear function (see Table 6), which may be different from the linear function of semantic interference accrual in latencies (e.g., Howard et al., 2006; Schnur, 2014). However, growth curve analysis (Mirman, 2014; also see Mirman, Dixon, & Magnuson, 2008), a multilevel modeling technique using third-order orthogonal polynomials, revealed that the pattern of accrual across ordinal position is significantly linear [$Estimate = .24$, $SE = .09$, $p = .01$] but not quadratic or cubic [$Estimate = -.02$, $SE = .09$, $p = .85$ and $Estimate = -.10$, $SE = .09$, $p = .29$, respectively]. In other studies demonstrating significant cumulative semantic interference in error rates (Navarrete et al., 2010; Rose & Abdel Rahman, 2017), the ordinal position effects also appear to be nonlinear (see Table 1 in Navarrete et al., 2010 and Table 1 in Rose & Abdel Rahman, 2017) but possible nonlinearities were not evaluated statistically in those studies. It may be of theoretical interest if the function of accrual of semantic interference in the continuous naming task differs for error rates versus latencies, but further research is required to address this issue.

Second, in Analysis III, it could be argued that the effect of semantic similarity of position 1 on semantic errors at position 2 is suspect because in Analysis II, the increase in semantic error rates from position 1 to 2 was minimal (i.e., .6% increase; Table 6). However, minimal increase in semantic error rates from position 1 to 2 does not preclude diagnosing a moderating variable (i.e., position 1 similarity) on semantic errors at position 2. Whereas the analysis approach in Analysis III was adopted to isolate the effect of similarity of position 1 on 2 in a controlled fashion, we acknowledge that it will be important in future work to manipulate the similarity of same-category exemplar sets systematically to provide a more direct test of how similarity influences the growth of interference, and thus incremental learning. Furthermore, it may also be informative to explore whether the semantic errors are in fact previously presented exemplars (i.e., perseveration errors; as in Hsiao et al., 2009), and that their incidence increases as a function of the preceding exemplar's similarity value. Finally, in the current study our measure of semantic similarity reflects the degree to which a given exemplar is similar to other exemplars in the category. In Analysis III, the similarity value of position 1 is assumed to capture the extent to which that exemplar may invoke the co-activation of other category exemplars (including the exemplar at position 2). We assume such category-wide co-activation produced the similarity effect that was observed. However, in future work, it will be important to explore whether the similarity between individual

exemplars impacts semantic interference differently than an exemplar's similarity to its category more generally.

Conclusion

The current research provides support for the view that incremental learning underlies semantic interference in speech error data in a manner that parallels evidence found in chronometric studies. Critically, that semantic interference manifested specifically in semantic errors is consistent with the assumption that the effect has its locus at the lexical-semantic stage of retrieval. Lastly, the results revealed that semantic similarity impacts the degree to which prior naming impacts subsequent naming from the category indicates that this factor modulates the extent of experience-dependent changes in semantic-to-lexical connections weight strengths. Together, these findings shed light on how incremental learning impacts lexical retrieval success, indicating that the degree of similarity between an item and with those that have yet to be named may represent the “relevance” of prior naming experience and in turn have greater consequences for future retrieval.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Dan Mirman, Gary Dell, and Myrna Schwartz for consultation on analysis, design and statistical approach. We also thank Adelyn Brecher, Maureen Gagliardi, and Kelly Garvey for assistance in coding data for this study.

Funding

This work was supported by National Institutes of Health (NIH) research grants [R01 DC015516 and R03 DC012426] and the Albert Einstein Society, Albert Einstein Healthcare Network, Philadelphia, PA awarded to Erica L. Middleton. Denise Y. Harvey was supported by the NIH [T32 HD071844].

References

- Alario FX, & Moscoso del Prado Martín F (2010). On the origin of the “cumulative semantic inhibition” effect. *Memory & Cognition*, 38(1), 57–66. [PubMed: 19966239]
- Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Barr DJ, Levy R, Scheepers C, & Tily HJ (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Belke E (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin & Review*, 15(2), 357–363. [PubMed: 18488652]
- Belke E (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *Journal of Memory and Language*, 69(3), 228–256. 10.1016/j.jml.2013.05.008
- Belke E, Brysbaert M, Meyer AS, & Ghyselinck M (2005a). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition*, 96, B45–B54. [PubMed: 15925568]
- Belke E, Meyer AS, & Damian MF (2005b). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology*, 58(4), 667–692. [PubMed: 16104101]

- Belke E, & Stielow A (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *The Quarterly Journal of Experimental Psychology*, 66(11), 2135–2160. 10.1080/17470218.2013.775318 [PubMed: 23521507]
- Biegler KA, Crowther JE, & Martin RC (2008). Consequences of an inhibition deficit for word production and comprehension: Evidence from the semantic blocking paradigm. *Cognitive Neuropsychology*, 25(4), 493–527. [PubMed: 19086200]
- Brodeur MB, Dionne-Dostie E, Montreuil T, & 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(5), e10773 10.1371/journal.pone.0010773 [PubMed: 20532245]
- Brown AS (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 204–215.
- Brysbaert M, & New B (2009). Moving beyond Ku era and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990. [PubMed: 19897807]
- Buchsbaum BR, Baldo J, Okada K, Berman KF, Dronkers N, D’Esposito M, & Hickok G (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory—an aggregate analysis of lesion and fMRI data. *Brain and Language*, 119(3), 119–128. [PubMed: 21256582]
- Caramazza A (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14(1), 177–208.
- Caramazza A, & Hillis AE (1990). Where do semantic errors come from? *Cortex*, 26(1), 95–122. [PubMed: 2354648]
- Chen Q, Middleton EL, & Mirman D (2018). Words fail: Lesion-symptom mapping of errors of omission in post-stroke aphasia. *Journal of Neuropsychology* 10.1111/jnp.12148
- Crowther JE, & Martin RC (2014). Lexical selection in the semantically blocked cyclic naming task: The role of cognitive control and learning. *Frontiers in Human Neuroscience*, 8(9). 10.3389/fnhum.2014.00009
- Damian MF & Als LC (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1372–1384.
- Damian MF, Vigliocco G, & Levelt WJM (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, 81(3), B77–B86. 10.1016/S0010-0277(01)00135-4 [PubMed: 11483172]
- Dell GS (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 27(2), 124–142.
- Dell GS, Schwartz MF, Martin N, Saffran EM, & Gagnon DA (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801–838. [PubMed: 9337631]
- Goldrick M, & Rapp B (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, 102(2), 219–260. [PubMed: 16483561]
- Gordon JK, & Cheimariou S (2013). Semantic interference in a randomized naming task: Effects of age, order, and category. *Cognitive Neuropsychology*, 30(7–8), 476–494. [PubMed: 24499271]
- Fridriksson J, Yourganov G, Bonilha L, Basilakos A, Den Ouden DB, & Rorden C (2016). Revealing the dual streams of speech processing. *Proceedings of the National Academy of Sciences*, 113(52), 15108–15113.
- Harvey DY, & Schnur TT (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. [PubMed: 25880795]
- Harvey DY, & Schnur TT (2016). Different loci of semantic interference in picture naming vs. word-picture matching tasks. *Frontiers in Psychology*, 7(710). 10.3389/fpsyg.2016.00710
- Hillis AE, Rapp B, Romani C, & Caramazza A (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, 7(3), 191–243.
- Howard D, Nickels L, Coltheart M, & Cole-Virtue J (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, 100(3), 464–482. 10.1016/j.cognition.2005.02.006 [PubMed: 16413014]

- Howard D, & Patterson K (1992). *Pyramids and palm trees: A test of semantic access from pictures and words* Bury St Edmunds, UK: Thames Valley Test.
- Hsiao EY, Schwartz MF, Schnur TT, & Dell GS (2009). Temporal characteristics of semantic perseverations induced by blocked-cyclic picture naming. *Brain and Language*, 108(3), 133–144. [PubMed: 19138794]
- Hughes JW, & Schnur TT (2017). Facilitation and interference in naming: A consequence of the same learning process? *Cognition*, 165, 61–72. [PubMed: 28501548]
- Jefferies E, & Lambon Ralph MA (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, 129(8), 2132–2147. [PubMed: 16815878]
- Kertesz A (1982). *Western aphasia battery* New York: Grune and Stratton.
- Kroll JF & Stewart E (1994). Category interferences in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174. 10.1006/jmla.1994.1008
- Levelt WJM, Roelofs A, & Meyer AS (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75. 10.1017/S0140525X99001776 [PubMed: 11301520]
- Mahon BZ, Costa A, Peterson R, Vargas KA, & Caramazza A (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 503–535.
- McCarthy RA, & Kartsounis LD (2000). Wobbly words: Refractory anomia with preserved semantics. *Neurocase*, 6(6), 487–497.
- Middleton EL, Schwartz MF, Rawson KA, & Garvey K (2015). Test-enhanced learning versus errorless learning in aphasia rehabilitation: Testing competing psychological principles. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(4), 1253–1261.
- Middleton EL, Traut HJ, & Stabile M (2018). The influence of semantic context on naming treatment in aphasia Manuscript in preparation.
- Mirman D (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective, and Behavioral Neuroscience*, 11(1), 32–43.
- Mirman D (2014). *Growth Curve Analysis and Visualization Using R* Boca Raton, FL: Chapman and Hall / CRC Press.
- Mirman D, Chen Q, Zhang Y, Wang Z, Faseyitan OK, Coslett HB, & Schwartz MF (2015a). Neural organization of spoken language revealed by lesion–symptom mapping. *Nature Communications*, 6, 6762 10.1038/ncomms7762
- Mirman D, Dixon JA, & Magnuson JS (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4), 475–494. [PubMed: 19060958]
- Mirman D, & Graziano KM (2013). The neural basis of inhibitory effects of semantic and phonological neighbors in spoken word production. *Journal of Cognitive Neuroscience*, 25(9), 1504–1516. [PubMed: 23647518]
- Mirman D, Zhang Y, Wang Z, Coslett HB, & Schwartz MF (2015b). The ins and outs of meaning: Behavioral and neuroanatomical dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. *Neuropsychologia*, 76, 208–219. [PubMed: 25681739]
- Navarrete E, Del Prato P, & Mahon BZ (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology*, 3(38). 10.3389/fpsyg.2012.00038
- Navarrete E, Del Prato P, Peressotti F, & Mahon BZ (2014). Lexical selection is not by competition: Evidence from the blocked naming paradigm. *Journal of Memory and Language*, 76, 253–272. [PubMed: 25284954]
- Navarrete E, Mahon BZ, & Caramazza A (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica* 134(3), 279–289. 10.1016/j.actpsy.2010.02.009 [PubMed: 20347062]
- Novick JM, Trueswell JC, & Thompson-Schill SL (2010). Broca's Area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, 10, 906–924.

- Oppenheim GM, Dell GS, & Schwartz MF (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, 114(2), 227–252. 10.1016/j.cognition.2009.09.007 [PubMed: 19854436]
- Rabovsky M, Schad DJ, & Abdel Rahman R (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, 146, 240–244. [PubMed: 26468758]
- Rapp B, & Goldrick M (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, 107(3), 460–499. [PubMed: 10941277]
- Riès SK, Karzmark CR, Navarrete E, Knight RT, & Dronkers NF (2015). Specifying the role of the left prefrontal cortex in word selection. *Brain and Language*, 149, 135–147. [PubMed: 26291289]
- Riley E, McMahon KL, & de Zubicaray G (2015). Long-lasting semantic interference effects in object naming are not necessarily conceptually mediated. *Frontiers in Psychology*, 6(578). 10.3389/fpsyg.2015.00578
- Roach A, Schwartz MF, Martin N, Grewal RS, & Brecher A (1996). The Philadelphia naming test: Scoring and rationale. *Clinical Aphasiology*, 24, 121–133.
- Rose SB, & Abdel Rahman R (2016). Cumulative semantic interference for associative relations in language production. *Cognition*, 152, 20–31. [PubMed: 27015349]
- Rose SB, & Abdel Rahman R (2017). Semantic similarity promotes interference in the continuous naming paradigm: Behavioural and electrophysiological evidence. *Language, Cognition and Neuroscience*, 32(1), 55–68.
- Runnqvist E, Strijkers K, Alario FX, & Costa A (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language*, 66(4), 850–869. 10.1016/j.jml.2012.02.007
- Schnur T (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language*, 75, 27–44. 10.1016/j.jml.2014.04.006
- Schnur TT, Schwartz MF, Brecher A, & Hodgson C (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, 54(2), 199–227.
- Schnur TT, Schwartz MF, Kimberg DY, Hirschorn E, Coslett HB, & Thompson-Schill SL (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences USA*, 106, 322–327.
- Schwartz MF, Dell GS, Martin N, Gahl S, & Sobel P (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, 54(2), 228–264.
- Schwartz MF, Faseyitan O, Kim J, & Coslett HB (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, 135(12), 3799–3814. [PubMed: 23171662]
- Schwartz MF, Kimberg DY, Walker GM, Brecher A, Faseyitan O, Dell GS, ... Coslett HB (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences USA*, 108(20), 8520–8524.
- Schwartz MF, Kimberg DY, Walker GM, Faseyitan O, Brecher A, Dell GS, & Coslett HB (2009). Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia. *Brain*, 132(12), 3411–3427. [PubMed: 19942676]
- Szekely A, Jacobsen T, D'Amico S, Devescovi A, Andonova E, Herron D, ... Bates E (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247–250. 10.1016/j.jml.2004.03.002 [PubMed: 23002322]
- Van Overschelde JP, Rawson KA, & Dunlosky J (2004). Category norms: An updated and expanded version of the norms. *Journal of Memory and Language*, 50(3), 289–335.
- Vigliocco G, Vinson DP, Damian MF, & Levelt W (2002). Semantic distance effects on object and action naming. *Cognition*, 85(3), B61–B69. [PubMed: 12169413]
- Vigliocco G, Vinson DP, Lewis W, & Garrett MF (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48(4), 422–488. 10.1016/j.cogpsych.2003.09.001 [PubMed: 15099798]
- Vitkovitch M, & Humphreys GW (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology*, 17(4), 664–680. 10.1037/0278-7393.17.4.664

- Walker GM, Schwartz MF, Kimberg DY, Faseyitan O, Brecher A, Dell GS, & Coslett HB (2011). Support for anterior temporal involvement in semantic error production in aphasia: New evidence from VLSM. *Brain and Language*, 117(3), 110–122. [PubMed: 20961612]
- Wei T, & Schnur TT (2016). Long-term interference at the semantic level: Evidence from blocked-cyclic picture matching. *Journal of Experimental Psychology*, 42(1), 149–157. 10.1037/xlm0000164 [PubMed: 26213834]
- Wheeldon LR, & Monsell S (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language*, 33(3), 332–356
- Wilshire CE, & McCarthy RA (2002). Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology*, 19(2), 165–186. [PubMed: 20957536]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1. Predicted (lines) and actual (dots) probabilities for semantic errors (black) and other errors (gray) as a function of ordinal position within a category. Error bars represent 95% confidence intervals.

Table 1.

Participant Demographics and Neuropsychological Characteristics.

Participant	Age	Years Ed	Gender	MPO	WAB AQ	Aphasia Subtype	Speech Apraxia	PNT Acc	Nonverbal Comp	Word Comp	Word Rep
1	55	16	M	10	75.5	TCM	None	59	85	90	96
2	77	14	F	70	83.1	C	Mild	65	98	90	80
3	50	14	M	35	65	B	Mild	55	90	98	97
4	64	16	M	136	56.8	B	Mild	59	90	81	90
5	67	12	F	69	92.9	A	None	78	88	98	94
6	59	15	M	58	88.2	A	None	79	87	93	94
7	72	12	F	20	73.2	TCM	Mild/Mod	77	92	94	95
8	55	18	M	75	55.6	B	Mild/Mod	78	87	96	100
9	61	10	F	12	88.7	A	None	79	96	86	91
10	56	12	M	99	71.2	C	None	70	94	98	94
11	66	16	M	145	48.4	B	Mod	43	92	90	91
12	62	19	M	8	84.7	A	None	83	92	95	98
13	49	14	M	41	73.7	A	None	83	81	95	95
14	55	16	M	10	76.7	A	None	64	96	92	87
15	56	21	M	53	89.8	A	None	90	96	95	99
Mean	60	15		56.1	74.9			71	91	93	93
Min,Max	49,77	10,21		8,145	48.4,92.9			43,90	81,98	81,98	80,100

Note. MPO = months post-stroke onset; WAB AQ = Western Aphasia Battery Aphasia Quotient, a measure of aphasia severity (Kertesz, 1982); Aphasia Subtype, where A = anomic; B = Broca's; C = conduction; TCM = transcortical motor; PNT = Philadelphia Naming Test (Roach et al., 1996) performance, where Acc = accuracy in percentages; Nonverbal Comp = an associative picture-picture matching task of nonverbal comprehension, in percentages (Howard & Patterson, 1992); Word Comp = a spoken word-picture verification task of word comprehension, in percentages (Roach et al., 1996); Word Rep = a test of immediate word repetition, in percentages (Philadelphia Repetition Test; Dell et al., 1997).

Table 2.

Characteristics of the 615-Item Picture Corpus.

Variable	615-Corpus Mean (SD)	Related Items Mean (SD)	Unrelated Items Mean (SD)
Name agreement	.93 (.06)	0.93 (.06)	0.92 (.06)
Log frequency/million	1.06 (0.57)	1.01 (.60)	0.95 (.59)
Visual complexity [†]	2.68 (0.76)	2.68 (.76)	2.68 (.74)
# of phonemes	5.20 (1.99)	5.22 (1.97)	5.06 (2.12)

Note.

[†] Five-point scale where 1 = “image is very simple” to 5 = “image is very complex.” # of phonemes = length. Standard deviation (SD) in parentheses.

Table 3.

Categories in the 615-item Corpus with Mean and Standard Deviation (SD) Item Similarity Norms and Sample Exemplars for the Related Categories.

Category	Typicality Norm [†]	Mean	SD	# of Items
Accessories (earring, purse, cast)	2.26	0.77	24	
Body Parts (finger, toes, mustache)	1.55	0.52	25	
Clothing (underwear, jacket, pocket)	1.70	0.49	23	
Food (bagel, cookie, tea bag)	1.87	0.40	37	
Fruits & Vegetables (lime, carrot, peanut)	1.50	0.34	34	
Furnishings (dresser, picture, barrel)	2.28	0.86	27	
Kitchen Items (fork, jar, bottle cap)	2.18	0.68	31	
Mammals (elephant, skunk, unicorn)	1.49	0.43	32	
Musical Instrument (piano, banjo, jukebox)	1.80	1.05	18	
Nature (lightning, mountain, smoke)	1.93	0.43	32	
Non Mammal (owl, octopus, dragon)	2.36	0.72	33	
Office Supplies (pen, eraser, magazines)	2.06	0.66	30	
Parts of Buildings (attic, sink, mailbox)	1.84	0.46	22	
Structure (lighthouse, airport, Africa)	2.27	0.70	25	
Toiletries (toothpaste, tweezers, needle)	1.81	0.66	16	
Tools & Hardware (hammer, flashlight, cage)	2.00	0.65	35	
Toys & Games (baseball, yoyo, trophy)	2.06	0.68	33	
Types of People (dentist, cowboy, mummy)	2.13	1.02	37	
Vehicles (bus, unicycle, roller coaster)	1.88	0.59	23	
Unrelated	-	-	78	

Note.

[†]Five-point scale where 1 = “exemplar is very similar” to 5 = “exemplar is very dissimilar.” Sample exemplars are drawn from the distribution of item similarity ratings (maximum, median, and minimum) per category, where “maximum” and “minimum” refer to most and least similar exemplars, respectively.

Table 4.

Counts and Percentages per Response Type as a Function of Category Type.

Category Type	Response Type			
	Semantic Error	Other Error	Accurate	Total
Related	866 (13.2%)	1903 (29%)	3797 (57.8%)	6566
Unrelated	227 (9.7%)	834 (35.8%)	1268 (54.4%)	2329

Table 5.

Analysis I Model Comparison Results for Semantic Errors (left panel) and Other Errors (right panel).

Semantic Errors							Other Errors						
Model		logLik	dev	Chisq	df	p-value	Model		logLik	dev	Chisq	df	p-value
Error ~ Freq + NA + Vis Comp + (1 + Cat Type Subject)		-3173.0	6346.0				Error ~ Freq + NA + Len + Vis Comp + (1 + Cat Type Subject)		-4899.8	9799.7			
Error ~ Cat Type + Freq + NA + Vis Comp + (1 + Cat Type Subject)		-3168.8	6337.6	8.36	1	.004	Error ~ Cat Type + Freq + NA + Len + Vis Comp + (1 + Cat Type Subject)		-4892.1	9784.2	15.45	1	<.001

Note. The first row of each panel illustrates the covariates included in the respective base models, with subsequent rows illustrating the model comparison results after adding the fixed effect of interest highlighted in bold. Cat Type = category type; Chisq = chi-squared; dev = deviance; df = degrees of freedom; Freq = frequency; logLik = log-likelihood; Len = length; NA = name agreement; Vis Comp = visual complexity; (1 + Cat Type | Subject) = random effects structure representing the inclusion of a by-participant random intercept and slope for the effect of category type.

Table 6.

Counts and Percentages per Response Type for the First 6 Ordinal Positions within a Category and the Mean (SD) Item Similarity and Lag between Successive Categorized Items.

Response Type	Ordinal Position within Category					
	1	2	3	4	5	6
Semantic Error	129 (11.7%)	135 (12.3%)	128 (11.6%)	162 (14.8%)	156 (14.4%)	156 (14.4%)
Other Error	309 (28.0%)	328 (29.8%)	333 (30.3%)	306 (27.9%)	328 (30.2%)	299 (27.7%)
Accurate	664 (60.3%)	639 (58.0%)	639 (58.1%)	628 (57.3%)	602 (55.4%)	625 (57.9%)
Total	1102	1102	1100	1096	1086	1080
Mean (SD) Item Similarity	1.92 (0.32)	2.00 (0.34)	1.93 (0.32)	1.95 (0.31)	1.87 (0.31)	1.97 (0.29)
Mean (SD) Lag		21.5 (22.7)	20.2 (19.9)	20.7 (22.2)	21.2 (21.9)	19.9 (20.1)

Table 7.

Analysis II Model Comparison Results for Semantic Errors (left panel) and Other Errors (right panel).

Semantic Errors						Other Errors					
Model	logLik	dev	Chisq	df	p-value	Model	logLik	dev	Chisq	df	p-value
Ordinal Position (1 through 6)											
Error ~ Freq + NA + Vis Comp + (1 Subject)	-2449.2	4898.3				Error ~ Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-3543.7	7087.3			
Error ~ Sim + Freq + NA + Vis Comp + (1 Subject)	-2445.3	4890.6	7.70	1	.006	Error ~ Sim + Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-3531.5	7063.0	24.35	1	<.001
Error ~ Pos + Sim + Freq + NA + Vis Comp + (1 Subject)	-2442.0	4884.0	6.64	1	.010	Error ~ Pos + Sim + Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-3527.4	7054.9	8.11	1	.004
Ordinal Position (2 through 6)											
Error ~ Freq + NA + Vis Comp + (1 Subject)	-2070.1	4140.2				Error ~ Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-2964.5	5929.0			
Error ~ Sim + Freq + NA + Vis Comp + (1 Subject)	-2067.3	4134.5	5.65	1	.017	Error ~ Sim + Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-2954.2	5908.4	20.59	1	<.001
Error ~ Pos + Sim + Freq + NA + Vis Comp + (1 Subject)	-2065.1	4130.1	4.40	1	.036	Error ~ Pos + Sim + Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-2951.6	5903.1	5.27	1	.022
Error ~ Lag + Pos + Sim + Freq + NA + Vis Comp + (1 Subject)	-2065.0	4129.9	0.21	1	.648	Error ~ Lag + Pos + Sim + Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-2950.8	5901.6	1.47	1	.226
Error ~ Lag*Pos + Sim + Freq + NA + Vis Comp + (1 Subject)	-2064.5	4128.9	1.22	2	.545	Error ~ Lag*Pos + Sim + Freq + NA + Len + Vis Comp + Trial + (1 Subject)	-2949.4	5898.7	4.39	2	0.111

Note. The first row of each panel illustrates the covariates included in the respective base models, with subsequent rows illustrating the model comparison results after adding the fixed effect of interest highlighted in bold. Chisq = chi-squared; dev = deviance; df = degrees of freedom; Freq = frequency; Len = length; logLik = log-likelihood; NA = name agreement; Pos = ordinal position; Sim = item similarity; Trial = trial within a single session; Vis Comp = visual complexity; (1 | Subject) = random effects structure representing the inclusion of a by-participant random intercept.

Table 8. Analysis III Model Comparison Results for Semantic Errors (left panel) and Other Errors (right panel).

Model	Semantic Errors					Other Errors					
	logLik	dev	Chisq	df	p-value	Model	logLik	dev	Chisq	df	p-value
P2 Error ~ P2 NA + (1 Subject)	-399.0	798.0				P2 Error ~ P2 Freq + P2 Len + P2 Vis Comp + (1 Subject)	-611.5	1222.9			
P2 Error ~ P2 Sim + P2 NA + (1 Subject)	-399.0	797.9	0.05	1	.829	P2 Error ~ P2 Sim + P2 Freq + P2 Len + P2 Vis Comp + (1 Subject)	-610.5	1221.0	1.91	1	.167
P2 Error ~ P1 Sim + P2 Sim + P2 NA + (1 Subject)	-396.9	793.9	4.05	1	.044	P2 Error ~ P1 Sim + P2 Sim + P2 Freq + P2 Len + P2 Vis Comp + (1 Subject)	-609.8	1219.5	1.51	1	.219

Note. The first row of each panel illustrates the covariates included in the respective base models, with subsequent rows illustrating the model comparison results after adding the fixed effect of interest highlighted in bold. *Chisq* = chi-squared; *dev* = deviance; *df* = degrees of freedom; **P1** = position 1; **P2** = position 2; *Freq* = frequency; *Len* = length; *logLik* = log-likelihood; *NA* = name agreement; *Sim* = item similarity; *Vis Comp* = visual complexity; (1| Subject) = random effects structure representing the inclusion of a by-participant random intercept.