Journal of Memory and Language 140 (2025) 104570

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Retrieval-induced semantic interference

Channing E. Hambric [a,*], Pádraig G. O'Séaghdha [a,b]

[a] Department of Psychology, Lehigh University, Bethlehem, PA, USA
[b] Cognitive Science Program, Lehigh University, Bethlehem, PA, USA

ARTICLE INFO

Keywords:
Semantic interference
Retrieval-induced forgetting
Adaptive learning
Semantic Memory
Language Production
Lexical access

ABSTRACT

The long-term negative effect of semantic retrieval on the subsequent accessibility of related material has been extensively studied in separate memory and language production literatures. Though ostensibly studying the same phenomenon, these literatures have remained separated by different framings and methodologies. We argue for integration of the two research streams in an adaptive learning perspective and present a bridging experiment as a proof of concept of this approach. The experiment implemented a multiphase *retrieval-induced forgetting* (RIF) design (with generation and memory assessment phases) in combination with the use of naming latency measures and the temporal analysis of interference featured in language production research. The generation phase, typically unanalyzed in the memory literature, examined generation time for category-stem completions as a function of ordinal positions of related items. There was strong cumulative interference in generation latencies in the first pass through the structured list, showing that memory is already affected in this phase. After a retention interval, accessibility of new items from previously activated categories, and unactivated controls, was assessed using continuous picture naming rather than aggregate memory measures. Crucially, there was a picture naming cost to previously activated (but not generated) category members relative to the control condition, a RIF effect. This cost was supervenient on new cumulative interference and was evident only in the beginning of the assessment phase, underlining the value of the positional analyses. The findings add important detailing to the processes underlying retrieval-induced costs in memory research while also showing that retrieval-induced semantic interference transfers from stem-completion to picture naming retrieval tasks. This format-independence is consistent with a conceptual basis of semantic interference but does not preclude a locus of adaptive learning in conceptual-lexical links. Overall, we show that the memory and language production fields indeed provide different but complementary perspectives on the same semantic interference phenomenon. Combining the fields promises to be productive.

Introduction

The compartmentalization of scientific knowledge is so often lamented that it could be concluded that this seemingly undesirable property is an inevitable concomitant of how focused scientific discovery operates (see Merton, 1961). But when the overlap between separate fields of investigation becomes evident, integration of the complementary explorations is called for. This article examines a particular case, long-term semantic interference resulting from generative activity, where parallel research efforts in language production and memory retrieval research sectors have remained stubbornly separate despite their seeming complementarity. Although this compartmentalization has not prevented important and fundamental discoveries in each field, we argue that separation is no longer justifiable. The segregated discoveries are overdue for integration into a larger memory-language framework that accommodates the discoveries of each subfield, aligns their frameworks, and points to future synthetic directions. We present a proof-of-concept experiment that adopts language production procedures in a multi-phase *retrieval induced forgetting* design in support of the case for an integrated adaptive learning approach to semantic interference.

Semantic interference has been studied independently in subfields of memory and language production for decades using framings and procedures that obscure the underlying commonality between the fields. Semantic activation is principally facilitatory, such that concepts and words that are related to generated or comprehended meanings are coactivated through spreading activation processes (e.g., Collins & Loftus, 1975; Neely, 1977). Although simple priming paradigms show

* Corresponding author.
E-mail address: channinghambric@gmail.com (C.E. Hambric).

https://doi.org/10.1016/j.jml.2024.104570
Received 6 February 2024; Received in revised form 6 September 2024; Accepted 11 September 2024
Available online 17 September 2024

that facilitation from such coactivation is fleeting, other work has shown that semantic coactivation can induce long-term facilitation under certain conditions (e.g., Joordens & Becker, 1997; Becker et al., 1997), or selectively strengthen the accessibility of particular conceptual-lexical mappings (Rodd et al., 2004; Rodd et al., 2013; Rodd et al., 2017). These findings suggest that strengthened access in comprehension is the result of long-term learning mechanisms. Crucially, in circumstances involving meaning-based generation, long-term adaptation results in both strengthening of selected meanings (or their lexical connections) and downregulation of unselected, coactivated meanings (Norman et al., 2007; Oppenheim et al., 2010). Specifically, generation strengthens access to retrieved meanings and induces interference to related material in tasks such as memory retrieval or picture naming in which a single option is selected from among multiple coactivated conceptual-lexical alternatives (e.g., Anderson et al., 1994; Norman et al., 2007; Schnur et al., 2009; Oppenheim et al., 2010). Separate memory (retrieval-induced forgetting) and language production (especially continuous picture naming) literatures show that the accessibility of conceptual-semantic or linked lexical representations varies as a function of recent retrieval experience (e.g., Brown, 1981; Johnson & Anderson, 2004; Howard et al., 2006; Norman et al., 2007; Schnur et al., 2009; Oppenheim et al., 2010; Anderson & Hulbert, 2021). Although a number of language production researchers have pointed to parallels between the two fields (e.g., Oppenheim et al., 2010; de Zubicaray et al., 2015; Navarrete et al., 2021), and both fields ostensibly describe modulations within conceptual-lexical networks, they focus on different aspects of the problem space and remain largely segregated. In the memory literature, accounts of retrieval-induced forgetting address an undifferentiated semantic memory stratum and generally omit consideration of lexical retrieval and selection, though these are clearly implicated in the verbal tasks that experiments employ (Anderson et al., 1994; Jonker et al., 2013; Norman et al., 2007). One influential account explicitly postulates adaptations within this memory stratum as the basis of semantic interference (Norman et al., 2007). Conversely, language production researchers for the most part ascribe long-term semantic interference in continuous naming to weakened links between conceptual representations and pre-phonological words (*lemmas*), with some researchers proposing that the links are to higher level *lexical concepts* (Belke, 2013; Doring et al., 2022). Regardless of whether conceptual features link to lexical concepts or to lemmas, these language models rely on oversimplified or underspecified conceptual-semantic representations and do not address the possibility of the dynamic changes within them that are postulated in some memory accounts. In our view, there is no strong reason to exclude either possibility. Based on current evidence, semantic interference may entail dynamic changes within conceptual representations (e.g., Norman et al., 2007), adaptations to conceptual-lexical links (e.g., Oppenheim et al., 2010), or both. Importantly, in relation to the theme of the special issue on the language-memory interface, we thus do not view memory and language as separate spheres but rather as intrinsic components of a conceptual-linguistic system in which long-term conceptual memory representations are often evoked through language. We now consider aspects of the relevant language production and memory literatures in more detail.

Semantic interference in language production

In language production research, production of words is often studied through timed picture naming tasks that require generation of names from memory. Pictures are assumed to activate conceptual knowledge and corresponding sectors of the lexicon from which the most felicitous, pragmatically fitting name is selected. Picture naming is thus a proxy for the general process of conceptually-driven word retrieval. When speakers produce an individual picture name, they first activate the concept they wish to convey. Semantic concepts, often construed as distributed semantic features (see Kumar, 2021; Lambon-Ralph et al., 2017 for reviews), link to corresponding abstract lexical representations, such as lexical concepts and lemmas (e.g., Levelt et al., 1999). As a result, words that share semantic features become coactive as activation spreads throughout the conceptual-lexical network (e.g., Dell & O'Seaghdha, 1992; Levelt et al., 1999; Rapp & Goldrick, 2000). This coactivation is inherently facilitatory: It strengthens later access to selected targets, and in some circumstances, such as naming pictures of members of the same category in immediate succession, it results in short-term facilitation of responses (Neely, 1977; Lupker, 1988; Navarrete et al., 2014; Scaltritti et al., 2017). However, in other circumstances such as continuous picture naming where related items are not contiguous, later access to unselected coactivated words is impaired (Howard et al., 2006; Schnur et al., 2009). The *semantic interference effect* in language production is thus the phenomenon in which it takes longer to retrieve and produce words after retrieving and producing related words (e.g., Vitkovitch & Humphries, 1993; Wheeldon & Monsell, 1994; Damian et al., 2001; Howard et al., 2006; Oppenheim et al., 2010; Abdel Rahman & Melinger, 2011; Schnur, 2014; Gordon et al., 2018).

Semantic interference is observed in picture-word procedures (e.g., Schriefers et al., 1990) where printed distractor words are superimposed on pictures, and in cyclic naming, where participants repeatedly name small sets of related or unrelated pictures in rapid succession (e.g., Damian et al., 2001). Critically, interference is also observed in the continuous picture naming paradigm where there are substantial, constructed lags between related items. In this paradigm, following an initial familiarization procedure, participants name pictures in a single stream that has been internally structured so that related items are dispersed throughout the sequence (Howard et al., 2006). This means that speakers are generally not aware of the manipulations, that short term facilitation between related items is absent, and that the cumulative negative effects of ordinal position within sequences are plausibly due to long-term adaptive mechanisms rather than to momentary selection difficulty (Howard et al., 2006; Oppenheim et al., 2010). We focus on this paradigm in the current work because it lends itself most naturally to fusion with memory procedures and because the picture-word and cyclic naming procedures may involve strategies and control processes that complicate data interpretation (Mahon et al., 2007; Belke & Stielow, 2013; Navarette et al., 2014; Hambric & O'Séaghdha, 2023). In the continuous naming task, experiments reliably show a linear increase in naming time over the ordinal positions of related items, such that the time it takes to name a picture increases substantially with each ordinal position (Howard et al., 2006; Oppenheim et al., 2010; Schnur, 2014). Given that interference increases with each related production, this pattern of semantic interference is referred to as *cumulative interference.*

What drives cumulative semantic interference? In some accounts, interference is attributed to an increase in the concurrent activation of previously selected lexical concepts at the moment of current target selection (e.g., Roelofs, 2018; Abdel Rahman & Melinger, 2019). To account for the long-term effects in the continuous naming paradigm, the heightened accessibility of recently selected lexical concepts was implemented as a bias term in WEAVER++ model simulations by Roelofs (2018). The bias is temporary and needs to be maintained by frequent reuse within a picture naming procedure (Roelofs, 2018). In contrast, and of particular interest here, adaptive learning accounts hold that word accessibility is shaped by an error-driven learning mechanism (Howard et al., 2006; Oppenheim et al., 2010). According to the earlier account of Howard et al. (2006), cumulative interference is the result of modulation to the mappings between concepts and lemmas. When a word is selected for production, the links between the target's conceptual representation and lemma are strengthened, thus making this word more competitive when a relative is later generated. That is, long-term strengthening of previous targets, rather than bias, is at the root of cumulative interference. Crucially, Oppenheim et al.'s (2010) "Dark Side" model extends the Howard et al. (2006) account by proposing that whereas links to selected words are strengthened, links to activated but unselected words are also weakened during name generation.

Specifically, this simple two-layer, feed-forward model has a single learning algorithm with two "sides": one that increases the connection weights from active semantic features to selected words, and one that decreases weights from shared features to unselected coactivated words. In sum, according to this model, a single learning mechanism strengthens access to chosen targets and weakens access to unselected competitors (Oppenheim et al., 2010). In this view, selection difficulty arises due to previous downregulation of semantic competitors and not primarily from increased competition from strengthened relatives. Indeed, though strengthened access to previously generated words occurs, Oppenheim et al. show through simulations that it may not play a role in long-term semantic interference (see Oppenheim & Nozari, 2024; Oppenheim, 2024 for additional discussion). Thus a key claim of Oppenheim et al. (2010) is that since competitive lexical selection is not necessary to capture semantic interference in continuous picture naming, long-term interference is primarily driven by adaptive learning processes. We note here that in making this claim, Oppenheim et al. (2010) pointed to the general parallels between semantic interference in language production and retrieval-induced forgetting and to the correspondences between theoretical debates in the two fields. As discussed below, in the field of retrieval-induced forgetting, some researchers posit that forgetting may be due to strengthened generated items "occluding" retrieval of ungenerated relatives, while others attribute it to mechanisms that directly affect the representations of semantically related activated items during retrieval (see Anderson, 2020, for review).

Returning to the broader characterization of language production accounts of semantic interference, models such as that of Oppenheim et al. (2010) focus on the accessibility and selection of lexical representations without consideration of processes internal to the level of conceptual representation (see Oppenheim & Nozari, 2024, for a related point). The work by Abdel Rahman & Melinger (2011), which posits that relatedness of thematically related concepts is construed "on the fly" through activation of situational knowledge, may be considered an exception. But, to our knowledge, no current account considers the possibility that change *within* the conceptual representations themselves either drives or contributes to interference phenomena. In the next section, we will review an influential account of semantic interference in memory (Norman et al., 2007) in which modulations within a conceptual (semantic memory) stratum drive long-term changes in semantic accessibility.

Retrieval-induced forgetting in semantic memory

Retrieval-induced forgetting (RIF) is the phenomenon whereby retrieval of material from memory leads to "forgetting," or later retrieval difficulty, for semantically related material (e.g., Anderson et al., 1994). This definition makes the affinity with cumulative semantic interference in language production evident, but also instigates a different framing. The retrieval-induced forgetting literature disregards language processes, effectively treating them as mere transducers of underlying memory operations. Related to this disregard, the retrieval processes that lead to forgetting are construed as *generation* but the how and what of generation are largely left unspecified (e.g., Norman et al. 2007; Ritvo et al., 2019). Moreover, the term forgetting itself may be a misnomer in that it suggests a categorical process instead of the continuous modulation of accessibility revealed in the language production literature.

These characteristics are exemplified in the standard tool of retrieval-induced forgetting research, the *retrieval practice* paradigm (Anderson et al., 1994). In this design, participants are familiarized with an extensive list of category-exemplar pairs (e.g., FRUIT-apple…. INSTRUMENT-violin…. FRUIT-orange) with multiple exemplars in each of several categories. In a subsequent Generation phase (usually referred to as the *retrieval practice phase* in the memory literature), some items from some of the familiarized categories are generated using category-stem completion cues (e.g., FRUIT-ap___). Generation is repeated several times to increase the rate of successful retrieval and potentially the strength of interference. Participants then engage in a distractor task (typically 15–20 min) before completing a final memory assessment phase involving all items (typically also using category-stem or category cues). There are two key conditions in this phase.[1] The Activated condition contains new members of the same categories that were generated in the preceding phase (e.g., FRUIT-or___, when "orange" was not previously cued for the category FRUIT). The Control condition contains members of categories that were not engaged at all during the Generation phase (e.g., INSTRUMENT-vi(olin) when no instruments were invoked during the Generation phase). The *retrieval-induced forgetting effect* is the finding that stem completion is typically less successful in the Activated condition than in the Control condition. A Generated condition (containing items from the Generation phase) is often included as well. This condition reliably shows greater retrieval success than the others (e.g., Anderson et al., 1994; Anderson, 2003; Hulbert et al., 2012). Note that throughout the paper we use transparent, vernacular terms to facilitate broad understanding in place of the procedure-specific acronyms of the retrieval-induced forgetting literature.

Retrieval-induced forgetting (RIF) suggests that prior generation impedes subsequent memory access to categorically related but not generated items. We propose that the pattern of impaired access to related material following generation reflects the same underlying process as the semantic interference observed in language production research. A variety of explanations have been proposed to account for interference/forgetting in the memory literature. The early accounts of Anderson and colleagues proposed that Activated items are inhibited during the Generation phase making them less accessible than controls at later testing, (Anderson et al., 1994; Anderson & Spellman, 1995). The long-term mechanism of this inhibition was left underspecified. Other accounts attribute retrieval-induced forgetting to a blocking process in which previously generated items preclude access to ungenerated competitors from the same category (see Murayama et al., 2014 for an extensive review). According to blocking accounts, encountering category cues used in the Generation phase during the final memory assessment activates the Generation phase context, thereby promoting access to previously Generated items at the expense of related Activated items that are not linked to that context (Jonker et al., 2013; Jonker et al., 2015). However, contrary to this account, RIF has been replicated across a variety of tests specifically designed to limit the influence of context, including variations in which new cue words without any relation to the practiced category cues are used to elicit retrieval during the final assessment phase (e.g., Anderson, 2003; Hulbert et al., 2012; Hanczakowski & Mazzoni, 2013). We suggest that while context certainly plays a role in some scenarios (see Norman et al., 2007; Murayama et al., 2014, for reviews), the blocking accounts do not explain the full range of semantic interference findings.

One other account of retrieval induced forgetting, the computationally implemented adaptive learning account of Norman et al. (2007; see also Norman et al., 2006), is of particular interest from the point of view of rapprochement with the language production literature. This account is partly a successor to the earlier Anderson et al. inhibition account but implemented through specific neurally-inspired mechanisms. The model contains episodic and semantic layers to model both intraconceptual and contextually mediated learning. The semantic aspect is composed of two layers, including fully connected "associative"/category and "item" layers, which enable the representation of the category-item pairs used in RIF procedures. In the model, stem-cue retrieval is implemented as pattern completion, and adaptive learning processes act on the connectivity between microfeatures within conceptual representations. When a particular memory representation is

[1] In the retrieval-induced forgetting literature, the corresponding terms are Rp+ (retrieval practiced) for Generated, Rp- for Activated, and Nrp (no retrieval practice) for the Control condition.

retrieved amidst coactivated competitors, an algorithm identifies weak aspects of the target representation and overly active aspects of semantic competitors by raising network-wide inhibition above and below a predetermined threshold. A learning mechanism then strengthens the connections of the weakly activated target features and weakens the connections to overactive features within competing representations, thereby making them more and less accessible, respectively, in future use (Norman et al., 2007). Thus the incremental adjustment of access to competing items matches the downregulation proposed by Oppenheim et al. (2010) to occur in language production, albeit with a different proposed locus and implementation.

Bridging the literatures

Thus far, we have argued that retrieval-induced forgetting and cumulative semantic interference are both situated in an interconnected conceptual-lexical network, that both may be attributed to adaptive learning mechanisms in the network (e.g., Hughes & Schnur, 2016; Oppenheim et al., 2010; Navarrete et al., 2010; Hulbert et al., 2012; Norman et al., 2007), and that they may indeed be the same phenomenon (Navarrete et al., 2021; Jeye et al., 2021, but see de Zubicaray et al., 2015). As noted above, other researchers have remarked on the intersection of semantic interference in memory and language production. As well as Oppenheim et al. (2010), Navarrete et al. (2010) made direct connections between retrieval-induced forgetting and cumulative semantic interference. However, these earlier discussions primarily entailed juxtaposing findings in language production research to theoretical debates in the retrieval-induced forgetting literature. As discussed earlier, Oppenheim et al. (2010) related the roles of strengthening versus downregulation to the "occlusion versus inhibition" debate in the RIF literature. Navarrete et al. (2010) suggested that the lack of cumulative interference in word naming is in line with the assumption of retrieval specificity proposed by Anderson et al (2000). More recent work by Navarrete et al. (2021) also found that cumulative interference did not depend on successful retrieval, which they noted corresponds to similar findings in the RIF literature by Storm et al. (2006).

Given these occasional discussions of parallels between the fields, why has there been so little integration of the fields or direct comparison of findings across them? Contributing factors may include the use of different dependent measures and very different experimental designs and procedures. Moreover, as discussed, each literature proposes a different locus of interference, and there is a differential consideration of episodic versus semantic memory processes. We now address each of these issues in turn.

Methodological considerations

First, underlying commonalities may be masked by the different research practices of memory and language production researchers. Language production researchers are thoroughly entrenched in picture-naming procedures, whereas the RIF literature is dominated by the valuable but restrictive retrieval practice paradigm, even to the point of adopting domain-specific labels that may impede connective thinking. Here, we emphasize that just because a task is used to study one particular type of processing, this does not mean that other processes are not engaged. For example, category-stem cueing is not simply a "memory task," but a task that actually involves both semantic retrieval and word production. The most important element may be that the different framings of each literature have led to the use of different dependent measures, specifically the exclusive use of continuous latency measures in language studies and of aggregate retrieval success measures in memory experiments.

Indeed it appears that the siloing of these practices occurred despite early evidence of their common ground. A seminal study of semantic interference by Brown et al. (1981) measured both latency and aggregate retrieval success for both category-stem prompts (Experiment 1) and pictures (Experiment 2). Moreover, Brown (1981, Experiment 4) was the first to implement the internal structuring of category members within sequences that would later become the norm in continuous picture naming studies. Brown observed both longer latencies and higher probability of retrieval failure over successive probes (ordinal positions) within categories. Subsequent work by Brown et al. (1985) continued this tradition of measuring both latencies and errors in word generation, and again found robust evidence that interference accumulates over semantically related word completions (see also Brown et al., 2005).

The Brown (1981) study was cited both in initial studies of retrieval-induced forgetting (e.g., Anderson et al., 1994) and in studies of interference in word production (e.g., Kroll & Stewart, 1994; Howard et al., 2006). It is notable that, since then, retrieval-induced forgetting studies have almost exclusively relied on aggregate retrieval success as a dependent measure. In doing so, they have lost the ability to examine how semantic interference unfolds within generative and memory assessment phases of experiments, as well as the sensitivity of latency measures to modulation of accessibility among retrievable items in different conditions. On the other side, although they do routinely log naming errors, including omissions, language production studies rarely consider memory failures. One reason for this is that omissions and other errors are infrequent in picture naming. In any case, unification of memory and language production contributions may require a return to Brown's (1981) early diligence in using both latency and memory measures.

Related to the previous issue, the two fields also use different ways of eliciting generation, picture naming (language production) and category-stem cueing (memory). We suggest that category-stem cued retrieval and picture naming share the key generative process that is implicated in semantic interference. Specifically, both require the same two key stages of processing: retrieval from semantic memory and retrieval and selection of a specific word. There are of course also differences between the tasks. Picture naming procedures necessarily recruit object recognition mechanisms within visual pathways (e.g., Jarret et al., 2022), whereas category-stem cueing procedures entail overt comprehension of category labels, constrained lexical search, and arguably, rely more heavily on episodic memory processes to match the stem to a specific familiarized exemplar (although such processes may be partly implicit). Ironically, given that lexical processes are rarely discussed, the category-stem cueing procedure may engage more lexical processes than picture naming. In line with these differences, we note that de Zubicaray et al. (2015) failed to find activation in brain areas typically associated with the Generation phase of retrieval-induced forgetting (i.e., hippocampus, right IFG, and ACC; see Kuhl et al., 2007) during continuous picture naming, and they concluded from this that the two phenomena are neurally distinct. However, we suggest that the differences in brain signatures may be due to modality-specific processes that are external to the necessarily shared processes engaged during conceptual-lexical retrieval. We also note that the lack of activation in the ACC is inconsistent with other picture naming studies that have found significant activity in the same regions (e.g., Indefrey & Levelt, 2004; Gauvin et al. 2016; Llorens et al., 2016). Overall, we retain our position that picture naming shares some key retrieval processes with category-cued stem completion.

Another methodological impediment to integration is that the study of cumulative semantic interference in language production has focused on the unfolding impact on latencies within a single continuous sequence, while retrieval-induced forgetting is studied using a multi-phase design in which performance in the final phase reflects learning in a prior phase. Each of these approaches has distinct benefits and limitations. In continuous picture naming, the primary focus is on examining how each retrieval instance dynamically impacts access to categorical relatives later in the same sequence. The Generation phase in the retrieval practice paradigm often follows a quite similar design to the continuous naming task in that it requires iterative conceptual-lexical retrieval of interspersed category members within a single sequence. However, critical elements of this presentation method are never considered, including the role of lags between related items in promoting long-term interference rather than short-term facilitation (Navarrete et al., 2014). Given that adaptive learning accounts (see prior discussion) assume that downregulation of competitors occurs during the Generation phase, it is puzzling that the field has focused on measuring the retrieval of related items in the final assessment phase, rather than also investigating the internal dynamics of the Generation phase itself (see Brown, 1981; Brown et al., 1985; and Kuhl et al., 2007, for exceptions). The present work addresses this crucial weakness in the memory research.

At the same time, our research addresses a significant weakness of language production studies, the lack of a true baseline, by adopting a *retrieval-induced forgetting* design. The final phase of this design assesses the effect of retrieving related category members in the prior generation phase by comparing retrieval of additional members of the categories to a true baseline (control) condition. Because continuous picture naming does not have such a baseline measure, there is some ambiguity as to whether progressive slowing of latencies is exclusively due to semantic interference. The assessment of memory after a retention interval in the retrieval practice paradigm necessitated comparison to a baseline control condition and we adopt that design feature here. Thus, the practices in each field have benefits and limitations, and combining them provides the opportunity to adopt the best features of both.

Assumed locus of interference

A key issue that separates the literatures is the presumed location of interference in long-term semantic memory representations versus in conceptual-lexical links. This issue is complex and there is not full agreement on the question of locus in either literature. We have argued above that, in both literatures, adaptive learning accounts can best explain the existing findings. Therefore, we suggest that the issue of locus can be provisionally reduced to the contrasting assumptions of the two influential computational models of Norman et al (2007) and Oppenheim et al (2010). A more detailed consideration of these two models illuminates both the possible benefits and challenges of bringing them into correspondence. In Norman et al. (2007), adaptations are proposed to operate on intra-conceptual configurations of features, whereas in the model proposed by Oppenheim et al. (2010), adaptations modulate the strength of conceptual-lexical links. Since memory and lexical retrieval are inherently intertwined in current generative tasks, it is impossible to discern which process drives interference. For example, one could argue that the Norman et al. (2007) model implicitly incorporates lexico-conceptual representations in the hidden layers that bind collections of semantic features. Therefore, bindings between these wholistic representations and features could potentially be altered by generative activity. Likewise, the Oppenheim et al. (2010) model could be modified to allow dynamic changes in conceptual ensembles as well as changes in conceptual-lexical bindings. Pending new evidence on this matter, semantic interference may be attributable to intra-conceptual change, to change in conceptual-lexical connections, or some combination of both.

Lastly, whereas semantic interference in language production is studied as a semantic phenomenon, retrieval-induced forgetting accounts often include episodic memory processes. This issue is exemplified by the previously discussed contextual blocking accounts, which have posited that "forgetting" results from using cues that are episodically bound to the Generation phase context (e.g., Jonker et al., 2012). In addition, work by Ciranni & Shimamura (1999) suggests that episodic relatedness is sufficient to elicit interference. However, other work contradicts this conclusion. New cues without any intrinsic relatedness to the target (i.e., purely episodic cues) do not produce interference (e.g., Perfect et al., 2004). Moreover, we suggest (in line with the integration efforts of the present work) that the success of semantic accounts of interference in the language production literature may indicate that most existing RIF findings can be accounted for by semantic memory processes alone. In any case, we do not focus on the semantic-episodic distinction in the current work, a position that is in line with modern critiques of over-segregated conceptualizations of episodic and semantic memory (Yee et al., 2018; see also McRae et al., 2021).

In sum, the bridging of these literatures and their perspectives is both desirable and feasible. We suggest that aligning both the theoretical claims and empirical evidence in the cumulative interference and retrieval-induced forgetting literatures will deepen understanding of retrieval-based adaptations and clarify the division of labor between processes in conceptual memory substrates and in lexical pathways. In the current study, we seek to combine the strengths of both language production and memory perspectives in order to overcome their respective weaknesses, thereby paving the way for future integrative work.

Experiment: Cumulative semantic interference in a retrieval-induced forgetting design

In light of these considerations, we designed an experiment that brings the memory and language production literatures into dialogue. Although there are many ways in which integration of the two literatures could be implemented, in this initial investigation we sought to make the case that cumulative interference and retrieval-induced forgetting (RIF) entail the same or similar learning mechanisms acting on the same conceptual-lexical representations. We did not seek to identify the specific mechanisms that drive interference. Rather, this initial proof-of concept exploration leaves open a range of possibilities for future integrative accounts. In service of this aim, we measured latencies in a structured retrieval practice design to investigate if the cumulative interference that is typically found in the language production literature occurs similarly in a category-stem cue Generation procedure. Critically, we also assessed if this generative activity impaired later accessibility of related picture names relative to controls, implementing a retrieval-induced forgetting design in a continuous picture naming procedure.

The experiment reported below follows the four-part structure of a standard retrieval-induced forgetting experiment. The design entailed an initial familiarization procedure (standard in both literatures), a Generation Phase using category-cued stem-completion, a filled retention interval, and a Final Assessment Phase using continuous picture naming. In the initial familiarization procedure, participants were presented with each picture paired with its category and name, which both aids in instantiating the specific category-target combination for the Generation Phase and reduces errors associated with "exuberant responding" in the later Picture Naming Phase (Ferreira et al., 2008). In the Generation Phase, fidelity was maintained to the repeated category-stem completion procedure of the retrieval-induced forgetting literature. However, an internal quartile structure was implemented in which related items were distributed evenly across the sequences. This is a much more controlled presentation than typically used in RIF studies. By measuring both generation latencies and errors, we sought to demonstrate that cumulative semantic interference occurs in the same way whether generation is elicited by category-stem prompts or by pictures. The prompts were presented in three passes to ensure adequate exposure, although past work has suggested that adaptive learning processes may primarily occur in the first pass (Kuhl et al., 2007).

An earlier study with a similar design used single letter cues (i.e., INSECTS – b__) in the generation phase, but retrieval success was lower than desirable (see Hambric & O'Séaghdha, 2022). The current experiment therefore used two-letter stem cues. The Generation Phase was followed by a 15-minute retention interval in which participants completed a medium-level sudoku puzzle. Participants then completed the Final Assessment Phase, which followed the typical design of a continuous picture naming task with name generation latencies as the primary dependent measure. In this phase, we only included the critical Activated and Control conditions. This was partly because the positive effect of prior generation is uncontroversial and highly reliable. More importantly, given that cumulative interference unfolds across encounters with categorically related items, and that Generated and Activated items are drawn from the same categories, we chose to implement a balanced design comprising only the never-generated Activated and Control conditions. This mitigates the possibility of contamination of the Activated condition by previously generated items and thus provides a clean test of the influence of past related generative activity.

The predictions were as follows. In the Generation Phase, we predicted the cumulative interference typically found in language production studies, which manifests as an increase in generation latencies over Ordinal Positions. In the Final Assessment Phase, we also expected typical cumulative interference in both the Activated the Control conditions. Critically, we predicted that, on aggregate, it would take longer to retrieve and produce picture names in the Activated condition relative to the Control condition. This Memory Condition effect would replicate the pattern of hindered access to items related to previously generated items that constitutes the retrieval-induced forgetting effect, but here manifested primarily as increased generation latencies rather than only as decreased retrieval rates. We did not predict a particular pattern of interaction between Memory Condition and Ordinal Position, that is, between new and old interference, but the design can capture any such interaction.

Method

Participants and power considerations

The cross-paradigm nature of the present study made estimation of predicted effect size and required power difficult. Rather than conducting a power analysis, we estimated the appropriate power using the guidelines proposed by Brysbaert & Stevens (2018). According to their recommendations, a reaction time experiment with repeated measures should have a minimum of 1,600 observations per condition. Given that there were 20 target trials in each condition of the critical picture naming phase, 96 participants were recruited from Lehigh University in return for either $8 US (N=13) or class credit (N=83) in an introductory psychology course, resulting in up to 1920 measures per condition.

Materials and apparatus

Ten groups of eight taxonomically related items (80) were selected from the Bank of Standardized Stimuli (Brodeur et al., 2010; Brodeur et al., 2014). In addition, five unrelated fillers were presented at the beginning and end of the Generation phase (see Design) to mitigate primacy and recency effects. Ten additional fillers were presented as task practice at the beginning of the continuous Picture Naming Phase to reduce data loss on experimental trials. The relatedness of items in each taxonomic group was established using normative data from the Bank of Standardized Stimuli (Brodeur et al., 2010; Brodeur et al., 2014) and additional frequency data were acquired from the LexOPS data base to ensure selected items were as equivalent as possible (Taylor et al., 2020). Both average name agreement (87 %) and average category agreement (94 %) were high and always exceeded 60 %. The complete list of materials and selection criteria can be found on OSF (https://doi.org/10.17605/OSF.IO/UASFM). Pictures were presented centrally and sized to approximately 4 in x 4 in (Brodeur et al., 2014). Text was presented in 14-point Corsolas font. The filler task was a printed medium level sudoku puzzle.

The experiment was programmed using E-Prime 3.0 and was conducted using a Dell OptiPlex 7020 computer with a 1908FPt 75 DPI monitor with a refresh rate of 80 Hz. Participants were seated approximately 16 in. from the screen. Voice onset times were recorded using a Psychology Software Tools 200A SR box interfaced to an ATR20 Audio-Technica microphone, situated approximatelty12 inches away from the participant, and audio for the entire session was recorded using a SONY ICD-PX720 recorder.

Design

The experiment comprised four sub-procedures, including an initial familiarization procedure, a Generation Phase, a filler task, and a final Picture Naming Phase. The two critical phases comprised two distinct but linked designs (see below). Assignment of categories between the Generation Phase and the Control condition was balanced across two experiment versions, and assignment of items to Generated vs. ungenerated (Activated) status within each version was further counterbalanced over two sub-versions. To equate the number of items in the Activated and Control conditions, only half of the items in the Control condition were presented, and the half presented was counterbalanced in the sub-versions.

Generation phase

This phase comprised the "retrieval practice" phase typically implemented in retrieval-induced forgetting experiments (e.g., Anderson et al., 1994). Participants underwent category-stem completion retrieval for half (4) of the items from half (5) of the categories. One item from each category was presented in each of four quartiles, thus distributing related items evenly. There was a minimum lag of 1, a maximum lag of 8, and an average lag of 4.5 unrelated items between members of the same category. The order of quartiles was balanced across participants via a Latin Square, and items were randomly presented within quartiles. Following standard practice, this sequence was repeated three times, each Repetition implementing a new ordering of quartiles and items and with no immediate repetition of quartiles across Repetitions. Filler items were presented only once, before the first Repetition and at the end of the third. In this phase, the key variables were the within-subject factors Ordinal Position (1–4) and Repetition (1–3).

Continuous picture naming phase

In this phase, pictures corresponding to items in the Activated and Control conditions (see above) were presented in a continuous stream. Prior to the experimental trials, 10 unrelated filler pictures were presented to reduce data loss associated with task orientation. As in the Generation phase, one item from each Activated and Control category was randomly assigned to each quartile. The order of presentation within each quartile was randomized, and the order of quartiles was counterbalanced across experiment versions. Because immediate repetitions from the same category were not allowed, there was a minimum lag of 1 and a maximum lag of 18 between members of the same category, with an average of 9.5. Thus, there were two within-subject factors: Memory Condition (Activated versus Control) and Ordinal Position (1–4).

Procedure

The experiment comprised four separate procedures (see Fig. 1).

Familiarization

Following standard practice in both the memory and language production literatures, participants were first introduced to the materials. They were told that they would be naming a series of everyday objects and that they would be using the information in subsequent tasks. All items used in later phases, including fillers, were presented as pictures with their category and name in a randomized sequence. Presenting the names with the pictures eliminates the need for active lexical retrieval but makes the designated names (as well as the category-name combinations) available for the experimental phases. Thus, both the category-

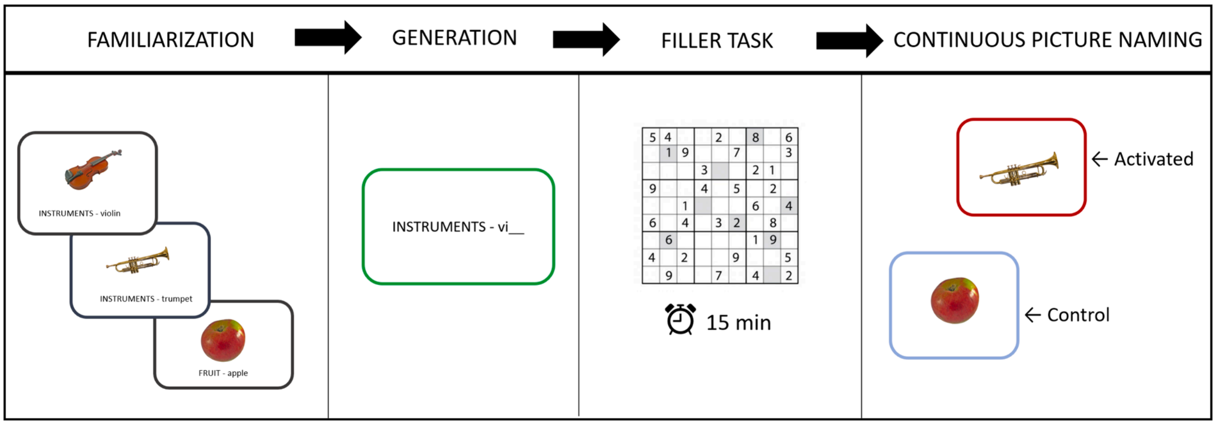

Fig. 1. Schematic of the four-part experimental procedure.

exemplar pairs and corresponding pictures were introduced in a single pass.

On each trial, a fixation cross appeared for 200 ms, followed by the presentation of a picture with the category-name combination (e.g., INSTRUMENTS – violin) below it for 1500 ms. To ensure that the information was registered, participants read each category and picture name aloud. The experimenter provided feedback on standard pronunciation on the few occasions where it was necessary. No data were collected.

Generation Phase: Category-stem cues

In this phase, participants were instructed (with examples) that they would be presented with category-stem cues for some of the familiarized items on each trial and that they should try to retrieve the word matching the cue as quickly and accurately as possible. They were told that each category-stem cue would appear multiple times and that they were to remain silent if they did not recall the word. The stems were always the first two letters of the target name. Each category-stem cue was assigned to a quartile in a continuous sequence, and the whole sequence (with new randomization) was repeated three times (see Generation Phase Design). Five fillers were presented at the beginning and at the end of the phase to alleviate possible primacy and recency effects. On each trial, a fixation cross appeared for 200 ms, followed by the category name and the two-letter stem of the target (e.g., INSTRUMENTS – vi___). The display lasted for 8 s or until a naming response was registered. In order to encourage effortful retrieval across repetitions, no feedback was provided during the Generation phase. The next trial began 1000 ms after the conclusion of the previous one.

Filler task

Participants worked on a printed medium difficulty level sudoku puzzle for 15 min. They were instructed to try to complete the puzzle to the best of their ability, but that there would be no formal assessment of their sudoku performance. The attention-demanding, numerically oriented sudoku task was selected to prevent participants from engaging with/rehearsing experimental materials during the interval.

Assessment Phase: Continuous picture naming

Participants named the target pictures in one continuous sequence internally segmented into quartiles (see Design). Ten unrelated filler pictures preceded the experimental trials. On each trial, a fixation cross was presented for 200 ms followed by the printed category (e.g., INSTRUMENTS) of the upcoming picture for 750 ms. Category labels are not usual in picture naming procedures but were used here for consistency with the Generation phase. The target picture (without label) was then presented for up to 1500 ms or until a naming response was registered. A tone sounded and a "too slow" message appeared for 200 ms if the time limit was exceeded. Another trial began 1000 ms after the previous one. Following the completion of this phase, participants were thanked and received an educational debriefing.

Results

Audio recordings of each session were coded off-line by the first author and trained research assistants. Trials in which an environmental sound set off the voice key and trials in which the reaction time was less than 250 ms were coded as invalid and excluded from all analyses. In both the Generation and Continuous Picture Naming phases, trials in which a participant used an incorrect or incomplete name, responded dysfluently, or did not respond, were coded as salient naming errors and excluded from latency analyses. Both latencies and salient errors were analyzed.

Separate linear mixed model analyses were conducted in R version 4.0.3 using the *afex* package (Singmann & Kellen, 2019) for the Generation and Picture Naming Phases. Sum-zero contrasts were set for the relevant categorical variables and F-tests for the fixed effects were conducted with Satterthwaite approximations applied to estimate the degrees of freedom. There was no meaningful effect of the counterbalancing factor and it was therefore excluded from analyses.

Generation phase

Latency analysis

Descriptively, generation time decreased across Repetitions ($Mean_{R1} = 1441$, SE=38; $Mean_{R2} = 1242$, SE=38; $Mean_{R3} = 1110$, SE=38). More importantly, a 182 ms increase in generation time across ordinal positions was observed in the first Repetition. There was little effect of Ordinal Position in subsequent Repetitions (see Fig. 2).

The maximally converging linear mixed model for latencies included the fixed main effects of Ordinal Position, Repetition, and their interaction, as well as random intercepts for both subjects and items. There was a main effect of Repetition, in which generation times decreased more than 300 ms over Repetitions, $F(2,3988) = 113.20$, $p < .001$. The main effect of Ordinal Position was also significant ($F(3,3997) = 4.51$, $p = .004$): Generation times increased with each newly retrieved category member. The effect of Ordinal Position was qualified by a significant interaction between Repetition and Ordinal Position ($F(6,4024) = 6.50$, $p < .001$), which reflects the concentration of cumulative interference in the first Repetition (see Fig. 2). This is in contrast to some past research, such as that by Navarrete et al. (2010), who found that cumulative interference in picture naming was stable even when the entire sequence was repeated four times (see also Costa et al., 2009). A possible interpretation is that category stem-cueing is strongly generative only until the solution is found. Overall, these data indicate that cumulative interference arises in category-stem completion in the same way as in previous picture naming studies, but that it may not accumulate further

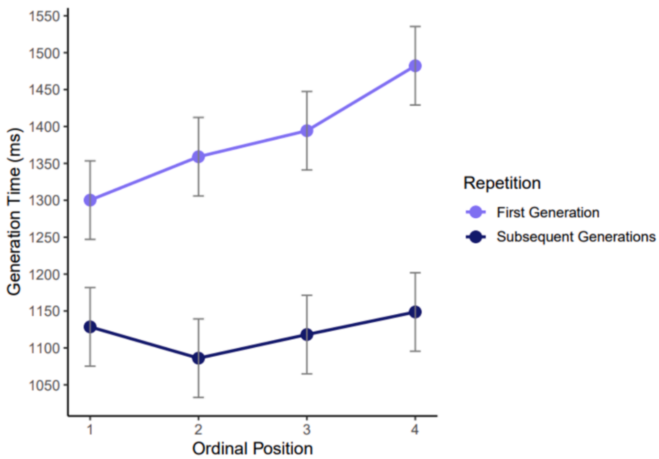

Fig. 2. Cumulative interference in production time for targets in the Generation phase. Subsequent Generations are aggregated over second and third repetitions.

in repeated passes through materials.

Error analysis

The total error rate in this phase was 19 %, which is slightly lower than the retrieval failure rate found in the RIF literature. Descriptively, retrieval errors decreased only slightly with Repetitions (Mean $_{R1}$ = 22 %, Mean $_{R2}$ = 19 %, Mean $_{R3}$ = 17 %). In the first Repetition, the error rate slightly increased across Ordinal Positions, Mean $_{Q1}$ = 20 %, Mean $_{Q2}$ = 22 %, Mean $_{Q3}$ = 20 %, Mean $_{Q4}$ = 25 %. Extraneous/equipment based errors comprised 1.7 % of total trials and 9 % of total errors. Naming errors comprised 91 % of all errors, and of these errors, 249 were commission errors, such as semantic substitutions (e.g., "ha (ndkerchief)" for "ha(t)") while 964 were omissions in which the participant failed to give any response at all within the eight second response window. These omission errors most closely correspond to the retrieval failures reported in the RIF literature. It is notable that, both here and in the RIF literature more generally, retrieval failures are substantial despite the familiarization procedure.

The error analysis included salient naming errors as explained earlier. A logistic linear mixed model analysis was conducted on the error data. Due to error sparsity, the maximal model included only the fixed effect of Repetition and random intercepts for participants and items. We were not able to statistically assess the effect of Ordinal Position. There was a main effect of Repetition ($\chi^2 = 15.62, p = .001$) such that errors decreased across Repetitions. This reflects that participants occasionally generated items in later repetitions that were not retrieved previously. Thus, although adaptive learning processes may primarily operate in the first Repetition, subsequent repetitions may still be somewhat useful in eliciting additional retrievals.

Continuous picture naming phase

Latency analysis

As predicted, picture naming time in the Activated condition (Mean = 792, SE=13) was longer than in the Control condition (Mean = 769, SE=13), a retrieval-induced semantic interference effect. Naming times also increased over the first three Ordinal Positions (see Fig. 3), showing new interference within the Picture Naming phase and convergence of latencies in the two conditions over Ordinal Positions.

The maximal linear model included fixed main effects of Memory Condition and Ordinal Position as well as their interaction, and random intercepts for both subjects and items. The 23 ms main effect of Memory Condition, F (1,3011) = 9.66, p = .002, a retrieval-induced semantic interference effect, was significant. The increased interference over quartiles was also significant (F (3,3014) = 15.88, $p < .001$), indicating new cumulative interference over Ordinal Positions in both conditions. In addition, the interaction between Memory Condition and Ordinal Positions was marginally significant (F (3,3009) = 2.56, p = .053), suggesting that the steeper accrual of interference in the control condition is meaningful. Planned comparisons show that the effect of Memory Condition was significant only in the first (z = 3.23, p < .001) and second (z = 2.31, p = .02) Ordinal Positions. Latencies converged in the third and fourth Ordinal Positions. In contrast to the Generation phase, latencies in both conditions even decreased slightly in the fourth position. A follow up comparison indicated that the difference between the third and fourth positions was not significant, z = 1.21, $p > .05$. The important implication of these findings is that retrieval-induced interference from the Generation Phase may be stronger at the beginning of

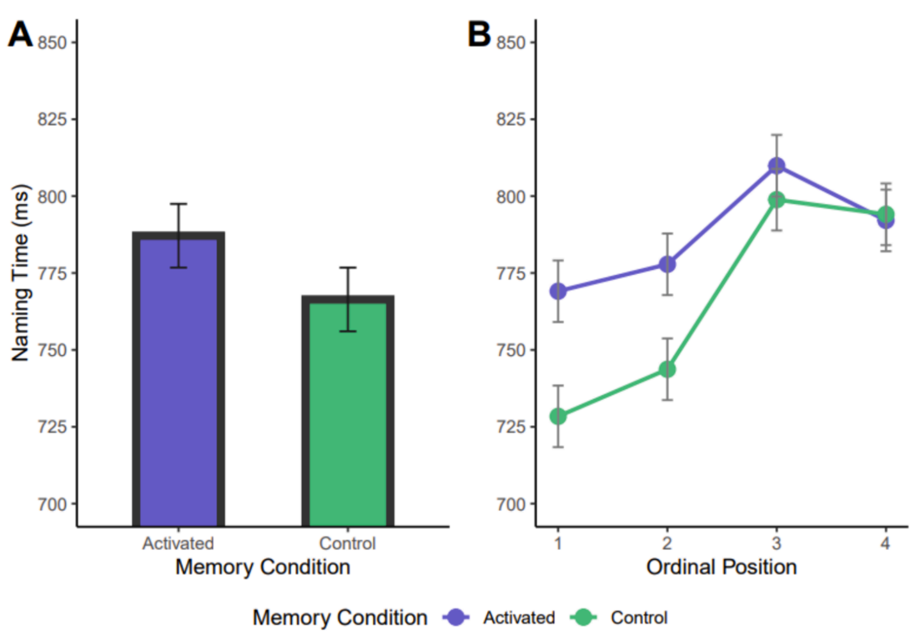

Fig. 3. Picture naming time as a function of Memory Condition. A) Significantly greater picture naming RTs for previously Activated categories relative to unactivated Controls. B) Cumulative interference in experimental and control conditions converges over Ordinal Positions.

the assessment phase and may even be completely absent as interference accrues more strongly in the control condition.

Error analysis

The overall error rate in this phase was 12 %. Extraneous/equipment based errors comprised 1.6 % of total trials and 14 % of total errors. Naming errors comprised 86 % of all errors, and of these errors, 205 were commission errors, and 262 were omission errors. The number of errors was similar across Memory Conditions, Activated = 159, Control = 150. As in the Generation phase, a logistic linear mixed model analysis was conducted on the target naming errors. The final model included the fixed factors of Memory Condition and Ordinal Position, their interaction, and random intercepts for subjects and items. The effects of Memory Condition ($\chi^2 = 0.66, p > .05$), Ordinal Position ($\chi^2 = 2.00, p > .05$), and their interaction ($\chi^2 = 3.39, p > .05$) were not significant. Thus, whereas there was a RIF pattern in naming latencies, there were no effects in errors. Picture naming errors in this study may thus be mostly due to considerations other than semantic interference.

Discussion

This research is part of a larger effort to bring independent but parallel language and memory accounts of retrieval-based long-term semantic interference into alignment. Separate cumulative semantic interference (language production) and retrieval-induced forgetting (memory) literatures both investigate decrements in later accessibility of activated but unselected concepts or words. Influential accounts within each literature propose that adaptive incremental learning processes are at the root of these decrements (Oppenheim et al., 2010; Howard et al., 2006; Norman et al., 2007), but synthesis of these complementary perspectives has yet to occur. The review of the respective literatures highlighted the strengths and limitations of current memory and language production approaches to semantic interference, and the experiment illustrated the value of an integrative approach. The findings showed lasting semantic interference in a continuous picture naming procedure following earlier related generative activity. The ways that semantic interference accrued over time within both stem-completion (Generation) and picture naming (Assessment) phases provided additional insights. We conclude that there is no difference in the underlying basis of *retrieval-induced forgetting* and *cumulative semantic interference,* and therefore propose that the literatures be integrated under the rubric *retrieval-induced semantic interference.*

Key findings

We now consider key findings in more detail. The experiment adopted a retrieval practice design from the retrieval-induced forgetting literature but employed latency measures to provide continuous assessment of semantic interference in both the Generation (category-cued stem completion) and Assessment (picture naming) phases of the experiment. The design provided a rigorous test of long-term interference relative to baseline and, by changing tasks between the crucial phases, showed that interference is a general process. The Generation phase, harking back to pioneering work of Brown (1981), measured latencies as well as errors in a structured category-cued stem completion task. The presentation sequence was structured such that one member of each category occurred in each of four quartiles. We observed cumulative interference in latencies only in the first of three reshuffled presentations of the sequence. Analysis of this phase thus provides insight into the emergence of interference in a typical memory research procedure and suggests, in agreement with Kuhl et al., (2007), that the standard practice of eliciting generation multiple times may have limited value when the generation phase uses category-stem completion. As noted earlier, this conclusion is task specific. Navarrete et al., (2010) showed that cumulative interference is robust over repetitions of picture naming sequences, so it appears that cumulative interference in picture naming does persist over repeated passes through the materials.

In the final picture naming phase, it took longer to name pictures corresponding to new items from previously used categories relative to controls. That is, the findings showed a retrieval-induced forgetting (RIF) pattern in picture naming latencies despite the change of format between the Generation and Assessment phases. The use of full counterbalancing between conditions, and over ordinal positions within them, bolsters this finding. The effect survived a 15 min retention interval, providing important evidence for the temporal durability proposed in both memory and language accounts that incorporate incremental learning (Norman et al., 2007; Oppenheim et al., 2010). This is solid evidence of long-term semantic interference relative to a true baseline – which we have noted is absent in most language production studies of semantic interference in continuous picture naming. In addition to the interference carried over from the Generation phase, additional interference accrued over ordinal positions in both the Activated and Control conditions of this phase, with a somewhat steeper accumulation observed in the Control condition. Planned comparisons showed that the difference between the conditions was significant only in the first half of the picture naming phase. Although this finding was not predicted, it underlines the value of examining cumulative interference rather than only using aggregate measures of successful retrieval. Additional within-task experimentation is needed to establish whether the long-term effect of retrieval-induced semantic interference is reliably stronger earlier than later in a sequence of retrievals.

The durability of semantic interference across tasks in the experiment provides novel evidence for cue-independence, a central tenet of long-term adaptation accounts of retrieval-induced forgetting (e.g., Anderson, 2003; Hulbert et al., 2012; Camp et al., 2007; Perfect et al., 2004). The cue-independence assumption states that retrieval-induced forgetting is independent of the particular cues used to elicit retrieval and so involves change within semantic memory (Anderson, 2003). The assumption has been tested by using different retrieval cues during initial and later retrievals (e.g., Hulbert et al., 2012). However, finding truly independent cues can be difficult, and even when this problem is solved participants may nevertheless use the originally assigned categories covertly (covert cueing hypothesis, see Jonker et al., 2012). In the present study, we used pictures as independent cues in the Assessment phase, and still observed a retrieval-induced forgetting effect in the picture naming latencies, supporting cue independence.

Extensions of present study

The study presented here used one particular design that combined elements of memory and language production methodologies, but of course many other variations are possible. First, in the present study we implemented the standard repeated category-stem cueing procedure during the Generation phase and used picture naming in the Assessment phase. While the switch from category-stem cue to picture naming highlights the robustness of semantic interference, it would be interesting to see if a similar pattern of results would be obtained if the tasks were flipped, such that the Generation phase entailed picture naming and the Assessment comprised category-stem cued retrieval. This would make it possible to assess memory measures as well as continuous retrieval latencies, providing more direct comparison to findings in the retrieval-induced forgetting literature. In other variations, both phases could use category-cued stem completion, as is done in the RIF literature (sometimes with minor changes in the cues between phases), or both phases could use picture naming. Each of these variations could provide additional insights.

Second, our findings call into question the value of repetitions in the Generation phase of the standard retrieval practice paradigm. Repeated production of category-stem cued targets is typical in the retrieval practice paradigm, but the present results suggest that adaptive learning processes are concentrated in the first generation of the items. Thus, the repetitions may merely serve as additional opportunities to retrieve items that were not retrieved in the initial pass. However, even this may have limited influence on semantic interference, given that work in both RIF and in language production has found that successful retrieval is not required to induce later interference (Storm et al., 2006; Navarrete et al., 2021). That is, downregulation of coactivated items may occur even if the target is not generated. On the other hand, semantic interference does persist through repeated passes of picture naming sequences (Navarrete et al., 2010), so quantifying how much repetition influences semantic interference in experimental variations with picture naming in the generation phase should be a priority. Varying the repetitions is easy to implement and will help inform under what conditions adaptive learning operates in both category-stem cued generation and picture naming.

Lastly, we chose not to repeat items from the Generation phase in the Assessment phase for reasons discussed earlier. It would be valuable to assess whether including a Generated condition affects the Activated condition, and especially the accumulation of interference over Ordinal Positions. Specifically, we reasoned that naming previously generated and strengthened items intermittently within the sequence could disrupt the linear accumulation of interference in the Activated condition. While this is at odds with the finding that repeating a single item did not disrupt the overall linear trajectory of cumulative interference in Oppenheim et al.'s (2010) Simulation 2, whether the same is true for an equal number of repeated and ungenerated items is far less certain and warrants further investigation. In sum, it would be useful to know whether inclusion of the Generated condition affects the strength of retrieval induced semantic interference, how this plays out over time, and how much the Generated condition itself shows new interference in a later phase.

Methodological considerations

We highlight four concrete methodological recommendations based on lessons learned from the present study. First, impaired access to Activated items relative to Control was observed in latency data. However, this RIF effect was evident only in the first half of the assessment phase. Thus, latencies revealed changes in relative accessibility that the aggregate retrieval measures that are typically reported in the memory literature do not. Therefore, we strongly recommend the use of latency measures in the assessment of semantic interference.

Second, we recommend that future work in retrieval-induced forgetting use latencies to assess the interference in the Generation as well as the Assessment phases of the retrieval practice paradigm. Specifically, as noted above, investigation into the internal dynamics of the Generation phase can help reveal how adaptive learning unfolds over time, and under what conditions.

Third, the design of the retrieval practice paradigm provides a baseline for assessing sustained interference that has been lacking in the language production literature. The lack of a true baseline is a weakness of the continuous picture naming paradigm, and implementation of a multiphase design such as we used here addresses this issue. Future cumulative interference studies should incorporate baselines. That said, the baseline we have adopted from the retrieval induced forgetting literature may itself have limitations. Convergence of latencies in Activated and Control conditions does not appear to indicate that interference from the generation phase is reduced over time (see Fig. 3). Instead, it could indicate that the validity of the baseline is reduced as it incurs strong within-phase interference.

Fourth, the present work differentiated between error types in both the Generation and Picture Naming phases. In the memory literature, errors are viewed as the converse of successful retrieval, and they are generally not analyzed further. Moreover, some responses that may be recorded as erroneous in the language production literature, such as disfluent responses, are not counted as errors in the memory literature. The language production literature has a long history of using the *type* of error to further elucidate underlying processes. For example, complete generation failures are termed omission errors, and other types of errors that entail dysfluent or inaccurate responses are termed commission errors. Such a nuanced approach to the analysis of errors might likewise provide added insight in studies of memory retrieval.

Theoretical considerations

Integration across literatures must happen at a theoretical level in addition to incorporating useful aspects of respective methodologies. In addition to the issue of cue independence discussed earlier, several other theoretical assumptions formalized by Anderson and colleagues may be fruitful areas of theoretical alignment and testing. The assumption of *retrieval specificity* states that explicit retrieval (rather than mere exposure) is necessary for semantic interference. This is evidenced by the absence of interference when activated items are presented (but not generated) prior to the assessment phase (e.g., Ciranni & Shimamura, 1999; see Murayama et al., 2014 for review). The same assumption is made throughout the language production literature (e.g., Navarrete et al., 2010; Belke, 2013) and is presumed by Oppenheim et al. (2010). Forthcoming work in our own lab confirms that cumulative interference is not found when reading semantically related words aloud rather than generating the names from pictures (O'Séaghdha et al., in prep). Further work is needed to more fully understand retrieval specificity, and the generality of retrieval-induced interference across different retrieval tasks.

The assumption of *interference dependence* states that the degree of observed forgetting is dependent on the degree of interference during the Generation phase (Anderson, 2003). Evidence concerning the factors that determine the degree of interference in language production is mixed. For example, Rose et al. (2019) found cumulative interference for related items with high feature overlap (e.g., *chimpanzee, orangutan*) but *not* for those sharing only a broad category membership (*orangutan, donkey*). In addition, Oppenheim (2018) found that newly acquired category members elicited the same degree of interference as those already incorporated in the lexicon. However, Rose et al. (2017) observed cumulative interference for thematic relations *without* any feature overlap (e.g., *pacifier* and *baby)*, suggesting that feature overlap cannot be the whole story. More research is needed to reconcile these findings.

Lastly, the assumption of *temporal insensitivity* explicitly proposed by Oppenheim et al. (2010) and implicitly assumed by Anderson et al. and Norman et al. (2007) is severely under-researched. According to Oppenheim et al. (2010), cumulative semantic interference is insensitive to time, such that it only dissipates with relevant production experience. The current evidence for this assumption is also mixed. While some work suggests that interference does dissipate across extensive lags of unrelated items (Schnur, 2014), the present study showed that interference persisted across a 15-minute retention interval. The longevity of semantic interference in memory has also been subject to debate with mixed results, with some work suggesting that forgetting vanishes after 24-hours (MacLeod & McRae, 2001), but other studies showing effects even after a one-week delay (Storm et al., 2012; Tandoh & Naja, 2007). Clearly, there is a strong overlap of theoretical interests across literatures, and we expect that these specific issues and others would be better addressed in a unified framework.

A unified conceptual-lexical approach

The initial integrative work reported here provides clear evidence of common ground between memory and language production approaches to semantic interference, but also leaves questions regarding the locus of semantic interference. The present research does not address where exactly along the conceptual-lexical interface semantic interference originates. As discussed previously, Norman et al. (2007) and Oppenheim et al. (2010) propose different loci of semantic interference. Specifically, Norman et al. (2007) assume that forgetting is ultimately traceable to changes to the connectivity of features within conceptual representations. However, this account neglects the potential role of the linguistic processes engaged in most recall tasks and conflates memory storage and access into a single layer. In contrast, the model of Oppenheim et al. (2010) adopts an opposite bias by ascribing semantic interference entirely to changes in the links between concepts and abstract words (lemmas). Like other language production accounts, Oppenheim et al. do not consider the possibility of dynamic changes in conceptual representations themselves, though it is uncontroversial to assume that adaptive learning process also occur in the conceptual stratum (e.g., Lambon-Ralph et al., 2017; Rodd et al. 2004; Oppenheim & Nozari, 2024). Theoretically, we see no basis for assuming either conflation. We suggest that this logic may, in principle, be extended to any phenomenon that requires both conceptual and lexical retrieval. For example, the tip-of-the-tongue phenomenon studied in production is often attributed to impoverished links between the relevant semantic features and the target lexical representation (e.g., Meyer & Bock, 1992; Gollan & Brown, 2006). However, if conceptual level adaptive learning processes also affect lexical accessibility, this effect could plausibly be explained by weakened connections among the conceptual features themselves. Overall, we do not yet know how much of interference in linguistically mediated generation occurs in the conceptual (semantic memory) stratum versus in mappings between concepts and lexical representations.

The final and most critical recommendation of this article is the development of a computational model that includes dynamically adjustable conceptual feature configurations (Norman et al., 2007) as well as links between the conceptual stratum and lexical units (as in Oppenheim et al., 2010). The successes of both types of models suggest that they capture important semantic interference processes, but we simply do not know what the division of labor is between intra-conceptual and conceptual-lexical processes, or how this may vary with different tasks and procedures. An elaborated model, in conjunction with directed empirical investigation, should serve as a primary goal of future work.

Conclusion

In summary, this paper makes a case for unifying the study of retrieval-based semantic interference in the fields of language production and memory. Both manifestations of interference reflect the long-term consequences of activating semantic competitors, and both fields are grounded in conceptual-lexical retrieval. We therefore argue that an integrated account of conceptual-lexical retrieval removes unnecessary obstacles to a more complete understanding of semantic interference. Such an integration sets the stage for addressing important unresolved issues such as determining the relative contributions of intra-conceptual adaptive processes and conceptual-lexical adaptations to semantic interference in a variety of circumstances.

CRediT authorship contribution statement

Channing E. Hambric: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pádraig G. O'Séaghdha:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All by-trial data, analysis code, and materials are available on OSF.

Acknowledgments

We thank Emily Wegrzyn, Cindy Tran, and Diane Marrow for their invaluable help with data collection and coding.

References

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

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 149–161.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1063–1087.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review., 102*(1), 68–100.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7*, 522–530.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49*(4), 415–445.

Anderson, M. C. (2020). Incidental forgetting. In *Memory* (pp. 277–313). Routledge.

Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *Annual Review of Psychology, 72*, 1–36.

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.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology, 66*(11), 2135–2160.

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS One1, 9*, e106953.

Brodeur, M. B., 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 One1, 5*, e10773.

Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory, 7*(3), 204–215.

Brown, A. S., Whiteman, S. L., Cattoi, R. J., & Bradley, C. K. (1985). Associative strength level and retrieval inhibition in semantic memory. *The American Journal of Psychology*, 421–432.

Brown, A. S., Zoccoli, S. L., & Leahy, M. M. (2005). Cumulating retrieval inhibition in semantic and lexical domains. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 496–507.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 1–20.

Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item- specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(5), 950–958.

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1403–1414.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences, 106*(50), 21442–21446.

Damian, M. F., Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition, 81*(3), B77–B86.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*(1–3), 287–314.

de Zubicaray, G., McMahon, K., & Howard, D. (2015). Perfusion fMRI evidence for priming of shared feature-to-lexical connections during cumulative semantic interference in spoken word production. *Language, Cognition and Neuroscience, 30*(3), 261–272.

Gauvin, H. S., De Baene, W., Brass, M., & Hartsuiker, R. J. (2016). Conflict monitoring in speech processing: An fMRI study of error detection in speech production and perception. *NeuroImage, 126*, 96–105.

Gollan, T. H., & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General, 135*(3), 462–483.

Gordon, J. K., Young, M., & Garcia, C. (2018). Why do older adults have difficulty with semantic fluency? *Aging, Neuropsychology, and Cognition, 25*(6), 803–828.

Hambric, C. E., & O'Seaghdha, P. (2022). Aligning language and memory accounts of semantic interference. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Hambric, C. E., & O'Séaghdha, P. G. (2023). The unseen, the seen, and the spoken: Latent and overt priming in cyclic picture naming. *Quarterly Journal of Experimental Psychology, 76*(10), 2410–2430.

Hanczakowski, M., & Mazzoni, G. (2013). Contextual match and cue-independence of retrieval-induced forgetting: Testing the prediction of the model by Norman, Newman, and Detre (2007). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 953–958.

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.
Hulbert, J. C., Shivde, G., & Anderson, M. C. (2012). Evidence against associative blocking as a cause of cue-independent retrieval-induced forgetting. *Experimental Psychology, 59*(1), 11–21.
Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1–2), 101–144.
Jarret, J., Ferré, P., Chedid, G., Bedetti, C., Bore, A., Joanette, Y., & Brambati, S. M. (2022). Functional network and structural connections involved in picture naming. *Brain and Language, 231*, Article 105146.
Jeye, B. M., Kark, S. M., Spets, D. S., Moo, L. R., Kensinger, E. A., & Slotnick, S. D. (2021). Support for an inhibitory model of word retrieval. *Neuroscience Letters, 755*, Article 135876.
Johnson, S. K., & Anderson, M. C. (2004). The role of inhibitory control in forgetting semantic knowledge. *Psychological Science, 15*(7), 448–453.
Jonker, T. R., Seli, P., & MacLeod, C. M. (2015). Retrieval-induced forgetting and context. *Current Directions in Psychological Science, 24*(4), 273–278.
Jonker, T. R., Seli, P., & MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review, 120*(4), 852–872.
Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1083–1105.
Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience, 10*(7), 908–914.
Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletin & Review, 28*, 40–80.
Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*(1), 42–55.
Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–38.
Llorens, A., Dubarry, A. S., Trébuchon, A., Chauvel, P., Alario, F. X., & Liégeois-Chauvel, C. (2016). Contextual modulation of hippocampal activity during picture naming. *Brain and Language, 159*, 92–101.
Lupker, S. J. (1988). Picture naming: An investigation of the nature of categorical priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 444–455.
MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12*, 148–152.
Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503–535.
McRae, K., Brown, K. S., & Elman, J. L. (2021). Prediction-based learning and processing of event knowledge. *Topics in Cognitive Science, 13*(1), 206–223.
Merton, R. K. (1961). Singletons and multiples in scientific discovery: A chapter in the sociology of science. *Proceedings of the American Philosophical Society, 105*(5), 470–486.
Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition, 20*(6), 715–726.
Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: A meta-analytic review of retrieval-induced forgetting. *Psychological Bulletin, 140*(5), 1383–1409.
Navarrete, E., Mahon, B. Z., & Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica, 134*(3), 279–289.
Navarrete, E., Del Prato, P., Peressotti, F., & Mahon, B. Z. (2014). Lexical selection is not by competition: Evidence from the blocked naming paradigm. *Journal of Memory and Language, 76*, 253–272.
Navarrete, E., Benavides-Varela, S., Lorusso, R., & Arfè, B. (2021). Cumulative semantic cost without successful naming. *Memory & Cognition*, 1–12.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*(3), 226–254.
Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review, 114*(4), 887–953.
Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114*(2), 227–252.
Oppenheim, G. M. (2018). The paca that roared: Immediate cumulative semantic interference among newly acquired words. *Cognition, 177*, 21–29.
Oppenheim, G. M. (2024). Lexical selection in language production. https://doi.org/10.31234/osf.io/rjezp.
Oppenheim, G. M., & Nozari, N. (2024). Similarity-induced interference or facilitation in language production reflects representation, not selection. *Cognition, 245*, Article 105720.
Perfect, T. J., Stark, L. J., Tree, J. J., Moulin, C. J., Ahmed, L., & Hutter, R. (2004). Transfer a ppropriate forgetting: The cue-dependent nature of retrieval-induced forgetting. *Journal of Memory and Language, 51*(3), 399–417.
Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*(3), 460–499.
Ritvo, V. J., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic plasticity: How memory retrieval drives learning. *Trends in Cognitive Sciences, 23*(9), 726–742.
Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28*(1), 89–104.
Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language, 68*(2), 180–198.
Rodd, J. M., Gilbert, R. A., & Betts, H. N. (2017). The role of learning mechanisms in understanding spoken words. *Cognition, 40*, 1095–1108.
Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition, 172*, 59–72.
Rose, S. B., Aristei, S., Melinger, A., & Abdel Rahman, R. (2019). The closer they are, the more they interfere: Semantic similarity of word distractors increases competition in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(4), 753–763.
Rose, S. B., & Abdel Rahman, R. (2017). Semantic similarity promotes interference in the continuous naming paradigm: Behavioural and electrophysiological evidence. *Language, Cognition and Neuroscience, 32*(1), 55–68.
Scaltritti, M., Peressotti, F., & Navarrete, E. (2017). A joint investigation of semantic facilitation and semantic interference in continuous naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(5), 818–823.
Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences, 106*(1), 322–327.
Schnur, T. T. (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language, 75*, 27–44.
Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*(1), 86–102.
Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. *In New Methods in Cognitive Psychology*, 4–31.
Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review, 13*, 1023–1027.
Storm, B. C., Bjork, E. L., & Bjork, R. A. (2012). On the durability of retrieval-induced forgetting. *Journal of Cognitive Psychology, 24*(5), 617–629.
Tandoh, K., & Naka, M. (2007). Durability of retrieval-induced forgetting. *Shinrigaku Kenkyu: The Japanese Journal of Psychology, 78*(3), 310–315.
Taylor, J. E., Beith, A., & Sereno, S. C. (2020). LexOPS: An R package and user interface for the controlled generation of word stimuli. *Behaviour Research Methods, 52*, 2372–2382.
Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language, 33*(3), 332–356.
Yee, E., Jones, M. N., & McRae, K. (2018). Semantic memory. *The stevens' handbook of experimental psychology and cognitive. neuroscience, 3*.