Cumulative Semantic Interference in People with Aphasia

Peter Hart (5559421)

Universiteit Utrecht

MA Thesis
Linguistics: The Study of the Language Faculty
Utrecht University
August 2017
Supervisors: Dr. Hayo Terband, Utrecht University
Prof. David Howard, Newcastle University, UK

Abstract

Purpose The purpose of this study was to investigate whether a well-studied interference effect, Cumulative Semantic Interference, occurs in a group of People with Aphasia, and to see whether it requires correct picture-naming or just exposure to a picture.

Method 25 people with Aphasia, all with some degree of word-finding difficulty, participated in a continuous naming task. Two sessions were performed by each participant; each session contained a total of 150 items from 5 semantic categories.

Results Following Linear Mixed Effects Modelling, the results did not generate the anticipated interference phenomenon for these participants, and a post-hoc reanalysis with narrower semantic categories also found no effect.

Conclusion The second hypothesis was expected to argue in favour of one of the three computational models of the effect. A range of possible explanations for this negative finding are offered, and potential adaptations to answer this question are offered.

Table of contents

1. Introduction

Lexical selection proceeds by selecting a single word from amongst a variety of competing alternatives (e.g. Dell, 1986; Levelt et al., 1999; for review see Spalek, Damian & Bölte, 2013). For healthy speakers, this system of retrieving and producing words is effortless, fast and impressively reliable. The picture of a dynamic system of competitive lexical selection is well-established in a variety of psycholinguistic settings. When, occasionally, the selection goes awry, the resulting errors are overwhelmingly likely to be semantically or phonologically related to the intended word, suggesting that these are strongly competitive or intrusive alternatives (Dell & Sullivan, 2003) In addition, the proximity of semantic alternatives predicts the time it takes to name an item (e.g. Vigliocco et al. 2002; Vigliocco et al., 2004; Costa et al., 2005; Aristei & Abdel Rahman, 2013).

A recent variation of competitive selection is that, when participants are required to name several semantically related items in a session, each consecutive item is named more slowly than the previous item (Howard et al., 2006; see also e.g. Schnur et al., 2006; Oppenheim et al., 2010). For example, if a participant has correctly named a picture of a *cup*, and is then presented with the semantically-related *plate*, the naming of *plate* will be slower than expected. This slowing effect does not hold for one item only, but is cumulative, such that the total slowing effect increases for each consecutive item. If a third item, *fork,* is to be named, it is even slower once again. The phenomenon is impervious to other external factors; the number of intervening items or amount of intervening time have no effect. This *Cumulative Semantic Inhibition* effect (CSI; Howard et al., 2006), is by now well-studied and appears robust (e.g. Howard et al., 2006; Schnur et al., 2006; Belke, 2013; Hoedemaker et al., 2017).

Interestingly, Schnur et al. (2006) demonstrated that the effect does not only hold for healthy participants; the responses of his participants with Aphasia (PWA) exhibit the CSI effect in a comparable way. Acquired Aphasia is a language disorder caused by stroke or neurodegenerative disease, which can have such a range of possible symptoms and subcategories that it usually makes little sense to discuss people with Aphasia as a unitary group (Baker, LeBlanc & Raetz, 2008). However, perhaps the most prominent and consistent characteristic across different types of Aphasia is word-finding difficulties (e.g., Goodglass & Wingfield, 1997). In picture-naming tasks, PWA are consistently slower to name items, and often name items incorrectly or fail to name items at all (e.g. Dell et al., 2004).

Since word retrieval requires the speaker to overcome competition, it is plausible that Aphasics' word-finding difficulty is the result of an impairment in this competitive selection process, and therefore that the Left Inferior Frontal Cortex (LIFG), or Broca's area, plays a role in selecting between competing alternatives (Schnur, Schwartz, Kimberg, Hirshorn, Coslett, Thompson-Schill, 2009).

The phenomenon of *Cumulative Semantic Inhibition* has informed the development of three computational models of lexical processing (Howard et al., 2006; Oppenheim et al., 2010; Belke, 2013) which are designed to capture the persistent nature of the effect by incorporating weight-changing mechanisms that do not dissipate as a function of time or unrelated input. For PWA, the effect is broadly similar; hence it is possible to manipulate the existing models to extend to the performance of impaired participants, as has been done by Oppenheim et al. (2010; simulation 4).

The assumption, then, is that whichever mechanisms are responsible for the effect in healthy participants also hold for those with impaired language systems. However, there have not been many additional studies into the effect in PWA; of all the published studies investigating the CSI effect, only Schnur et al. (2006) and Stielow & Belke (2015) included a language-impaired group of participants. The present experiment is the first large-scale study done within the experimental paradigm called *continuous naming task* (Howard et al., 2006).

The aim of the present study is twofold. The first aim is to extend the current experimental background for the effect in populations with Aphasia. The second hypothesis investigates whether the CSI effect is generated in the absence of successful lexical retrieval, i.e. when the picture is not successfully named. The motivation for this second hypothesis relies on a closer understanding of the different and incompatible predictions of the computational models, which are discussed in the theoretical background.

In this respect, the word-finding difficulties of PWA present a unique opportunity to test this aspect of the models. An understanding of whether semantic interference is generated by conceptual or lexical activation is relevant to the ongoing development of treatment plans for Aphasic patients.

In section 2, I present the theoretical background to the paper, describing the experimental paradigms and results that are the foundation for the relevant models of lexical processing. The hypotheses are formally restated in section 3, focusing on the underlying differences between these models, which generate competing predictions for the present experimental data. Section 4 then presents the experiment, with results, discussion and conclusion to follow in sections 5 and 6.

2. Theoretical Background

2.1 The Cumulative Semantic Interference effect

The Cumulative Semantic Interference effect (henceforth CSI) was first explicitly demonstrated in Howard et al. (2006), but has been established and extended by a series of supplementary experimental findings, lending support to the nature of semantic interference as a long-term, persistent effect. The key findings on the nature of the effect can be summarised as follows:

- The slowing effect on response times holds consistently, regardless of the number of unrelated words intervening between the two category exemplars (e.g. Howard et al., 2006).
- The effect also does not dissipate as a function of time, e.g. when there is a longer lag between items (Hsiao, Schwartz, Schnur & Dell, 2009) and interference persists despite a 1-hour break between two rounds of the experiment (Oppenheim et al., 2010)
- The effect is cumulative for multiple consecutive exemplars, i.e. each exemplar is named more slowly than the previous one: *cat* is slower than *bear*, and *dog* is slower again than *cat.*

Howard et al. (2006) demonstrated that this slowing effect was linear out to 5 items, generating an average slowing effect of ~30ms for each additional category exemplar (revised to ~26ms in the reanalysis of Alario & Moscoso Del Prado Martín (2010)). These features form the core of the CSI effect. A simple example should help to make the phenomenon more concrete: there are three items from the same semantic category (*animals: bear, cat, dog*), with intervening items from other semantic categories.

The CSI effect can be contrasted with another, better-known experimental effect: the *repetition priming* effect (Mitchell & Brown, 1988). This effect occurs when speakers are required to produce the same word on consecutive trials – in this case, the second item is named more quickly and accurately than expected (for review see Francis, 2014). So, the successful naming of a picture as *dog* means that the next time *dog* is named, it is faster than expected.

However, when the words are not identical, but are semantically related, there is a slower response for the second token (e.g. Wheeldon & Monsell, 1994; Belke, Meyer & Damian, 2005; Damian & Als, 2005), and an increased incidence of error (e.g. Schnur et al., 2006). So, where *cat* is followed by *dog,* the naming of *dog* is slower than expected, or more likely to be named incorrectly. This is the beginning of the Cumulative Semantic Interference (CSI) effect. It is very likely that these two phenomena, priming and inhibition, can be subsumed under a single, parsimonious model of lexical selection. Oppenheim et al. (2010; p.2) refer to this second, semantic inhibitory effect as "the dark side" of repetition priming, capturing the notion that the two phenomena, repetition priming and CSI, are "two sides of the same coin", i.e. explainable by one and the same set of processes.

2.2 Experimental Paradigms

Two types of picture naming tasks have been successfully used to generate the CSI effect: the Continuous Naming Task and the Blocked Naming Task. The Continuous Naming Task, first used by Brown (1981), is a picture naming task using a single continuous selection of pictures from a number of semantic categories. The original paradigm was adapted by Howard et al. (2006) to ensure that the position of a word within its category was not confounded with its overall position in the experiment, so that the data can be corrected for slowing throughout the experiment. In the experiment of Howard et al. (2006), there were five items from each of 24 semantic categories, in an order that ensures that there are either 2, 4, 6 or 8 intervening unrelated items. The data demonstrated a CSI effect: there was a ~30ms increase in naming latency for each subsequent item from the same semantic category, and the effect was consistent up to 5 items, the number of items tested. This was the first demonstration that the phenomenon is long-lasting (persisting until at least the end of the experiment), cumulative (the RTs of each item within the semantic category increase monotonically), and insensitive to the number of semantically unrelated items between two category exemplars.

In the Blocked Naming Task, participants name a series of pictures in either homogeneous or heterogeneous blocks in sets of 6 items. In the homogeneous set, all pictures are from the same semantic category (e.g. *cat, dog, horse, goat*...) while in the heterogeneous set items are from different categories (e.g. *bread, coat, dog, glass*…). Pictures are repeated multiple times within a block, and results show semantic facilitation for the first cycle, with faster latencies in the homogeneous condition, and semantic inhibition for subsequent cycles, with longer naming latencies in the homogeneous condition (Abdel Rahman & Melinger, 2011; Navarrete, Del Prato & Mahon, 2012). Again, intervening unrelated trials do no damage to the slowing effect (Hsiao, Schwartz, Schnur & Dell, 2009).

The experimental paradigms converge in producing the same broad phenomenon: object naming reliably interferes with the retrieval of subsequent semantically related items. Importantly, there is an increase in naming latencies for subsequent items within the category, and in each paradigm the phenomenon is resistant to the effects of intervening time or irrelevant trials.

2.3 Modelling Lexical Processing

Computational models of lexical processing are intended to predict and quantify the data from human experiments in a parsimonious way. There are three models that attempt to capture and explain the findings from CSI-inducing experiments: Howard et al. (2006), Oppenheim et al. (2010) and Belke (2013). Throughout the remainder of this thesis these models are respectively referred to as **HM**, **OM** and **BM** (with **M** standing for “model”)**.** The models can be seen as derivatives of longer-standing lexical processing models, simplified to capture the more specific phenomenon of CSI. Though narrower models such as these are under no obligation to capture a broader range of data, they should at least be compatible with empirical findings from other domains of lexical processing, and ideally should be plausible as a domain-general neuropsychological mechanism (e.g. Oppenheim et al., 2010; p.29).

The three computational models that will be considered are substantially similar in their core assumptions. The core mechanisms are rooted in influential models such as WEAVER++ (Levelt, Roelofs & Meyer, 1999) and Featural and Unitary Semantic Space model (FUSS; Vigliocco, Vinson, Lewis & Garrett, 2004). The first of these considerations: **stages of processing,** incorporates discussion of conceptual and lexical levels, alongside the mechanisms of **spreading activation,** and **competitive selection**. These are well-established aspects of pre-existing models of lexical processing that are borrowed for use in these more specific models. An additional component, **persistent priming**, is a less widespread feature of lexical processing models, though it has received some previous discussion (Damian & Als, 2005; Wheeldon & Monsell, 1994), and similar learning processes appear in cognitive models from other domains, such as Retrieval-Induced Forgetting (see Oppenheim et al., 2010).

Stages of processing. To capture the full story from semantics to lexical retrieval and production, models are divided into more-or-less encapsulated stages. More developed models have more levels, enabling them to account for a wider range of observed phenomena, or to capture the extended chain of processing from conceptualisation to articulation.

Lexical processing is usually concerned with two broad categories: *lexical selection* and *phonological retrieval* (e.g. Dell, 1986; Dell, Schwartz, Martin, Saffran & Gagnon, 1997; Foygel & Dell, 2000). However, in comparing the present models in relation to CSI, it will be more informative to look "upstream", to focus on a distinction between *conceptual processing* and *lexical selection*.

Conceptual processing. Early stages of conceptual processing are incorporated within lexical processing models if they are broadly construed. A system of *conceptual processing* can be divided to include two modes of semantic features: decomposed conceptual nodes and nondecomposed concepts. The nodes of this latter stage are also referred to as *lexico-semantic representations* (FUSS; Vigliocco, Vinson, Lewis & Garrett, 2004), or *lexical concepts* (WEAVER++; Levelt, Roelofs & Meyer, 1999), and they intervene between the decomposed conceptual nodes and the lexical level. Lexical concepts include information such as the concept's appearance, function, and other semantic properties (Friedmann, Biran & Dotan, 2013), and in most cases they have a one-to-one connection with a lexical representation (Levelt, Roelofs & Meyer, 1999; Levelt, 2001).

HM makes no strong commitment to the nature of semantic processing, since the computational model finds no functional difference between the two modes, while OM opts for decomposed semantics. BM assumes a full complement of decomposed semantics with an intervening stage of lexical concepts. In the model, decomposed semantic elements are bundled into nondecomposed lexical concepts prior to lexical encoding, as in the Featural and Unitary Semantic Space model (FUSS; Vigliocco, Vinson, Lewis & Garrett, 2004).

Note that these simplifications do not mean that, say, HM and OM are committed to the claim that only these levels exist in a full model of lexical processing. Recall that the models presented here are idealisations that capture a narrower subset of data; as such, they incorporate the smallest number of components necessary to capture the effect. Within a maximally parsimonious model of a single phenomenon, here the CSI effect, a smaller number of components may be sufficient to model the data.

Spreading activation. Aside from the structure of the system, another feature of processing is its functional workings. *Spreading activation* is stipulated to some extent in virtually all theories of word production (e.g. Bloem, van den Boogaard & La Heij, 2004; Caramazza, 1997; Dell, 1986; Roelofs, 1992; Levelt, Roelofs & Meyer, 1999), and appears essential to capture the likelihood of well-recognised phenomena such as semantically and phonologically related errors (e.g. Dell, 1986; Levelt et al., 1999). According to one possible conception of spreading activation, the activation of several semantic features at the conceptual level leads to activation and selection of multiple relevant lemmas at the lexical stage. The process is in fact deterministic in normal cases, since the word with the most corresponding features ought to receive the most activation. The same semantic features also activate the other lexical items to which they are connected, leading to the (lesser) activation of related semantic concepts. At the semantic level, activation is additive, so the lexical concept that is connected to the most conceptual features receives the most activation, and is selected for the next stage.

More specific evidence of spreading activation at the semantic level in fact comes from semantic interference tasks and CSI experiments: the degree of semantic relatedness predicts the size of the refractory effect (in a semantic interference task; Vigliocco, Vinson, Damian & Levelt, 2002; and in CSI; Alario & Moscoso Del Prado Martín, 2010; Rose & Abdel Rahman, 2017).

Lexical Selection. The *lexical* level stores the representations of corresponding lexemes and includes the mechanism responsible for selecting the appropriate lexical representation (Levelt, 1989). The mechanism of lexical selection is one of the most debated aspects of general models of lexical retrieval (for a recent review see Britt et al., 2016). In both HM and BM, there is competitive selection between multiple activated forms at the lexical level. Again, the authors make no commitment to the exact details of competition - HM notes that selection could be performed by a differential threshold used in WEAVER++ (Levelt et al., 1999), or by a mechanism of lateral inhibition that holds between all competitors, according to which a target word is irretrievable because it is inhibited by its activated competitors (e.g. Harley, 1993; Stemberger, 1985).

On the other hand, OM makes the claim that competitive selection is unnecessary to model CSI, provided that a model uses a mechanism of error-driven learning, explored in further detail below. A replacement non-competitive mechanism would select a winner once its activation reaches a certain threshold (e.g. Mahon et al., 2007). In this “winner takes all” model, selection is unaffected by the activation level of competitors. It should be noted that, in each case, the decision on whether to include competitive selection is contingent on the decision made “upstream” on the nature of persistent priming.

Persistent priming. Coactivation of items at the lexical level is normally used to explain Stroop-like effects in picture-word interference tasks (PWI; e.g. Damian & Bowers, 2003; for reviews see Spalek, Damian & Bolte, 2012; Navarette & Mahon, 2013). Given that activation levels are assumed to dissipate quickly, Howard et al. noted that primed activation levels alone are insufficient to model the persistence of CSI, and its insensitivity to intervening time and semantically irrelevant trials.

Some kind of long-term or incremental change is essential to generate the interference effect on the appropriate time-scale, and such a mechanism is advocated by each of the models. These persistent mechanisms receive more attention in the next section, but their role within the models can be briefly contrasted here. In HM, the weight-change is strengthening-only, and holds between semantic nodes and lemmas, while in OM, the weight-change is both strengthening and weakening. It is the nature, rather than the location, of the incremental learning procedure that forms the crux of the difference between HM and OM. In BM, the incremental learning mechanism is compressed into the conceptual level, holding between the decomposed conceptual nodes and the stage of lexical concepts. The change is strengthening-only, as in HM, but holds between decomposed semantic nodes and lexical concepts, rather than mapping between the conceptual and lexical levels.

The models do not entirely agree on their use of terminology, perhaps reflecting the inspiration of alternative prior models, though the underlying components described are similar or identical. It was already noted that BM's lexico-semantic representation can be equated with *lexical concepts,* which in turn equates to nondecomposed semantic representations. The decision to identify some components with one another across models ensures that the three models differ in fewer superfluous ways, allowing a more direct comparison of their integral differences. The core difference, following these assumptions, lies in the location and nature of the weight-changing mechanism used to generate the CSI effect. For reference, Table 1 provides a summary of the key components just discussed.

Model	Author(s)	Spreading activation	Levels	Level of weight change	Type of weight-change	Competitive lexical selection
HM	Howard et al. (2006)	✓	(Conceptual) Lexical concept, Lexical representation	conceptual - lexical	Increase only	✓
OM	Oppenheim et al. (2010)	✓	Conceptual, Lexical representation	conceptual - lexical	**Increase and decrease**	✕
BM	Belke (2013)	✓	(Conceptual) Lexical concept, Lexical representation	**(conceptual) - lexical concept**	Increase only	✓

Table 1: A summary of the most important aspects of models, with the various components of each model. Items in brackets are optional within the model, and the most important differences are boldface.

2.4 Modelling Cumulative Semantic Interference

Having outlined the core components of any potential model of lexical retrieval in a semantic context, it will be informative to present a detailed walkthrough of how picture-naming proceeds in each model, contrasting the way that CSI is generated in each case.

HM. A picture of a dog is presented, and activation spreads at the semantic level, such that when the semantic nodes for *dog* are activated, so are those for *cat* and to a lesser extent *rat* and *bear*. In a successful naming event, the target lemma receives the greatest activation, and is retrieved following a competitive selection process of some kind. Though Howard et al. do not specify a process of competitive selection, let us assume that a version of the Luce ratio from WEAVER++ applies: activation from the conceptual system is a continuous process, and activation spreads to the lexical level until the target word reaches beyond its peers by a critical difference (e.g. Piai et al., 2014). Following lexical retrieval, the selected lexical item is, optionally, articulated.

A successful selection event at the lexical level triggers the incremental learning mechanism: the mapping between the lexeme *dog* and the semantic nodes that activated it is strengthened, while the mapping from the same semantic nodes to the unsuccessful lexical competitors remains unchanged.

This weight-changing mechanism simultaneously captures both the priming and interference effects in picture naming tasks involving multiple items. When a new word from the same semantic category is required, activation spreads more rapidly to its competitors, which are semantically related to various degrees. Conversely, the next time the same item is named, the activation proceeds through the strengthened mapping more quickly, and the lemma is activated and selected more quickly, leading to faster naming.

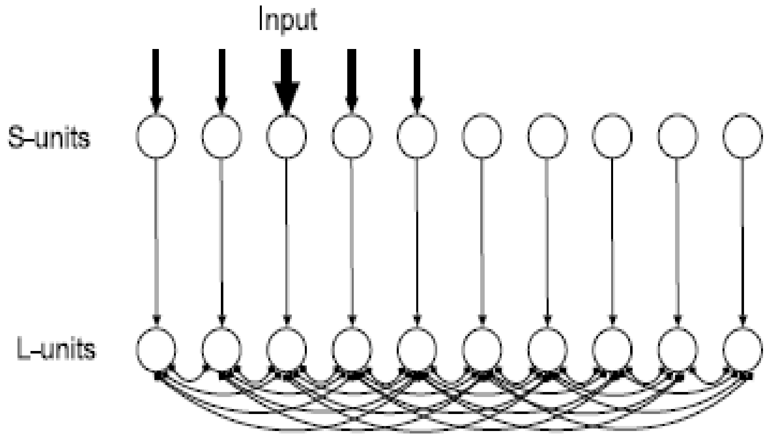

Fig. 1. The computational model of Howard et al. (2006). Note the direct correspondence from Semantic units to lexical units (here respectively S-units and L-units), and the lateral inhibition between lexical units.

OM. In Oppenheim et al.'s 'error-driven' model, activation spreads from semantic to lexical nodes in the same way as in HM, activating several items at the lexical level. Again, following lexical retrieval, the mapping between the semantic nodes and the target word is strengthened. The important modification here is that, at the same time, mapping between those semantic nodes and the incorrect lexical competitors is *weakened.* The degree of weight change, either strengthening or weakening, is determined by the difference between the activation level of the target lemma, and the 'desired' level, required for lexical selection.

Returning to the concrete example, successful retrieval of *dog* leads to the strengthening of the mapping between those semantic nodes and the lemma *dog,* and the simultaneous weakening of the mapping from the same semantic nodes to the lemma *cat.* If, in a later trial, *cat* is the target item, activation proceeds more slowly through the newly weakened mapping from the conceptual level to the lexical level, and naming is slowed.

To repeat, although this initial difference between HM and OM seems relatively subtle, the alternative conception of the nature of persistent priming has an important effect on the downstream mechanism of competitive selection – there is no longer any need to invoke a competitive lexical selection process to explain CSI (Oppenheim et al., 2010; simulations 5 & 6). The inhibition effect is carried by the incremental weakening of the mapping that underlies semantic competitors.

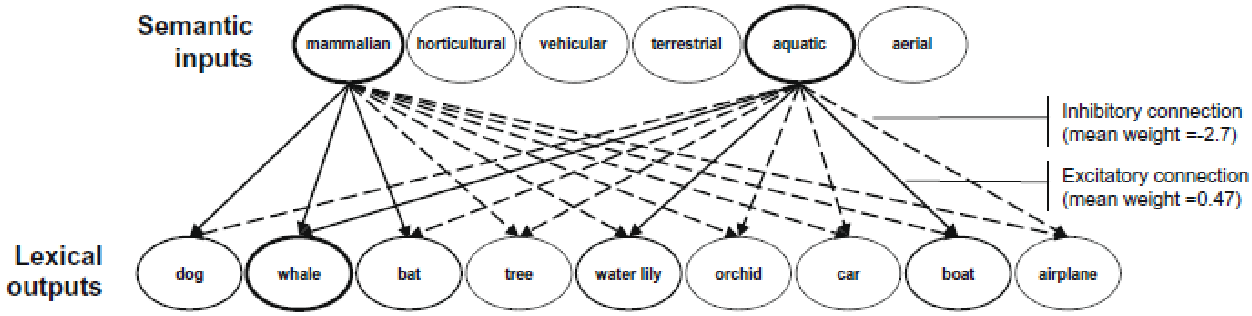

Fig. 2. The *error-driven* model of Oppenheim at al. (2010). Here the semantic units *mammalian* and *aquatic* are activated, and activation spreads to a variety of lexical units. Solid lines represent connections that are to be strengthened following this item-naming event, while dashed lines are weakened.

BM. The model of Belke (2013) makes a more radical departure from HM than does OM. In arguing for the Conceptual Accumulation account, according to which the origin of the CSI effect is contained entirely within the conceptual level, rather than at the conceptual-lexical interface. The conceptual level here encompasses decomposed semantic features leading in to lexical concepts, and a layer of lexical representations follows. The model is driven by the hypothesis that the incremental learning process is conceptually-mediated, rather than originating at the lexical level. Belke demonstrated this using a semantic classification task in which a cumulative *facilitation* effect was argued to be the result of conceptual-level residual priming (Experiment 1) and showed that semantic classification and object naming influence one another (Experiment 5; Belke, 2013). This view is consistent with some interesting additional findings: the CSI effect transfers across languages in bilingual participants (Runnqvist, Strijkers, Alario & Costa, 2012), which could suggest a shared origin of the effect at the conceptual level. In Hoedemaker et al. (2017), the authors performed continuous naming tasks with two participants who either named an object, or observed their partner naming an object. Though self-produced items were more likely to be recalled, there was no difference in the strength of the effect for self-produced versus other-produced items.

In a picture naming event, the model proceeds by changing the weights of connections from the distributed semantic nodes to lexical concepts. Alternatively, the same effect is achieved by including some mechanism of long-term priming that acts upon lexical concepts alone. This preserves the function of the distributed semantics stage while obviating the need to include its details within the model. Whichever method of weight-strengthening is used, repeated access to a semantic category generates “residual activation” at the conceptual level (Kroll and Stewart, 1994; Belke, Meyer & Damian, 2005). By the same process as in HM, in subsequent naming events, competing alternatives become activated more strongly at the lexical level. Since this weight-change mechanism is once again strengthening-only, as in HM but as opposed to OM. It follows that the need for a competitive lexical selection mechanism returns.

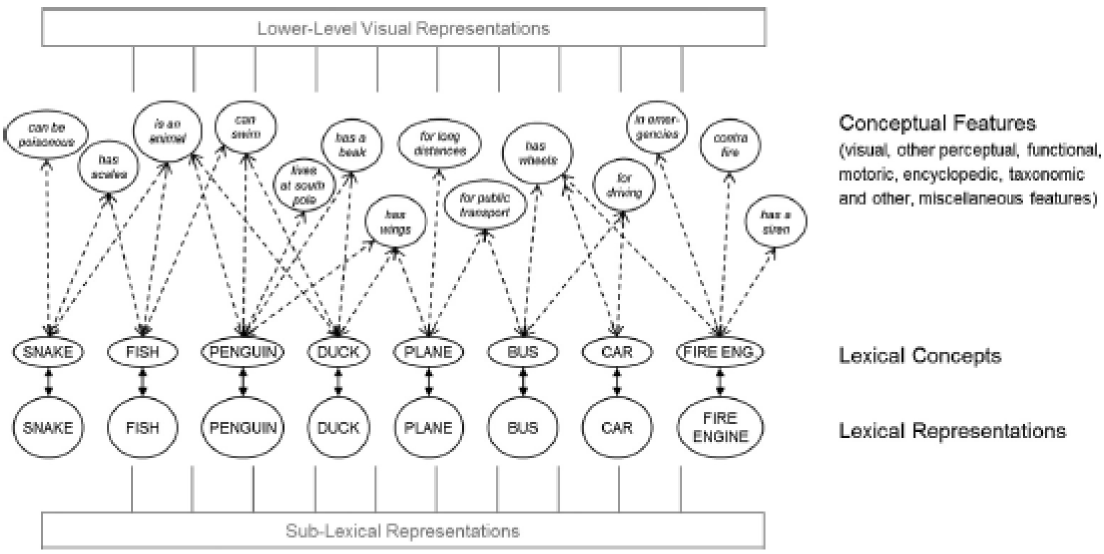

Fig. 3. The c*onceptual accumulation* model of Belke (2013). Incremental eight changes are strengthening-only, but now occur between semantic units (here "conceptual features") and *Lexical Concepts*, not seen in the previous two models. Most lexical concepts directly correspond to a lexical representation.

2.5 Semantic Interference in People with aphasia

As well as predicting the slowing of RTs in a continuous naming task, the models under consideration here can be extended to capture the performance of people with Aphasia (PWA). Almost all PWA experience some degree of word-finding difficulty, regardless of the classification of aphasia (Kohn & Goodglass, 1985; Goodglass & Wingfield, 1997). In picture-naming tasks, PWA experience more difficulty in resolving semantic interference (e.g. Biegler, Crowther & Martin, 2008; McCarthy and Kartsounis, 2000; Wilshire and McCarthy, 2002). Much recent research has focused on the role of the Left Inferior Frontal Gyrus (LIFG), often lesioned in people with Broca's aphasia, as responsible for competitive selection (e.g. Canini et al., 2016; Jefferies et al., 2008; Piai et al., 2014; Britt, Ferrara & Mirman, 2016).

To date, there are only two published experiments that apply CSI paradigms to people with Aphasia: in a blocked naming task, Schnur et al. (2006) investigated 18 PWA and found an increased incidence of semantic errors in the homogeneous condition compared to the mixed condition, without a corresponding increase in phonological or other errors (experiment 2), especially for the subgroup with Broca's aphasia. Stielow & Belke (2015) investigated 9 PWA and found that performance on a blocked naming task correlated with a participant's executive functioning, but found no such correlation in the continuous naming task.

The source of semantic errors and omissions. In tasks such as the CSI paradigms that generate a semantic interference effect, PWA can be expected to produce a greater proportion of semantic, phonological, mixed or omission errors.

Semantic errors are made more likely in semantically homogenous contexts due to excessive activation of competing lemmas, or underactivation of the target lemma (Schnur et al., 2006). Errors occasionally occur in healthy participants under more taxing experimental conditions, such as in a semantically homogeneous context, or if the time until cutoff is short (see Schwartz, 2006; Table 2). The greatly increased likelihood of semantic errors in people with Aphasia can be modelled by simply including noise within the system (Oppenheim et al., 2010). In the event of a semantic error, it is safe to assume that the item was recognised – the error would then be a result of erroneous selection at the lexical level.

However, it does not automatically follow that the correct lemma was not retrieved. If we are to claim this, then we will need to know that the target form was not retrieved at any (later) stage. Correct retrieval could happen in a cascading system, where there is simultaneous activation of multiple lemmas, and one incorrect form is sent to the phonological stage before the competition has been resolved. The speaker could soon realise that a mistake had been made, and retrieve the correct form - although too late. Nonetheless, the correct lemma is selected and, according to HM and BM, the weight-change should proceed as if it were a correct naming event.

Phonological errors suggest that processing has been successful up to and including lexical retrieval, and that the error has taken place at the phonological retrieval or motor programming stages. An additional logical possibility is that an error occurs at both semantic and phonological levels, with the accumulation of spreading activation from both types of similarity giving rise to a "mixed error" (Dell & Reich, 1981).

Lexical-level selection is clearly not the only possible route to an omission error. In the case of omissions suggesting a source of the error is particularly difficult with people with Aphasia, whose language systems may have varied and multiple underlying impairments. To list some of the alternative possibilities, a patient could succeed in all the steps of lexical retrieval, up to retrieving the (phonological) lexical representation, before being unable to retrieve the motor plan to be able to produce the word. At the earliest end of the system, patients could fail to even recognise the item, and therefore fail to fully activate the conceptual nodes or retrieve the lexical concept downstream.

Omissions are trials where no response is produced before the cut-off time (5000ms in the present experiment) – which includes hedges or verbal groping with no phonological relation to a word. Omissions are, by their nature, not very informative about the underlying impairment within a processing framework. Conceivably, the source of the error could be a failure to: recognise the salient features of an item (decomposed semantics), recognise the whole item (derive a lexical concept), retrieve the lemma (lexical selection), retrieve the phonological form of the lemma, and one of several stages at the motor level, as a result of dysarthria or dyspraxia. For the present purposes, the most important consideration will be whether the lemma was successfully retrieved at the lexical level.

3. Research Questions

There are two research questions, restated more formally below:

(1) Does the experiment produce the CSI effect in this group of PWA? A finding here relies on following the analyses of this experiment's predecessors, Howard et al. (2006), with reference to the mixed-effects reanalysis of Alario & Moscoso Del Prado Martín (2010)

(2) Is the mechanism underlying the CSI effect driven by lexical selection, or picture recognition (i.e. Conceptual activation)? These predictions are driven by the alternative conceptions of the three computational models we have examined, made explicit below.

The present experiment is a continuous naming task using people with Aphasia, and, with 25 participants, it is the first large-scale study of this type. The only previous study of this type, by Belke & Stielow (2015), replicated Howard et al.'s original paradigm with 9 participants with Aphasia.

The potential for a high incidence of omissions presents a good opportunity for testing the second hypothesis. Seemingly, it would be simpler to use healthy participants and ask them to view, but not name, a picture. However, such an experiment would fail due to the likelihood that healthy participants would involuntarily retrieve the correct lemma at the lexical selection stage, but opt not to produce it at a later stage.

3.1 Competing predictions of computational models

The models are primarily built upon the experimental reaction time trajectories in healthy participants, and the authors do not usually extend the models to make specific predictions about impaired participants, or the procedure in the event of omission or error responses. The exception here is Oppenheim et al. – the authors implemented additional noise at the lexical level to generate errors and omissions. Nonetheless, on the basis of the models' details, it is possible to derive divergent predictions about the effects of such responses on the workings of the model, generating different results for the overall CSI effect.

HM. The computational model explicitly requires lexical retrieval in order to perform connection-strengthening – it is correct lexical retrieval that drives the learning procedure. So, for omissions and semantic errors that do not involve lexical selection of the target word, we would expect there to be no connection-strengthening. The simple prediction, then, is that omission and semantic errors should not contribute to the CSI effect, and that they could be treated in the same way as, for example, semantically unrelated filler trials.

It should be noted here that HM is a relatively simple model, and was not intended to extend to modelling omissions and errors, so it's not clear that the authors would stand by these predictions. Furthermore, as mentioned in section 2.5, the model's prediction may not be so simple. The weight-changing mechanism requires lexical access, which can occur in the absence of lexical production, i.e. despite the response qualifying as an omission. Alternatively, following a semantic error, the system could realise that a mistake has been made and retrieves the correct lexical item. In this case, the effect of a semantic error or omission could even be the same as a normal event of correct naming. I shall return to this point following discussion of OM, in which the same issue applies.

Prediction: No effect of omission; possible effect of errors, depending on whether the appropriate lemma is in fact retrieved

OM. This is the only model to explicitly attempt to account for impaired processing, with the simple addition of noise in activation levels during lexical selection (simulation 4; Oppenheim et al., 2010). As always, the most highly activated item is selected, such that if the addition of noise causes a competitor to exceed the target word, then a semantic error follows. Unlike HM, OM does not stipulate that a lexical item need be correctly retrieved in order to produce a weight-change. Their model finds inspiration from the more general neuropsychological phenomenon *Retrieval-Induced*

Forgetting: they note that, in a Retrieval-Induced Naming task, erroneous or omission responses generated an equivalent effect (Anderson, Bjork & Bjork, 1994; Storm, Bjork, Bjork and Nestojko, 2006). Therefore, the authors make the specific prediction that the incremental change should "even accrue from naming trials that elicit omissions or errors" (p.28, Oppenheim et al., 2010).

In a personal communication, Gary Oppenheim elaborated the likely responses of the model in the face of erroneous and omission responses. In each case, the strengthening component of the mechanism requires as input the desired activation of the target word, and therefore needs to know what the intended target word should have been. This leads to some relatively straightforward predictions in the case of semantic errors and omissions. If the correct word is known to the system, the system proceeds much as it would if the correct item had been retrieved: connections to the target item are strengthened while those to its competitors are weakened. On the other hand, if the correct name is not known, there is weakening of the erroneous items, without corresponding strengthening of the target item. The result is that, for errors, regardless of whether items were "known" to the system, the mechanism always predicts at least *some* weight-changing in the event of an error.

For omissions, the case is less clear: by the same reasoning as in semantic errors, the system must know what the correct lemma was in order to strengthen and weaken connections appropriately. If the lemma is known to the system, there is a change in connection weights just as there would be for a correct naming event. Overall, these predictions translate to the expectation that semantic errors contribute to the CSI effect, though the appearance of errors in which the target word is not known would predict that the overall effect would be less strong than correct naming.

Prediction: No effect of omission; and a possible effect of errors (though less than for correct naming)

BM. When the same set of conceptual-level nodes are repeatedly activated, persistent activation accumulates within that subsection of the conceptual system. It follows that a semantic interference effect is generated regardless of whether the lemma is retrieved, provided that the item is "recognised", i.e. the lexical concept is retrieved. Where there is a semantic error, it is quite likely that item has been recognised, and the error has arisen at the level of lexical selection. If this assumption is true, this means that we can expect semantic errors to contribute towards the CSI effect in the same way as correct naming responses. The case of omissions is less clear, since we cannot know whether the item has been recognised. If the item is recognised but the lemma is not retrieved, the effect should the same as an instance of correct naming. If the omission results from a failure to recognise the item, there should be no effect.

Prediction: Full effect of semantic errors, possible effect of omissions, depending on whether the item is recognised

4. Experiment

The data for this analysis is taken from a previous study called SemaFoRe (Morris, Howard & Buerke, 2014), which compared two types of treatment for word production difficulties in PWA. The original study was longitudinal, taking 7 measurements of the task at various stages of the intervention program. Two base-level tests were taken prior to any treatment, and these are used in the present analysis.

4.1 Participants

Participants were people with aphasia who were at least 6 months post stroke. All patients had some difficulty with lexical retrieval; to qualify they needed to score between 10-60% on the Nickels Naming screen (Nickels, 1992). It was in theory possible to categorise participants into subgroups based on their aphasia diagnosis. However, the results of the clinical impression were not available for all patients, so this information is not included in the analysis. Average (and range) age was 68 (48-89) time post-onset 23 months (5 months – 10 years). A battery of assessments on language, executive function and memory was completed prior to testing, reproduced in Appendix A.

4.2 Materials

A largely new set of items and pictures were devised by the original authors for this test. The pictures were simple colour images of a total of 150 everyday items, with 21 items in each of four semantic categories, and a filler category comprising 45 miscellaneous items. The complete list of stimuli is reproduced in Appendix B.

4.3 Procedure

The experiment was performed by a trained speech and language therapist at Newcastle University, UK, in 2014. The presentation of the experiment was organised using DMDX (Forster & Forster, 2003). Each trial comprised 150 unique items, divided into two parts of 75 items each to prevent excessive fatigue. Each part of the test began with two filler items, always *sea* and *chocolate,* and each picture was presented for 5000ms before it disappeared. There were two trials for each participant, taken at least a week apart, and the two trials did not use the same order of presentation.

4.4 Coding protocol

IPA transcriptions of the response were taken by the experimenter at the time of the experiment. Subsequent analysis of responses was performed using CheckVocal (Protopapas, 2007). This reanalysis included checking the transcriptions, recording the latencies of response onsets and classifying errors, according to the taxonomy in Table 2. To ensure continuity and comparability with previous data, the protocol was developed to be consistent with protocols used in previous studies (e.g. Schnur et al. 2006, p.209; Oppenheim, 2010).

On the basis of responses, a list of acceptable alternative names for each item was also agreed, which were to be treated as correct responses (available from the author). Also included as correct were responses that contained a slight distortion of a single phoneme, and responses that were only partially produced but subsequently correctly completed. In these cases the onset of the latter, correct utterance was taken as the RT.

10 participants were diagnosed with mild or moderate dysarthria or apraxia of speech (numbers 2, 3, 4, 5, 8, 11, 12, 19, 21 and 24). Following Schnur et al. (2006), the responses from

these patients were coded leniently, using the protocol of the Philadelphia Naming Test (Roach, Schwartz, Martin, Grewal & Brecher, 1996).

Error type	Description	Example
Omission	No response attempted	"I don't know", or unrelated verbal groping
Semantic	A semantically related response (subcategories) **SC –** category relation **SS –** synonym, or picture part (i.e. meronym) **SH –** hypernym/superordinate **SH2 –** hyponym **SO –** associated	*lion* -> "tiger" *Beach* -> "sand" *yacht*-> "boat" *Ice-cream*->"Cornetto" *bench* -> "park"
Phonological-formal	A semantically unrelated real word, shares phonemes with target. Partial production of correct response	*Yacht -> "*watch" *Elephant* -> /ɛlə/
Phonological-nonword	Nonword response that shares phonemes with target	Chest -> /kɛs/
Mixed	Semantically related, shares phonemes with target	*Lemon*-> "Lime" *spider* -> "tiger"
Description	Characterisation of the object, without target word	*Bunk* -> "It's a pair of beds"
Other	Unrelated error, no semantic or phonological relation	*olives* -> "marble"

Table 2. A taxonomy of error types. The protocol was discussed and agreed with the original researchers, and examples of each type of error are provided.

4.5 Assumptions in analysis

In omissions, we cannot know whether the item was recognised. The issue of interpreting omissions is not unique to experiments with PWA, but applies more generally to investigations of lexical processing below the level of production. An item that is not named may not have even been recognised. On the other hand, items that are not overtly named may nonetheless be named covertly, with similar effects on the incremental learning system (see Hoedemaker et al., 2017; p. 56).

Since these alternatives cannot be easily distinguished, some assumptions are required in the present analysis. In a semantic error, it is probable that the target word is selected following the selection and production of the incorrect alternative. The errors, then, are assumed to be largely driven by the participants' impaired processing combined with the time-pressure in the experiment.

If it is found that the interference effect is in fact generated by omissions, there are in fact two possible conclusions: it could be that the omissions are surreptitiously named in the majority of cases, meaning that lexical retrieval occurs anyway. Alternatively, the lexical item is not retrieved, and it is conceptual activation that leads to the effect, in support of the *conceptual accumulation hypothesis* of BM. One aspect of the patients' background testing suggests that the latter is the more likely case: they all performed strongly on the Pyramids & Palm Trees Test (Howard & Paterson, 1992), a test of conceptual knowledge (range 81-98%, mean 91%).

For the purposes of the analysis, the assumption is therefore that the conceptual activation *always* occurs in the case of an omission. It is noted that this is an imperfect solution, but the ambiguity of the phenomenon makes it necessary.

4.6 Data processing

The experiment included 2 x 150 trials from 25 participants, giving a total of 7,500 trials. However, some of the data was unavailable or unsuitable due to experimental error or data corruption. Four participants were partly affected by these errors. A total of 606 trials were thus unavailable for the analysis (8%). Overall, it appears likely that the participants in this experiment had more severe word-finding difficulties than participants used in similar studies: with a 5000ms deadline, Schnur et al.'s (2006) blocked naming task yielded an average error rate of 28% (range 3-91%).

To analyse the effect of semantic competitors on reaction times, only correct responses are required. Thus, the data were further processed to remove all responses coded as omissions and errors, so that only correct responses remained. One participant was excluded entirely because there were too few correct responses (4%), leaving 1,894 correct responses (25.2% of the total trials) from 24 participants. The total proportion of erroneous or omitted responses was 73%, (range 59-90%).

All subsequent data processing and analyses were performed using the statistical program *R* (2012). Reaction times were log-transformed using the natural logarithm to approach a normal distribution, and visual inspection of quantile-quantile plots indicated that it was not necessary to remove any outliers. Log reaction time, ordinal number and trial number were centralised, so that the value of each data point was expressed as its difference from the mean for that value. For reaction times, the centralised value for a given word was taken as the mean for that word, when it was correct, across all participants. Since mixed effects models usually failed to converge using these centralised values, the values were also standardised, such that their scales were more comparable.

5. Results

Figure 5 shows log RT as a function of trial number for the remaining 24 subjects. The plot suggests that there is a steady slowing effect for some participants, such as A19, but no effect, or indeed a speeding effect, for others, such as A22. Since previous studies suggested that lag should have no effect on (Howard et al., 2006; Oppenheim et al., 2010; Belke, 2013; Hoedemaker et al., 2017), subsequent analyses only make use of ordinal number.

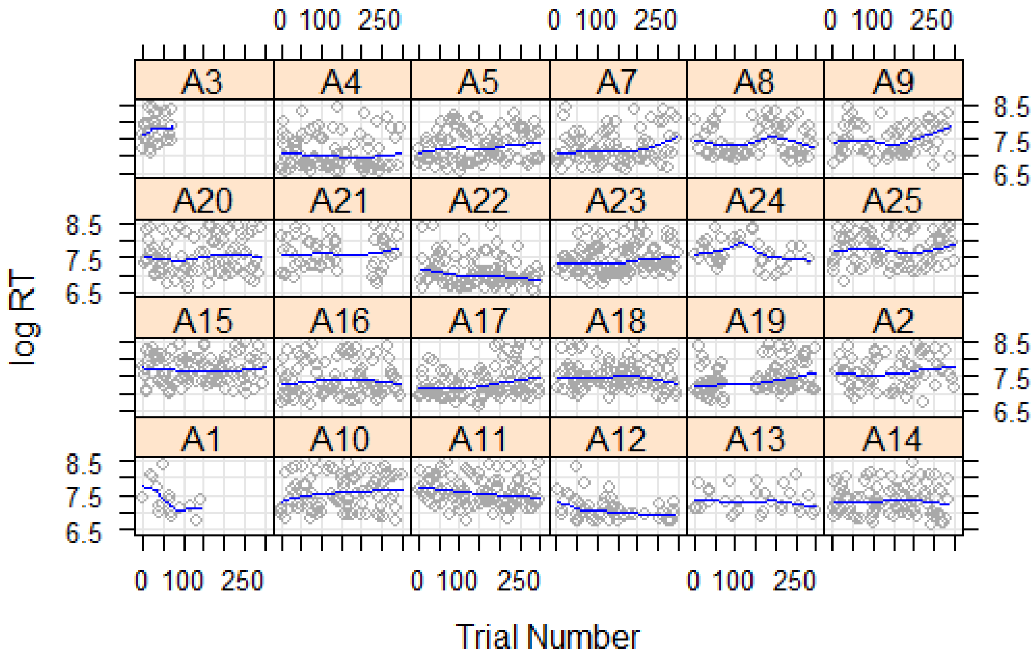

Fig. 5. Log-transformed and centred RT plotted by Trial number, to inspect slowing throughout the experiment.

5.1 Linear analysis

The first research question expected the dataset to replicate the effect from previous experiments using the continuous naming paradigm: a main linear effect of ordinal position on RTs, independent of the effect of trial position. Therefore, an initial mixed effects model (Baayen, Davidson & Bates, 2008) was implemented in *R* (2012) using the package *lme4* (Bates, Maechler, Bolker, & Walker, 2015). The maximal random effect structure supported by the model had trial number varying across participants and items, a fixed effect of ordinal position and trial position, and random effects for subject and word. The mixed effects model therefore repeated the structure of Alario & Moscoso Del Prado Martín (2010; *HH-model 2*). Following previous analyses, the ordinal number included correct responses only. This is does not entail a commitment to the hypothesis that only correct responses have an effect: rather, it is a default option that maintains consistency with the literature. The results appear in the column Ordinal number (correct) in Table 3.

In addition, the second research question was concerned with whether omissions and errors, as well as correct responses, contribute to any CSI effect. In practice, this means that the position of an item within its semantic category (henceforth its *ordinal position*) can be determined by either (a) only trials that were correct responses, or (b) the number of preceding trials of *any response* (i.e. including omissions and error). Thus, two versions of ordinal number were generated to be specified as fixed effects: *ordinal number (correct)* and *ordinal number (all)*. Table 3 shows that naming latencies did not follow the hypothesised effect of ordinal number in either case.

	Estimate	Std. Error	*t*	*p*
Intercept	7.46	0.48×10^{-2}	156.707	<0.001
Trial number	2.4×10^{-2}	0.15×10^{-2}	1.685	0.104
Ordinal number (correct)	8.7×10^{-3}	0.21×10^{-2}	-0.426	0.67
Ordinal number (all)	1.5×10^{-2}	0.23×10^{-2}	0.649	0.516

Table 3. Results of the mixed-effects model with random effects for Trial number, Ordinal number (correct) and Ordinal number (all). The two varieties of Ordinal number respectively discard or include the erroneous or omitted trials when calculating the within-category position of the word.

An ANOVA between the two mixed effects models in Table 3, Ordinal number (correct) versus Ordinal number (all), revealed that the fit of the model using Ordinal number (all) did not significantly differ. Figure 6 shows, for each participant, the log RT as a function of ordinal number (with both values centred).

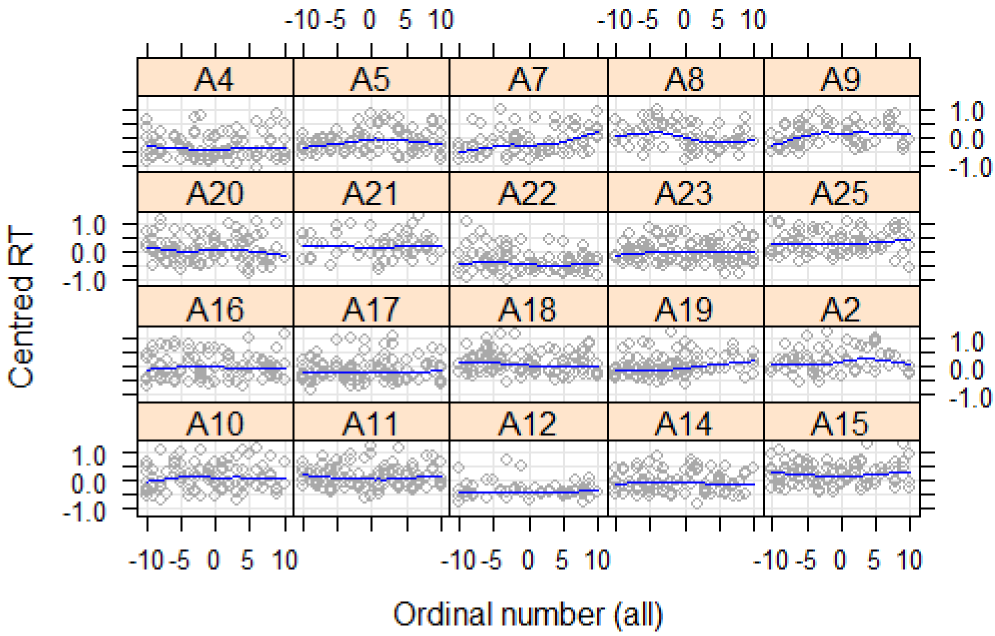

Fig. 6. Log-transformed and centred RT plotted by Ordinal number (all), which is the within-category position of the item, including omissions and errors. Numbers are the anonymised codes for each participant.

5.2 Intermediate discussion

The results of the analysis showed no significant effect of ordinal position on reaction times, and thus it was not possible to proceed with the original research question assessing whether omissions contribute to the CSI effect. The CSI effect is a robust and replicated finding, and so the present experiment was expected to comport with the results of Stielow & Belke (2015) and Schnur et al. (2006; experiment 2). However, there are several possible explanations for the absence of an effect here, relating to both the participants and aspects of the experiment. These are summarised in turn below.

Effect of aphasia: It is in principle possible that some additional aspect of acquired aphasia means that a cumulative effect is not generated in a continuous naming task, whereas it is generated in a blocked naming task. However, Belke (2013) suggests that, if anything, the opposite should be true:

PWA ought to display an interference effect in a blocked naming task in particular, since this is the paradigm that requires top-down competitive selection, a process that is frequently impaired in people with aphasia. This hypothesis would not extend to the continuous naming task, which does not require top-down selection between a small number of competing alternatives. Nonetheless, the absence of top-down selection does not suggest that PWA should differ from healthy participants

Another possibility is that the participants' level of impairment exceeded that found in predecessors. As was noted earlier, the present participants produced a far larger proportion of errors than in Schnur's dataset (73% versus 28%), despite having the same cutoff time of 5000ms. This difference in naming success is likely to reflect two differences in the present experiment: a relative increase in the average severity of aphasia, and an increase in the difficulty of the items to be named. This is reflected in the patient's average Aphasia Quotient according to the Western Aphasia Battery (Kertesz, 1982); on average the present participants scored 63.2 versus Schnur et al.'s 71.5, though both scores qualify as "moderate" aphasia. The severity of the patient's aphasia could be responsible for the high rate of errors and omissions. Although this feature presented the opportunity to study the effect of omissions, it appears that the large amount of missing data meant that the final dataset lacked sufficient power.

Semantically diverse categories. The stimuli in the present data comprised 5 semantic categories of 21 items each. This deviated from the general pattern of previous experiments, which favour smaller semantic categories, both in terms of the number of items within them and the semantic specificity of the category. For example, Howard et al.'s (2006) categories were the converse of the present experiment – 21 semantic categories of 5 items each. Alario & Moscoso Del Prado Martín's reanalysis suggests the possibility that an effect is detectable up to 10 items. It is not reasonable to expect that the cumulative interference suddenly cuts off at a specified number of items; but on the other hand the effect found in the original experiment cannot plausibly extend, linearly, to indefinitely many items.

One unavoidable result of larger categories is that the stimuli needs to include items that are less frequent, and therefore more likely to induce an omitted or erroneous response. Picture-naming tasks are usually designed using common, everyday objects that are easily identifiable. This was the aim in the present stimuli, but it was necessary that some more exotic items were included. In the final dataset, *car* was successfully named 43 times, while more exotic items such as *kiwi* and *pomegranate* were only correctly named twice each.

Breadth of semantic categories. In addition to including infrequent items, the categories also had to be semantically broad or general. One category in the present dataset, *animals*, conflates the categories *farm animals* and *zoo animals* from Howard et al.'s original stimuli. Several studies, in a variety of experimental paradigms, converge to suggest a "graded effect" of semantic closeness, according to which the degree of semantic relatedness predicts the strength of the interference effect. In a picture-word interference task, a closer semantic relation between target word and distractor amplifies the interference effect (Vigliocco et al. 2002; Vigliocco et al., 2004; Costa et al., 2005; Aristei & Abdel Rahman, 2013), and the effect is also demonstrated within a blocked naming task (Navarette et al., 2012; experiment 3a & 3b).

The effect of semantic similarity has also been explored in the continuous naming paradigm, with varying results. In support of an interference effect within narrower categories, Rose & Abdel

Rahman (2017) manipulated a continuous naming task to explicitly control for semantic closeness, and found that semantic closeness predicted a greater interference effect on subsequent RTs. Moreover, only those items that had a large degree of semantic overlap contributed to the effect.

However, not all previous literature has argued that categories must be narrower. Alario & Moscoso Del Prado Martín (2010) combined Howard et al.'s original categories into "supracategories", which were motivated on the basis of their sharing more "global" semantic features. Their reanalysis suggested that supracategories had an effect over and above the effect of original categories, indicating the possibility that a broad semantic relation can contribute to an interference effect.

5.3 Reanalysis with narrower categories

On the basis of these points, a subsequent analysis can investigate the possibility that the semantic categories used in the present experiment were too general. New, narrower versions of the semantic categories were created based on their sharing a particularly close semantic relation, presented in Table 4. Several of the categories are partly similar or identical to the original categories in Howard et al. (2006).

Category	Items	Number
Farm animals	horse, rabbit, cow, pig, sheep	5
Road vehicles	car, bus, van, motorbike, ambulance, fire engine, jeep	7
Zoo animals	tiger, giraffe, elephant, zebra, lion, kangaroo, polar bear	7
Furniture	cup, knife, fork, spoon, mug	5

Table 4: The recreated, narrower semantic categories used in the present reanalysis. The name of the category indicates the motivation for including the items.

It is possible that a large number of items intervenes between two items from these categories. Since irrelevant items do not affect cumulative interference (e.g. Oppenheim et al., 2010), this is not expected to matter, and is not accounted for. However, this ignores the possibility that items that have been omitted from these categories do in fact have some effect. The reanalysis is therefore far from perfect but should be indicative of the presence of at least some semantic interference.

This subset of data comprised 560 trials from 21 participants. The data were recoded according to the new categories, so that the new value of ordinal number was taken as the position of the item in the new category, up to a limit of either 5 or 7. The same mixed effects model was applied to the new data: ordinal and trial number were assessed as fixed effects, while subject and item were random effects. The analysis yielded the results displayed in Table 5. The reformed categories did not lead to the resolution of the CSI effect: neither of the fixed effects - ordinal number or trial number - reached significance in the new dataset.

	Estimate	Std. Error	*t*	*p*
Intercept	3.01×10^{-3}	6.14×10^{-2}	0.049	0.962
Ordinal number	8.19×10^{-3}	1.48×10^{-2}	-0.555	0.587
Trial number	3.08×10^{-4}	2.05×10^{-4}	1.502	0.134

Table 5: The results of the mixed effects model performed with the recreated semantic categories. The structure of the model is identical to that reported in Table 3.

6. Discussion & Conclusion

The aims of the present study were (a) to establish the Cumulative Semantic Interference effect in continuous naming task, with a new group of participants with Aphasia, and (b) to investigate the effect of omissions on the expected effect, which would be taken to indicate whether the phenomenon is mediated at the conceptual or lexical level.

The CSI effect has been demonstrated in a continuous naming task (e.g. Howard et al., 2006) and in People with Aphasia (e.g. Belke & Stielow, 2015; Schnur et al., 2006), and it was expected to be replicated in the present dataset. That is, the ordinal number of the word within its semantic category was expected to have a significant effect on the increase in response latency of that word, over and above the previous word in the category. No such effect was detected, and some possible reasons for the failure to find a significant effect were outlined in the intermediate discussion, including the severity of the participants' aphasia, which depleted the number of responses that could be used as data points in final analysis, perhaps reducing the dataset's power to such an extent that a significant effect could not be detected.

Ideally, an experiment using a continuous naming task that investigates the presence of a cumulative semantic interference effect in people with aphasia would make fewer alterations from the previous incarnation of the experiment – Howard et al.'s (2006). A successful replication of that experiment, using the same experimental setup with 9 PWA, was performed by Belke & Stielow (2015). The experimental setup used here has the potential to assess at least 3 core aspects of the CSI effect:

(1) *The number of category items to which the CSI effect extends.* Clearly the cumulative trend established in Howard et al. (2006) cannot extend linearly - although, without modification, models that implement a strengthening-only mechanism (HM & BM) would in fact generate this result. An experiment that extends the number of items for which an effect is found would shed light on the limits on the cumulative slowing effect.

(2) *The breadth of semantic relation that is possible while maintaining an interference effect.* The literature is inconsistent on the degree of relatedness that is required to generate the effect (Alario & Moscoso Del Prado Martín, 2010; Rose & Abdel Rahman, 2017). This may be a result of the difficulty of detecting the increasingly subtle interference effects that could, in theory, hold between distantly related items. The much larger semantic categories used in the present experiment could be used with a healthy control group to ensure that they are sufficient to generate the effect, before moving on to the more difficult case of PWA.

Ultimately, the large number of experimental changes behind the present experiment, in comparison to the background literature, may have contributed to the difficulty in assessing the more ambitious hypotheses that were set out here. An attempt at answering one of these research questions would therefore look to alter the previous paradigm more subtly, either by applying the current stimuli to healthy participants, or using a set of stimuli closer to that of Howard et al. (2006) with PWA.

Appendix

Appendix A. Some of the background tests taken prior to the experiment.

Partic -ipant No.	***WAB AQ***	***Pyramids & Palm Trees Test***	***Apraxia of speech screening***	***AoS rating***	***Wisconsin Card Sorting Test***	***WCST standard score***
P1	65.6	49/52 (94%)	No AoS	0	Invalid	N/A
P2	70.2	49/52 (94%)	Mild AoS & mild dysarthria	5	Average	101
P3	84.3	51/52 (98%)	Mild AoS, many phonological errors	3	Average	99
P4	62.6	51/52 (98%)	Not evident, but occasional phonological searching	2	Average	99
P5	65.7	50/52 (96%)	Not evident, but many phonological errors	3	Average	107
P6	23	44/52 (85%)	Not evident	0	Low Average	83
P7	36.5	47/52 (90%)	Not evident	0	Low Average	83
P8	53	48/52 (92%)	Not evident on screen, occasional groping/phonological searching	2	Borderline	78
P9	53.7	52/52 (100%)	Not evident	0	Low Average	88
P10	76.2	46/52 (88%)	Not evident	0	borderline	79
P11	74.6	50/52 (96%)	No/Mild AoS, frequent phonological errors	4	Very Superior	135
P12	44	49/52 (94%)	No/Mild AoS, frequent phonological errors	4	Average	98
P13	28.4	46/52 (88%)	Not evident	0	Failed to complete	70
P14	59.75	47/52 (90%)	No AoS	0	Average	98
P15	76	49/52 (94%)	No AoS	0	Failed to complete	70
P16	74.5	42/52 (81%)	No AoS	0	Failed to complete	70
P17	69	48/52 (92%)	No AoS	0	Average	90
P18	82.9	42/52 (81%)	No AoS	0	Average	102
P19	73.2	46/52 (88%)	No/Mild AoS	1	Low Average	87
P20	76	49/52 (94%)	No AoS	0	Average	91

P21	50.3	42/52 (81%)	No/Mild AoS	2	Average	90
P22	81.9	45/52 (86%)	No AoS	0	Very Superior	132
P23	80.4	51/52 (98%)	No AoS	0	Average	90
P24	31.3	46/52 (885)	Very mild AoS	1	Failed to complete	70
P25	89.8	51/52 (98%)	Not evident	0	High Average	114

Appendix B. Table of items used in the experiment, under their respective semantic categories

Filler	**Furniture**	**Vehicles**	**Fruit**	**Kitchen items**	**Animals**
watch	chair	Car	watermelon	peeler	horse
necklace	table	bus	kiwi	grater	rabbit
icecream	sofa	bicycle	avocado	microwave	tiger
beer	bed	train	pomegranate	scale	giraffe
rose	armchair	van	coconut	whisk	frog
pencil	chest of drawers	aeroplane	peanut	ladle	snake
clouds	wardrobe	motorbike	walnut	corkscrew	tortoise
ticket	piano	canoe	apple	cup	cow
cheque	lamp	wheelchair	orange	knife	pig
window	bath	tank	pear	fork	elephant
radio	television	balloon	strawberry	kettle	zebra
dart	clock	rocket	banana	spoon	seal
moon	toilet	submarine	raspberry	mug	shark
bread	bin	pram	grapes	pan	spider
umbrella	desk	lorry	peach	fridge	sheep
glasses	bunk	campervan	cherry	plate	lion
camera	stool	helicopter	lemon	teapot	polar bear
domino	cot	ambulance	blackberries	frying pan	camel
hammer	rocking chair	jeep	apricot	toaster	kangaroo
pizza	chest	fire engine	lime	chopping board	dolphin
computer	bench	yacht	pineapple	rolling pin	bat
postbox					
sandwich					
lipstick					
stamp					
jacket					
anchor					
suitcase					
telephone					
mountain					
toothbrush					

stick					
bucket					
pyjamas					
beach					
biscuit					
church					
razor					
tree					
castle					
candle					
baby					
nose					
book					
key					

References

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. http://dx.doi.org/ 10.1037/a0021208.

Alario, F. -X., & del Prado Martín, F. M. (2010). On the origin of the “cumulative semantic inhibition” effect. *Memory & Cognition, 38*, 57–66.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Baker, J. C., LeBlanc, L. A., & Raetz, P. B. (2008). A Behavioral Conceptualization of Aphasia. *The Analysis of Verbal Behavior*, *24*(1), 147–158.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67(1)*, 1–48.

Belke, E. (2013). Long-lasting semantic context effects on object naming are necessarily conceptually mediated. *Journal of Memory and Language, 69*, 228–256.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology, 58*(4), 667–692.

Biegler, K. A. l., Crowther, J. E., Martin, R. C. (2008). Consequences of an inhibition deficit for word production and comprehension: Evidence from the semantic blocking paradigm. *Cognitive Neuropsychology. 25*(4), 493–527.

Bloem I., Van den Boogaard S. & La Heij W. (2004), Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access, *Journal of Memory and Language 51*, 307-323.

Britt A. E., Ferrara C., and Mirman D. (2016). Distinct Effects of Lexical and Semantic Competition during Picture Naming in Younger Adults, Older Adults, and People with Aphasia. *Frontiers in Psychology, 7*(813), 1-17. DOI: 10.3389/fpsyg.2016.00813.

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

Canini, M., Della Rosa, P. A., Catricala, E., Strijkers, K., Branzi, F., Costa, A., & Abutalebi, J. (2016). Semantic Interference and Its Control: A Functional Neuroimaging and Connectivity Study. *Human Brain Mapping*, *37*(11), 4179-4196. DOI: 10.1002/hbm.23304

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*, 177–208.

Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1372–1384.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell G. S, & Reich P. A. (1981). Stages in Sentence Production - an Analysis of Speech Error Data. *Journal of Verbal Learning and Verbal Behavior, 20*(6), 611–629.

Dell, G.S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic patientes. *Psychological Review, 104*, 801–838.

Dell G. S. & Sullivan J. M. (2003). Speech Errors and Language Production: Neuropsychological and Connectionist Perspectives *Psychology of Learning and Motivation - Advances in Research and Theory*. 44: 63-108. DOI: 10.1016/S0079-7421(03)44003-6.

Foygel, D. & Dell, G. S. (2000). Models of Impaired Lexical Access in Speech Production. *Journal of Memory and Language*, *43*(2), 182-216. DOI: 10.1006/jmla.2000.2716

Francis, W.S. (2014). Repetition priming in picture naming: sustained learning through the speeding of multiple processes *Psychon Bull Rev 21,* 1301. DOI: 10.3758/s13423-014-0610-9

Friedmann N., Biran M., Dotan D. (2013). Lexical retrieval and its breakdown in aphasia and developmental language impairment, in *The Cambridge Handbook of Biolinguistics*, Eds: Boeckx C., Grohmann K. K., 350–374. Cambridge, MA: Cambridge University Press.

Goodglass, H. & Wingfield, A. (1997). “Word-finding deficits in aphasia: Brain-behavior relations and clinical symptomatology”. In *Anomia: Neuroanatomical and cognitive correlates*, Eds: Goodglass, H. and Wingfield, A. 3–27. New York: Academic Press.

Harley, T. A. (1993). Connectionist approaches to language disorders. *Aphasiology, 7*(3), 221–249.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition, 100*, 464–482.

Howard, D., & Patterson, K. (1992). Pyramids and palm trees: A test of semantic access from pictures and words. Bury St. Edmunds, UK: Thames Valley Test Company.

Hsiao E. Y., Schwartz M. F., Schnur T. T., & Dell G. S. Temporal characteristics of semantic perseverations induced by blocked-cyclic picture naming*. Brain and Language. 108,* 133-44. PMID 19138794 DOI: 10.1016/j.bandl.2008.11.003.

Jefferies E., Patterson K., Lambon Ralph M. A. (2008) Deficits of knowledge versus executive control in semantic cognition: Insights from cued naming. *Neuropsychologia, 46,* 649–58.

Kertesz, A. (1982). Western Aphasia Battery (1st ed.). San Antonio, TX: The Psychological Corporation.

Kohn S. E. & Goodglass H. (1985). Picture-naming in aphasia. *Brain & Language*; *24*(2), 266-83.

Kroll J. f. & Stewart, E. (1994) Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations*. Journal of Memory and Language, 33*. 149–174.

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

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access*. Proceedings of the National Academy of Sciences, 98,* 13464–13471.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production*. Behavioral and Brain Sciences, 22,* 1–75.

McCarthy R. A. & Kartsounis L. D. Wobbly words: Refractory anomia with preserved semantics. *Neurocase. 2000, 6*:487–497.

Morris, J., Howard, D. & Buerk, F. (2014.) SemaFoRe: Comparing word retrieval treatments for aphasia via a randomised crossover trial. Clinical Aphasiology Conference 2014, 44th, St. Simons Island, GA : May 27-June 1, 2014 (81).

Navarrete, E., Mahon, B. Z., & Caramazza (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica, 134(*3), 279–289.

Navarrete E., Del Prato P., & Mahon B. Z. (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology, 3*:38. [PubMed: 22363309]

Navarrete, E., Mahon, B. Z., Lorenzoni, A., & Peressotti, F. (2016). What can written-words tell us about lexical retrieval in speech production? *Frontiers in Psychology, 6*.

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.* 227–262.

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J.-M., & Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PLoS One, 9.*

R Development Core Team (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing (ISBN 3-900051-07-0).

Riley, E., McMahon, K. L., & de Zubicaray, G. (2015). Long-lasting semantic interference effects in object naming are not necessarily conceptually mediated. Frontiers in Psychology, 6, 578.

Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia naming test: Scoring and rationale. *Clinical Aphasiology, 24*, 121-134.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Rose, S. B., & Abdel Rahman, R. (2016). Cumulative semantic interference for associative relations in language production. *Cognition, 152,* 20–31.

Runnqvist, E., Strijkers K., & Costa A. (2014). Bilingual word access. *The Oxford Handbook of Language Production.*

Schnur, T. T. (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language, 75,* 27–44. http://dx.doi.org/10.1016/j. jml.2014.04.006.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language, 54*(2), 199–227. http://dx.doi.org/10.1016/j. jml.2005.10.002.

Spalek, K., Damian, M. F. & Bölte, J. (2013). Is lexical selection in spoken word production competitive? Introduction to the special issue on lexical competition in language production, *Language and Cognitive Processes, 28*(5), 597-614. DOI: 10.1080/01690965.2012.718088.

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 of the United States of America*. 106: 322-7. PMID 19118194 DOI: 10.1073/pnas.0805874106 .

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Vol. Eds.), Progress in the psychology of language: Vol. 1. London: Lawrence Erlbaum.

Stielow, A., & Belke, E. (2015). Interaction of Executive Functions and Word Retrieval in aphasic patients and healthy controls - a Psycholinguistic Investigation. *Stem-, Spraak- en Taalpathologie, 20*(Suppl.), 200-202.

Vigliocco, G., Vinson, D. P., Lewis, W., & Garret, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*, 422–488.

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language, 33*, 332-356.

Wilshire C. E. & McCarthy R. A. (2002) Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia*. Cognitive Neuropsychology. 19*, 165–186. [PubMed: 20957536]