The dual origin of lexical perseverations in aphasia: Residual activation and incremental learning

Christopher R. Hepner[a*] and Nazbanou Nozari[a]

[a] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

[*] Corresponding author. Department of Psychology, BH 446J, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA.
Email address: chepner@andrew.cmu.edu (C.R. Hepner).

Abstract

Lexical perseveration, the inappropriate repetition of a previous response, is common in aphasia. Two underlying mechanisms have been proposed: residual activation and incremental learning. Previous attempts to differentiate the two have relied on experimental paradigms that encourage semantically related errors and analysis techniques designed to detect perseverations over short distances, resulting in a bias towards detecting short-lag, semantically related perseverations that both mechanisms can account for. Two key predictions that differentiate these accounts remain untested: only residual activation can explain short-lag, semantically unrelated perseverations, whereas only incremental learning can explain long-lag, semantically related perseverations. In this paper, we used a large set of picture naming trials and a novel analysis technique to test these key predictions in a multi-session study involving six individuals with aphasia. We found clear evidence for both mechanisms in different individuals, demonstrating that either one is sufficient to cause perseveration. Importantly, perseverations due to residual activation were associated with more severely impaired systems than those due to incremental learning, suggesting that a certain degree of structural and functional integrity was necessary for incremental learning. Finally, the results supported a key prediction of the incremental learning account by showing perseverations over longer lags than have previously been reported.

Keywords: perseveration, aphasia, residual activation, incremental learning, language production

1 Introduction

Perseveration, or the repetition of a previous response when it is not appropriate for the current stimulus, is observed in a range of neurological disorders (Sandson & Albert, 1984). For example, frontal lesions are associated with perseverations in both verbal and non-verbal tasks. Individuals with such lesions are often impaired in tasks like the Wisconsin Card Sorting Test (Grant & Berg, 1948), where they produce incorrect responses that would have been correct in the previous phase of the test (e.g., continuing to sort the cards by color when they should now be sorted by number (Milner, 1963; Petrides & Milner, 1982). Perseverations of various kinds have also been reported in conditions as diverse as schizophrenia (Freeman & Gathercole, 1966), dementia (Bayles, Tomoeda, Kaszniak, Stern, & Eagans, 1985; Fuld, Katzman, Davies, & Terry, 1982; Miozzo, Fischer-Baum, & Caccappolo-van Vliet, 2013; Shindler, Caplan, & Hier, 1984), Parkinson's disease (Lees & Smith, 1983), traumatic brain injury (Fischer-Baum, Miozzo, Laiacona, & Capitani, 2016), and autism (Boucher, 1977).

One domain in which perseveration as a consequence of brain damage has been consistently observed is language production. More than half of individuals with aphasia produce these errors in tasks like picture naming and single-word repetition (Albert & Sandson, 1986; Allison & Hurwitz, 1967; Shindler et al., 1984; Yamadori, 1981). Moreover, the problem is not limited to a single modality, and has been reported in both spoken and written language (e.g., Buckingham, Whitaker, & Whitaker, 1979; Fischer-Baum & Rapp, 2012; Moses, Nickels, & Sheard, 2004; Nozari, 2019). Language also provides an excellent context for understanding the origins of perseverative behavior, because this behavior can be observed in simple tasks like picture naming that do not involve complex mental operations like rule switching. Moreover, since one can present a large number of pictures in a picture naming task without repeating any of the stimuli, it is easy to administer a task in which repetition is never prompted by the design. Finally, it is possible to trace a repetitive response to its origin in such tasks and to investigate the relationship between the response and the current target (e.g., their similarity) to further shed light on to the causes of perseverative behavior. For these reasons, we focus this study on an investigation of the origin of perseverations in language production.

Within the domain of language production, perseverations can occur at different levels of processing. *Lexical perseverations* involve the repetition of an entire word—for example, responding "dog" to a picture of a cat after saying "dog" in a previous trial. Segmental perseverations, on the other hand, involve the repetition of segments instead of whole words. In written production, the segments are letters, e.g., the perseverated L in "under" → UNDE<u>L</u> after MO<u>L</u>DE<u>L</u> (e.g., Fischer-Baum & Rapp, 2012, 2014). In spoken production, the segments are phonemes, e.g., the perseverated /b/ in car → "<u>b</u>ar" after "<u>b</u>utter" (e.g., Buckingham, Avakian-Whitaker, & Whitaker, 1978; Moses et al., 2004). Given the frequency with which they occur and their close relationship to specific underlying deficits, perseverations have been the target of both clinical interventions and theoretical investigations of the structure and operation of the cognitive processes involved in language production.

Two theoretical accounts have been proposed to explain perseverations in language. A classic account posits that perseverations are caused by the *residual activation* of the previous response (Cohen & Dehaene, 1998). In recent years, a competing account has surfaced, which attributes perseverations to *incremental learning* in the language production system (Oppenheim, 2018; Oppenheim, Dell, & Schwartz, 2010). The current paper is focused on disentangling these two mechanisms.

1.1 Residual activation vs. incremental learning

In normal (i.e., non-perseverative) production, semantic features of the concept to be named activate the corresponding lexical item, along with semantically related items (Figure 1a). Once produced, a lexical item is suppressed so that it will not be produced repeatedly (Dell, Burger, & Svec, 1997; Nozari, Dell, Schneck, & Gordon, 2015). This suppression prevents perseverations on subsequent trials. The most obvious explanation for perseverations is thus that they occur when the previous response has not been sufficiently deactivated and has retained enough activation to overtake the current target. Although the persistent activation of the previous lexical item would be augmented by input from semantic features shared with the current target (as in Figure 1c), perseverative errors are possible even when the items are not semantically related (as in Figure 1d). This explanation has been referred to as the *residual activation* account (Cohen & Dehaene, 1998). Because activation usually decays rapidly, perseverations resulting from residual activation are only expected to occur over short distances, as perseverative errors are generally found to do. Cohen and Dehaene (1998) observed that, across tasks (picture naming and single-word reading) and perseveration types (lexical and segmental), the likelihood of repeating an item decreased exponentially over successive trials. To summarize, then, the residual activation account predicts perseverations over short lags, in response to both semantically related and unrelated targets.

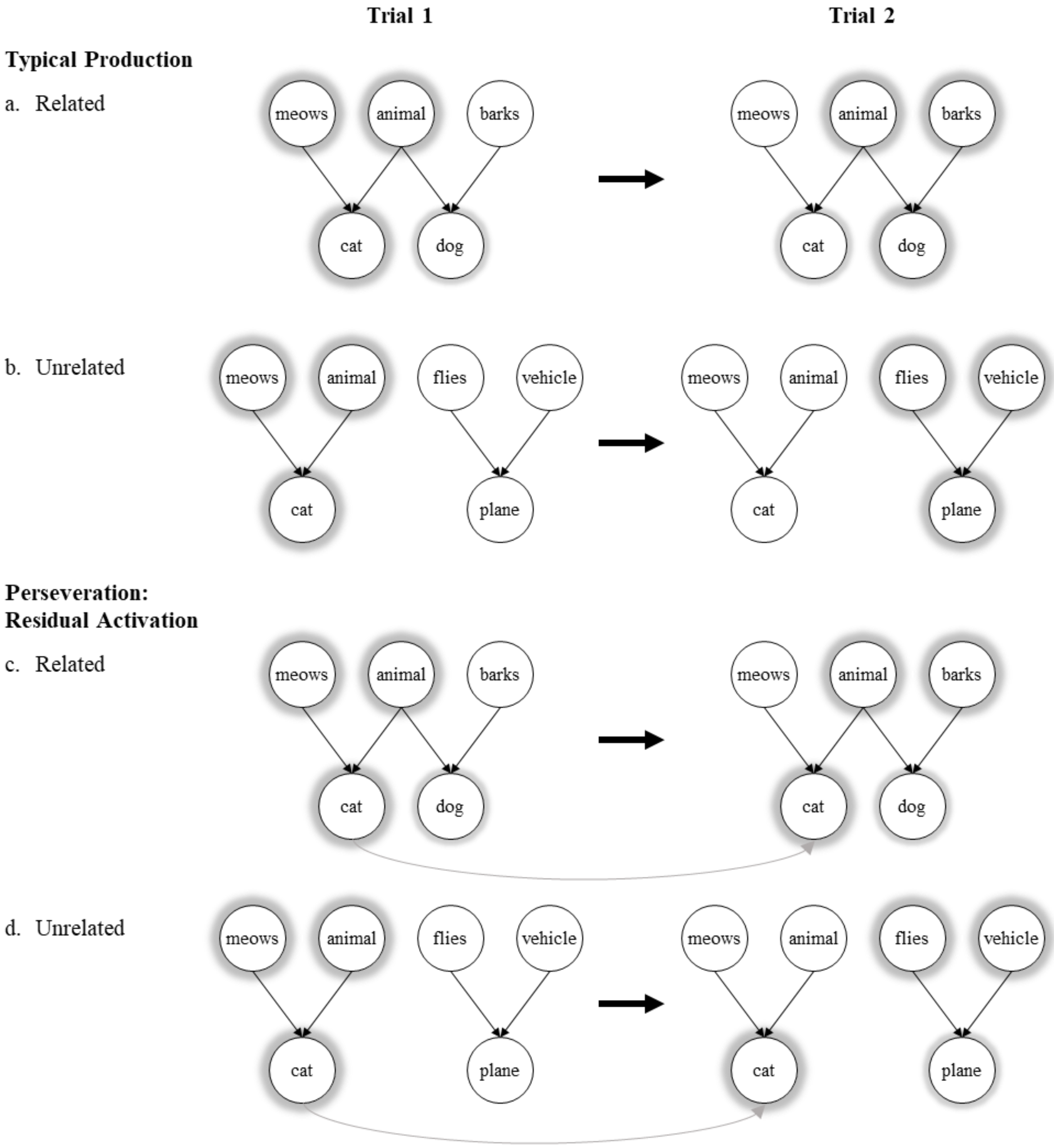

Figure 1. *Activation in semantic and lexical units in successive trials in typical production (a-b) and perseveration due to residual activation (c-d). In typical production, a lexical item is suppressed after it is produced on the first trial. If the next trial is semantically related, the previous item gains some new activation through shared features, but rarely enough to overtake the target (a). If the next trial is unrelated, the previous item's activation remains close to zero (b). In perseverations, according to the residual activation account, the lexical item* cat *retains some residual activation after "cat" is produced in response to a picture of a cat in the first trial. If the target in the second trial is semantically related (e.g., "dog"),* cat *will receive*

additional activation from shared semantic features, potentially resulting in a semantically related perseveration (c). However, even without this additional input, residual activation alone may be enough to cause semantically unrelated perseverations, as in (d). The circumference of the halo indicates the level of activation.

Residual activation was the dominant account for perseverations until a new model emerged to explain semantic interference in production. This interference manifests as greater difficulty in naming a picture in the context of semantically related compared to unrelated pictures (e.g., Belke, Meyer, & Damian, 2005; Oppenheim, 2018; Schnur, 2014; Schnur, Schwartz, Brecher, & Hodgson, 2006; see Nozari & Pinet, 2020 for a review). The residual activation account attributed this phenomenon to the same mechanism that caused perseverations: the residual activation of the previously named item interfered with the selection of the current item. This was more likely when the previous and current responses were semantically related because the shared semantic features would send additional activation to the previously selected item, further contributing to the probability of its perseveration. However, the residual activation account ran into problems when it was demonstrated that the semantic interference effect can persist even when intervening items are inserted between two semantically related pictures (see Schnur, 2014 and references therein). The persistence of interference over time and several intervening items pointed to a more lasting mechanism, like a structural change. This led to the proposal of *incremental learning*, a mechanism that made small but cumulative changes to the language production system as a result of naming a given picture (Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Oppenheim et al., 2010; Vitkovitch & Humphreys, 1991; Vitkovitch, Humphreys, & Lloyd-Jones, 1993).

According to the incremental learning account, naming a picture of a cat as “cat” strengthens the connections between the lexical item *cat* and its semantic features. A version of the incremental learning account (Oppenheim, 2018; Oppenheim et al., 2010) additionally posits that the strengthening of the target weights is accompanied by a weakening of the weights between the competitors and the semantic features that activated those competitors (implemented via the delta rule). Whether or not the weights of the connections to competitors are decreased, a recently produced target is more likely to be produced on a subsequent trial if some of its semantic features are activated by the new target, as depicted in Figure 2a. These changes are hypothesized to be structural and thus long-lasting, enabling them to account for the robustness of semantic interference to the passage of time and the insertion of unrelated items (Hsiao, Schwartz, Schnur, & Dell, 2009; Schnur, 2014). Note that the situation is very different for unrelated competitors. In normal (or only mildly damaged) systems unrelated competitors are rarely activated along with the target. There will thus be no reason for weights to be adjusted (Figure 2b). In more severely damaged systems, it is not unusual for unrelated items to be activated along with the target, leading to the production of lexical errors that bear no relationship to the target (e.g., Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). The source of this activation is not clear, but it may be caused by the random activation of semantic features unrelated to the target. Such a situation is depicted in Figure 2c. In these cases, the high activation of the unrelated item in the first trial could trigger weight adjustments as depicted in the figure. However, these changes would not be sufficient to cause perseverations. Strengthening the weights of the connections between the target and its semantic features only confers an advantage to the lexical item when those features are activated. If the next target is semantically unrelated, the semantic features of the previous item will not be activated, so the strengthened connection weights will not contribute to the activation of its lexical representation and no perseveration would be expected. To summarize, the incremental learning model can account for both short- and log-lag perseverations, resulting from structural changes to the connections, as long as the new target is semantically related to the item being perseverated.

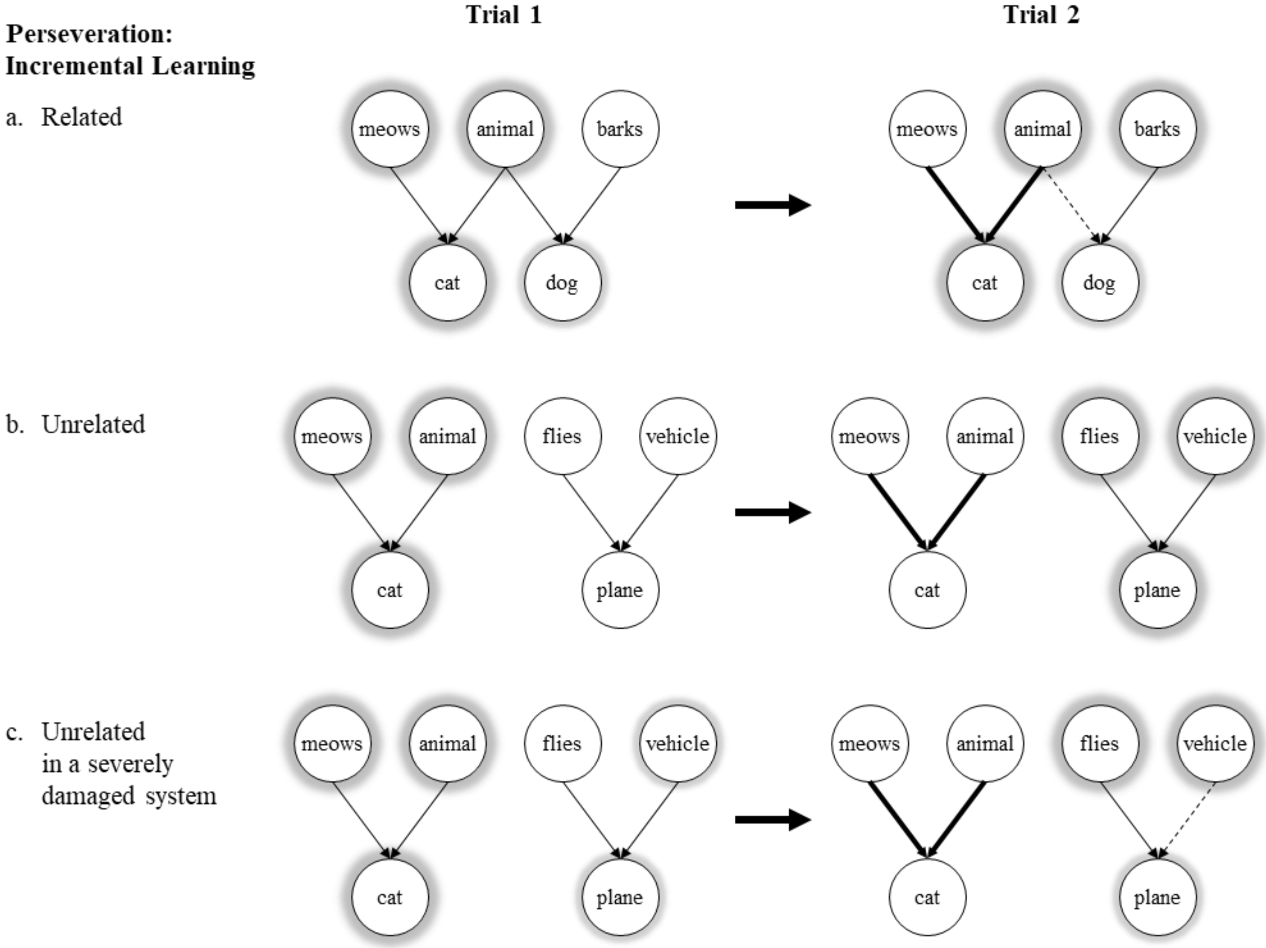

Figure 2. Perseverations under the incremental learning account. (a) After producing "cat" in response to a picture of a cat, the connections between cat *and its semantic features are strengthened, while the connections between those features and semantic competitors like* dog *are weakened. When naming a picture of a dog in the next trial,* cat *thus receives more activation from shared features than* dog*, potentially resulting in a semantically related perseveration. (b) By contrast, producing "cat" has no effect on the connections between semantically unrelated words like* plane *and their semantic features, since there is no overlap between the two. In this case,* cat *would receive no additional activation in the second trial. (c) In damaged systems, some semantic features and lexical items unrelated to* cat *may be activated in the first trial due to random noise (sometimes leading to unrelated lexical errors in trial 1). This random activation may induce weight changes that are not normally observed in non-damaged systems. However, even if such changes take place, they would not result in a semantically unrelated perseveration, because the semantic features of* cat *are not activated by the new unrelated target, so the strength of their connections to the lexical item* cat *is irrelevant. The circumference of the halo indicates the level of activation. The thicknesses of the connections reflect weight changes, with thicker connections denoting strengthening and dashed connections denoting weakening.*

1.2 Which account is right?

Table 1 summarizes the predictions of the residual activation and incremental learning accounts. As described above, both can account for semantically related perseverations over short distances. However, their predictions do differ in two key areas. First, only the residual activation account allows for completely semantically unrelated perseverations, because the persistent activation of a previous target can cause a perseveration even in the absence of input from shared semantic features. According to the incremental learning account, on the other hand, perseverations critically depend on the activation of shared semantic features, so semantic similarity is vital for generating perseverations. In this sense, the incremental learning account is more restrictive than the residual activation account. On the other hand, the second difference between the predictions of the two models is that only the incremental learning mechanism can account for long-lag perseverations. The structural changes posited by this account can persist over much longer intervals than the transient activation of individual nodes posited by the residual activation account. In this sense, the residual activation account is the more restrictive of the two.

Table 1. Summary of the key predictions of the residual activation and incremental learning accounts. The gray cells indicate the critical predictions that differentiate the two models.

	Semantically Related		**Unrelated**	
	Short-Lag	*Long-Lag*	*Short-Lag*	*Long-Lag*
Residual Activation	✓	✗	✓	✗
Incremental Learning	✓	✓	✗	✗

As can be seen in Table 1, neither model has a clear advantage in its ability to explain a wider range of empirical phenomena. Residual activation can account for perseverations over a wider variety of target-error relationships, whereas incremental learning can account for perseverations over a wider range of lags. Accepting or rejecting either of these accounts is thus a matter of determining the characteristics of perseverations in aphasic speech that need to be accounted for. For example, if it turned out that individuals with aphasia rarely perseverate over long lags, but they produce both semantically related and unrelated perseverations, then the residual activation account would provide the best explanation for the observed pattern of perseveration errors. If, on the other hand, it turned out that long lag perseverations were common in aphasia, and semantically unrelated perseverations were rare on unrelated targets, then the incremental learning account would provide the best explanation.

Although there have been a number of studies of perseveration in language production, which have contributed substantially to our understanding of the nature of linguistic representations (e.g., Fischer-Baum, McCloskey, & Rapp, 2010; Fischer-Baum & Rapp, 2014), there has been little research focused on directly testing the predictions of the residual activation vs. incremental learning accounts in aphasia. One exception has been a study by Hsiao et al. (2009), who used a blocked cyclic naming task with 6 items in each cycle to specifically test the viability of the incremental learning account. There were two key manipulations. First, the items in each block were either from the same semantic category (homogeneous blocks; e.g., dog, horse, cat, etc.) or from different semantic categories (mixed blocks; e.g., tree, cat, nose, etc.). Short-lag perseverations could thus be semantically related or unrelated, depending on the block (homogeneous or mixed, respectively). Second, trials were separated by either 1 or 5 seconds. Their findings were generally consistent with the incremental learning account: semantically related perseverations were much more common than unrelated ones, and the amount of time between trials made no difference in the distances over which perseverations occurred. However, Hsiao et al. (2009) did not conclusively address two of the key predictions that differentiate the competing accounts. On the one hand, while they did find more semantically related perseverations in homogeneous blocks, they also found above-chance rates of (presumably unrelated) perseverations in mixed blocks, which are difficult for incremental learning to explain. On the other, a real test for long-lag perseverations (i.e., over longer distances than the residual activation account would predict) has been missing. The blocked cyclic naming paradigm is ill suited for this purpose, since a small set of items is repeated over short intervals. Moreover, repeatedly naming a small set of pictures may artificially promote perseverations that are exacerbated by semantic similarity. If incremental learning is truly the underlying cause of lexical perseveration, it should be possible to observe it in a task without any demand for repeating items, and perseverations should occur over much longer lags (e.g., 10+ trials).

In short, the current body of literature on perseverations in aphasia does not paint a clear picture of their characteristics. Well-controlled studies that have manipulated the intervals between trials have found no effect of time, supporting the incremental learning account (Gotts, della Rocchetta, & Cipolotti, 2002; Hsiao et al., 2009), but, due to the limited number of items, have been unable to confirm its prediction of long-lag perseverations. Conversely, the presence of significant numbers of semantically unrelated perseverations in these and earlier studies (e.g., Hirsh, 1998) is better explained by the residual activation account. However, these two accounts need not be mutually exclusive. It may, for example, be the case that individuals with aphasia show a mixture of the error types listed in Table 1, providing evidence that a hybrid model is appropriate at the individual level. It is also possible that different individuals show distinct patterns of perseveration, compatible with one model or the other but not both. The current study was designed to distinguish between these possibilities.

1.3 Current study

As discussed above, both lexical items and segments can perseverate in the speech (or writing) of individuals with aphasia. We target lexical, as opposed to segmental, perseverations because of our focus on disentangling the residual activation vs. incremental learning accounts. Specifically, we focus on two key areas in which the residual activation and incremental learning accounts diverge: short-lag, semantically unrelated errors, and long-lag, semantically related errors. While the former can be studied using both lexical and segmental perseverations, the latter does not lend itself easily to an investigation of long-lag perseverations, even though incremental learning accounts of phonological interference, like semantic interference, have been proposed (Breining, Nozari, & Rapp, 2016, 2019; Nozari, Freund, Breining, Rapp, & Gordon, 2016). The reason is that the number of segments in language are limited, and repeating the same segment becomes inevitable after only a few words (hence the small windows chosen for such analyses; e.g., 5 trials in McCloskey, Macaruso, & Rapp, 2006). Lexical items, on the other hand, are numerous, and it is simple to create a picture naming task with a large number of items without repeating any of them. We thus chose to use lexical perseverations in order to be able to fully test the predictions of the incremental learning account. Additionally, the potentially unlimited window over which lexical perseverations can be assessed allows us to investigate the maximum lag of the effect. In a series of experiments investigating cumulative semantic interference in neurotypical adults using a continuous naming task, Schnur (2014) found no interference when semantically related trials were separated by as few as 8 unrelated trials, except when short lags (2 unrelated trials) were also present in the sequence. However, when short lags were present in at least a subset of the semantic categories, interference was observed at longer lags (8-14). A finding that there is an upper bound to the lags of lexical perseverations (at least without very short-lag semantically related items as reminders) would place constraints on the degree to which the changes to connection weights caused by incremental learning persist.

We chose simple picture naming tasks to avoid repeating items within a session or introducing the extra complexity due to task switching, arbitrary rules, and demands on working memory associated with many of the tasks used to probe perseverative behavior, like the Wisconsin Card Sorting Task (Grant & Berg, 1948), or even blocked cyclic naming (Belke & Stielow, 2013). The items in the task were distributed so as to enable us to test the key predictions summarized in Table 1. There are two critical types of perseveration errors: long-lag, semantically related perseverations and short-lag, unrelated items. The inclusion of semantically related items with short lags between them, as previous studies using blocked naming paradigms have done, would be undesirable, as perseverations on these trials would be equally compatible with both theoretical accounts. Semantically related items were thus separated by longer lags (an average of 22 intervening items), so perseverations on these trials would be unequivocal evidence for the incremental learning account. Conversely, any short-lag perseveration would necessarily be semantically unrelated and would only be compatible with the residual activation account. The design was thus optimal for disentangling the two proposed mechanisms of perseveration.

A non-trivial challenge in studies of perseveration is in disentangling perseverative and non-perseverative repetitions. For the purposes of this study, we were specifically interested in cases in which producing a response changed the state of the language production system such that the same response was repeated in a subsequent trial. We will refer to these as *true perseverations*. However, not all repeated responses are true perseverations. Some may occur simply by random chance. Imagine a computer program that labels pictures using words from a limited vocabulary. It operates according to the following rule: If the correct label is known, output that label; otherwise, output a label selected at random from the vocabulary. Each time the program encounters a picture it does not know the correct label for, there is a chance that the label it selects at random will be a repetition of a previous response and thus appear to be a perseveration. Non-perseverative responses may also occur due to a *predilection*, or a bias in the system towards producing a particular response. If, instead of selecting a label at random, the rule was for the program to output “cat” whenever it doesn’t know the correct label, it would look like the response “cat” is being perseverated. However, neither of these two cases is a true perseveration, because producing a response in one trial has no effect on the probability of producing the same response in subsequent trials. The repeated responses produced by each participant are a combination of these non-perseverative repetitions and, potentially, true perseverations. Only if the total number of repeated responses significantly exceeded the number of non-perseverative responses expected by chance would it be possible to conclude that true perseverations were present. It was thus critical that we develop a method for calculating the probabilities of non-perseverative errors. This method, described in detail in the Methods section, was inspired by McCloskey et al. (2006), but was adapted to fit the purpose of the current study so that we could distinguish between short- and long-lag perseverations.

Six individuals with chronic (> 6 months post-onset) aphasia resulting from stroke were included in the study, after initial profiling described in the Methods section. The main criterion for selection was a pattern of predominantly lexical (as opposed to segmental) errors. This pattern indicates that the main locus of the problem in the language production system is in mapping semantic features to lexical items (i.e., the first stage of processing in two-stage models of language production; e.g., Dell, 1986; Garrett, 1975), and not in mapping lexical items to segments. If they produced perseverations, these individuals would be expected to produce lexical, rather than segmental, perseverations, which makes them a suitable target population for the study. Additional tests were administered to rule out major problems in semantic processing (see Methods). We did not screen participants for specific types of lexical errors (e.g., semantic vs. unrelated) or overall disease severity. In fact, we intentionally included a range of severity and distributions of error types in order to investigate the potential relationship between specific mechanisms of perseveration and the state of the language production system. To increase statistical power, we administered a large number of trials over multiple sessions to each participant.

Recall that, because we designed the task to have relatively long lags between semantically related items, short-lag errors can only be caused by residual activation, and long-lag errors by incremental learning. Thus, if both residual activation and incremental learning mechanisms cause perseverations, we should observe above-chance perseverations over both short and long lags. Additionally, if both residual activation and incremental learning occur in parallel in a given system, we should observe both error types within the same participants. If, on the other hand, the two types of perseverations are tied to different states of the language production system, we would expect individuals to produce primarily one error type or another, but not both.

2 Methods

2.1 Participants

Six individuals (1 female, 5 male; ages 45-66) with chronic aphasia resulting from a stroke were recruited from the Snyder Center for Aphasia Life Enhancement (SCALE; https://www.leagueforpeople.org/scale) in Baltimore, MD. All participants presented with word-finding difficulty, but their language deficits were otherwise varied in severity and extent. Testing was conducted 4-10 years post-stroke. Basic information about each participant is presented in Table 2. Participant QD has been described previously in Nozari (2019). All participants were consented under an IRB protocol approved by Johns Hopkins University and received monetary compensation for their participation.

Table 2. Participant information.

Participant	**Age**	**Gender**	**Premorbid Handedness**	**Education**	**Years Post-Stroke**
QD	64	male	right	high school	10
GK	51	male	right	college	6
FJ	45	male	right	high school	5
OW	57	male	right	high school	7
CE	65	female	unknown	college	4
EG	66	male	right	unknown	4

2.1.1 Background tests and participants' profiles

Participants were administered a battery of tests to profile their language abilities. A non-linguistic semantic comprehension task (28 trials) probed semantic knowledge by requiring participants to pick which of three pictures was most closely related to a target picture. This task is, in spirit, similar to Pyramids and Palm Trees or Camels and Cactus, but improved in two ways: a) all materials have been normed to be familiar to North American participants without higher education, and b) half of the trials probe taxonomic relationships (e.g., carrot-celery) and the other half thematic relationships (e.g., key-door). Foils are semantically unrelated to the target (Nozari, 2019). Two word-to-picture matching tasks measuring lexical comprehension, one with unrelated foils and one with semantically and/or phonologically related distractors, assessed auditory word comprehension. In the task with unrelated foils, participants heard each of the 175 picture names from the Philadelphia Naming Test (PNT; Roach, Schwartz, Martin, Grewal, & Brecher, 1996) and selected the matching picture from a set of three. The purpose of this task was to assess basic lexical comprehension without the additional demand of distinguishing between close (semantic or phonological) neighbors. In the more challenging task with related distractors (30 trials), participants selected the target picture (e.g., for "plate") from a set of four that included a semantically related distractor (a fork), a phonologically related distractor (a plane), and an unrelated image (a cactus). The PNT, a picture naming task, and the Philadelphia Repetition Test (PRT), a word repetition task using the same words, each comprising 175 trials, were used to assess single-word naming and repetition abilities, respectively. A nonword repetition task comprising 60 pseudowords of increasing length was used to assess nonword repetition (Nozari, 2019). Finally, a modified version of the Category Probe task (Nozari, 2019; adapted from Freedman & Martin, 2001) was used to assess semantic working memory. Table 3 summarizes each participant's performance on these tasks. Additional information on participants' working memory and inhibitory control abilities is provided in Appendix A.

Table 3. Background language assessments.

	Sem. Comp. (28)	**Auditory Word Comp.**		**PNT (175)**		**PRT (175)**	**Nonword Repetition (60)**	**Category Probe**
		Unrel. (175)	*Rel. (30)*	*Correct*	*% Sem. Rel.*			
QD	28	174	27	134	88%	168	40	4.6
GK	27	174	28	104	96%	170	41	2.7
FJ	24	175	25	74	80%	150	7	2.7
OW	28	172	24	71	61%	155	23	1.7
CE	22	165	0	55	41%	119	11	—
EG	28	175	26	35	25%	155	44	2.4

The numbers in parentheses next to the names of the tests indicate the total number of items. Percentages of semantically related errors on the PNT were calculated relative to the total number of commission errors. Scores on the Category Probe task are reported as the length of the longest sequence at which the participant met the criterion for continuing (75% accuracy or higher), followed by a decimal indicating the proportion of the next block answered correctly (e.g., 2.5 means the maximum length at which criterion was met was at length 2, and 50% of responses at length 3 were correct). CE became unavailable for testing towards then end of the study due to health reasons, and thus did not complete the Category Probe task. Sem = semantic(ally); Comp = comprehension; Unrel = unrelated; Rel = related; PNT = Philadelphia Naming Test; PRT = Philadelphia Repetition Test.

QD has been described previously in Nozari (2019). He had good semantic comprehension and auditory word repetition abilities, relatively good nonword repetition, and relatively good semantic working memory. He had mild-moderate impairment in picture naming with a predominance of semantic errors. In the related lexical comprehension task, all of his errors were phonologically related to the target. GK had excellent semantic comprehension, but poorer picture naming and semantic working memory than QD, while his auditory repetition was comparable. FJ had a mild semantic and lexical comprehension impairment, with errors semantically related to the target in the latter. He had a more pronounced naming problem than the previous two participants, and also had some difficulty repeating words, with the majority of his errors in the PRT being nonwords. He also had the lowest score on nonword repetition. His very poor performance on that task is unlikely to reflect a hearing problem, since he only picked a phonologically-related error in place of the target once in the lexical comprehension task with related foils, even though all trials contained phonological lures that differed from the target in at most two phonemes. We suspected that his impaired ability in nonword repetition might instead reflect a problem in phonological working memory. He was thus tested with a variant of the Rhyme Probe task (R. C. Martin, Shelton, & Yaffee, 1994), in which he indeed scored very low (0.5, significantly lower that the capacity of a control sample, indicating that he could not even hold onto one phonological form; $t = 47.82$, $p < .001$).

OW showed intact semantic comprehension without close neighbors and mild impairment in the presence of related foils. All his errors in the related lexical comprehension task were semantically related to the target. While his overall accuracy on the PNT was comparable to FJ and both made predominantly semantic errors, OW made many more unrelated errors than FJ. He also had the lowest nonword repetition ability in the group after FJ. Similar to FJ, his poor performance on the repetition tasks, especially nonword repetition, cannot be easily attributed to hearing problems, because he never picked a phonological foil in place of the target in the related lexical comprehension task, pointing instead to a problem in phonological working memory. His score on the Rhyme Probe task supported this view (0.25; $t = 50.06$, $p < .001$).

CE's performance showed a profound impairment in lexical comprehension in the presence of foils—she was never able to pick the correct target. All of her errors were words that were phonologically related to the target. Not surprisingly, she also performed poorly on both the PRT and nonword repetition tasks, producing the largest number of phonologically related errors in the PRT of any participant. Together with the absence of any semantically related errors in her responses in the lexical comprehension task and her much better performance in the unrelated lexical comprehension task (which did not include phonologically similar foils), these findings are most compatible with a moderate to severe hearing impairment. Her picture naming performance was poorer than both FJ and OW, and a greater percentage of her commission errors were unrelated words. Finally, EG showed very good semantic and lexical comprehension abilities, along with a moderate lexical and nonlexical repetition impairment, but a severe naming impairment. He made the largest number of unrelated word responses in the PNT of any participant.

To summarize, semantic knowledge and language comprehension were generally relatively intact in the participants. This is important in evaluating the origin of lexical perseverations, because it allows us to rule out the possibility that participants in this study may simply not have had the knowledge necessary to identify items correctly. Word repetition varied from moderately impaired to intact. Importantly, even in the participants with the lowest PRT scores, repetition performance was still much better than naming performance. This finding is relevant to our study in that it shows that the locus of the problem in picture naming was not primarily in mapping lexical items to phonemes, but instead in mapping semantic features to lexical items; i.e., in lexical retrieval (Nozari & Dell, 2013; Nozari, Kittredge, Dell, & Schwartz, 2010). Nonword repetition scores were lower than word repetition scores across all participants, which is typical of the relationship between the two tasks. Word repetition can be accomplished both lexically and sublexically, whereas nonword repetition can only be accomplished sublexically. Sublexical processing requires strong phonological working memory, so nonword repetition is expected to be particularly poor in participants with a phonological working memory impairment (Nozari & Dell, 2013). This was true for two participants in our sample. One participant had a pronounced hearing impairment, leading to her particularly poor performance on both word and nonword repetition tasks. Since the focus of the study is on lexical perseverations in picture naming, a task which does not rely on hearing, a hearing impairment does not pose a problem.

Poor phonological working memory could potentially present a problem by introducing a second mechanism by which short-lag, semantically unrelated perseverations could occur. Previous studies have found that deficits affecting phonological processing result in perseverations of individual phonemes (Cohen & Dehaene, 1998; Moses et al., 2004). If participants with poor phonological working memory were found to produce significant numbers of short-lag, semantically unrelated perseverations, it would thus be necessary to examine those errors more closely to determine whether they were indeed caused by residual activation of lexical items rather than a mechanism (either residual activation or incremental learning) at the level of individual phonemes. This issue will be addressed in the Discussion.

Most relevant to the topic of the current study is performance on the PNT. We deliberately selected a range of performances on the PNT (ranging from 20-77% correct) to determine whether the overall severity of a naming impairment is related to the mechanism for producing lexical perseverations. All participants produced predominantly lexical (as opposed to segmental) commission errors. In addition, their much better performance on auditory word repetition compared to picture naming provided additional support for the assumption that all of them had problems with mapping semantic features to lexical items. This is precisely the kind of deficit that is expected to give rise to lexical perseverations, and thus the sample is well suited for investigating the origins of such perseverations. As expected, the pattern of lexical errors in participants with semantic-to-lexical mapping deficits shifts from predominantly semantic to predominantly unrelated word errors as disease severity increases (Dell, Schwartz, et al., 1997). This is because semantic errors reflect the preservation of at least some of the system's original architecture and functioning (e.g., it may not correctly map the four-legged pet who meows onto "cat", but it maps it onto a close neighbor like "dog"), whereas unrelated errors reflect a greater dominance of noise in the process of semantic-to-lexical mapping (e.g., instead of activating "cat" or other four-legged pets, a random lexical item like "bus" gets activated). We will investigate whether this has a bearing on the mechanisms of perseveration.

3 Materials and Procedures

Two picture naming tasks were used to elicit lexical perseverations: the PNT, with 175 items, and a large-set picture naming task, with 444 items (Nozari, 2019). The large-set task was designed with two goals in mind: to obtain a sufficiently large sample to detect perseverations over long distances, and to enable semantically related and unrelated perseverations to be detected and distinguished in the same task. Images were color photographs obtained from online repositories. Two lists with different pseudorandom orders were administered on separate occasions, for a total of 888 trials. Each list consisted of items from 22 semantic categories (e.g., fruits, vegetables, desserts, vehicles, clothing items, body parts, tools, etc.). The average gap between an item and the nearest item from the same semantic category was 22 trials (SD = 5.4). The choice to use long gaps between the semantically related items served two purposes. First, it was meant to decrease the chance of the semantic blocking effect, which has been shown to depend, to some extent, on the gap between the related items (Schnur, 2014). The second reason was to provide opportunities for observations that could adjudicate between the two theoretical possibilities (long-lag semantically related vs. short lag unrelated perseverations), as explained earlier and shown in Table 1.

The task was split into 3-7 sessions lasting up to 1.5 hours, with 45-250 trials per session, to accommodate individual participants' abilities. All participants completed the PNT first. Four participants also named all 888 pictures in the large-set picture naming task. In two cases (CE and EG), administration of the long task was discontinued after the first list (444 trials) when the participants expressed frustration with the task. Participants had up to 20 seconds to respond to each image, and no feedback except general encouragement or a reminder to respond with a single word was given. If participants used the superordinate category name (e.g., "animal") to refer to a picture, they were encouraged to provide the basic-level label (e.g., "What kind of animal is it?"). The audio from each session was recorded for offline transcription.

Both the PNT and the large-set picture naming task were coded in the same way. Participants' responses were transcribed from the audio recordings. The audio for 80% of the large-set task was double-transcribed for greater reliability. For each trial, the first completed response, if any, was identified. Subsequent responses were not included in our analyses. Responses were coded according to the scheme described in Table 4, following Nozari (2019).

Table 4. Coding scheme for the picture naming tasks.

Coding	**Examples**	**Description**
Correct	cat → "cat" cat → "cats" cat → "a lovely cat"	Response is the same lexical item as the target and is pronounced correctly or with only minor aberrations. Plural forms and elaborations containing the correct word are also counted as correct.
Semantic	cat → "dog" saddle → "horse"	Response is taxonomically or thematically related to the target.
Phonological	cat → "cap"	Response shares at least one phoneme with the target in the same position, or at least two in any position.
Mixed	cat → "rat"	Response is both semantically and phonologically related to the target.
Compound	waterfall → "watermelon"	Response replaces one of the morphemes in a polymorphemic target.
Unrelated	cat → "pencil"	Response is a lexical item with no clear semantic or phonological relationship to the target.
Nonword	cat → "zat"	Response is a nonword that may or may not be phonologically related to the target.
Non-commission	cat → no response cat → "I don't know" cat → "it walks around in my yard"	Response does not contain a clear single-word response. This includes omissions (no response produced), descriptions, circumlocutions, and unrelated comments.

3.1 Analysis of perseveration errors

A case-series design is used in this study. Statistics are computed at the level of individual participants, and convergent or divergent patterns are explored across the group. To test for the presence of true perseverations, the observed rate of repetitions must be compared to the probability of a non-perseverative repetition, taking both predilections and random chance into account. The data were coded to allow this distinction to be made.

All lexical errors, including semantic, phonological, mixed, compound, and unrelated errors, were coded as lexical substitutions. For each lexical substitution, the total number of instances of the same response within the session were counted and the closest previous instance of that response, if there were any, was identified. In addition, the relationship between repeated responses and the target was coded as semantically related or unrelated according to the same rules as in Table 4.

3.1.1 True perseverations vs. non-perseverative repetitions

As discussed in the Introduction, the crux of the problem in analyzing perseverations lies in distinguishing true perseverations from other kinds of non-perseverative repetitions (e.g., those due to random chance or a predilection). If some of the repetitions were true perseverations, then the response produced on a particular trial would depend on the responses produced in previous trials (e.g., “cat” is more likely after “cat” than “tree”). The probability of producing a repetition of “cat” in that trial would thus depend on the number of times “cat” had been produced in *previous* trials, not just the overall rate of “cat” responses. If, on the other hand, all of the repetitions were non-perseverative, then the order of the trials would have no effect on the responses produced in the current trial. Under the latter assumption, the responses prior to a given trial would represent a random sample drawn without replacement from the set of responses produced in the session.

For an individual trial, the probability of a non-perseverative repetition occurring by chance can thus be determined using a hypergeometric distribution. If x is the response produced on the current trial and X is the number of other instances of the same response within the sample (in this case, a session), X follows a hypergeometric distribution with the following probability mass function:

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Here, K is the total number of *x* responses produced in the current session excluding the current trial, *k* is the number of *x* responses on previous trials, *N* is the total number of trials in the current session excluding the current trial, *n* is the number of previous trials, and

$$\binom{a}{b} = \frac{a!}{b!\,(a-b)!}$$

is the binomial coefficient. Thus, under the null hypothesis that none of the repetitions are true perseverations, the probability $P(X > 0)$ that preceding responses contain at least one instance of *x* can be calculated as $1 - P(X = 0)$.

This can be illustrated using an example from CE. On trial 11 of 150 in her second session of the large-set picture naming task, she responded “walrus” to a picture of a dolphin, a response she produced three other times over the course of the session. The probability of at least one of these two responses appearing before trial 11 by chance could thus be calculated as follows:

$$1 - \frac{\binom{3}{0}\binom{149-3}{10-0}}{\binom{149}{10}} = .189$$

In other words, the probability of observing a non-perseverative repetition on this specific trial was .189. Note that this calculation takes both predilections and random chance into account, through the *K* and *n* parameters, respectively. If a particular response is produced many times during a session, the increase in *K* will result in an increase in the calculated probability of a non-perseverative repetition, $P(X > 0)$. For example, if CE had produced "walrus" six other times instead of three, it would have increased from .189 to .346. Similarly, the calculated probability would increase as a function of the number of preceding trials to reflect the additional opportunities for observing another instance of the same response. For example, if CE had produced "walrus" on the 100th trial instead of the 11th, it would have increased from .189 to .964.

The probability of a non-perseverative repetition was calculated, as described above, for each individual lexical substitution using the `phyper` function in R (version 3.6.3; R Core Team, 2020), resulting in a vector of probabilities $p_1, p_2, \ldots, p_n$. Aside from rare cases like the repetition of an item exactly once in the first several trials of a long session, few if any of these probabilities are likely to reach statistical significance. In general, it is not possible to reliably determine which *individual* trials are true perseverations. However, the sum of these probabilities provides an estimate of the expected number of non-perseverative repetitions, which can be compared to the observed number of repeated responses in order to determine whether true perseverations are present. Under the null hypothesis that none of the observed repetitions represent true perseverations, the number of repeated responses should be equal to the sum of the probabilities calculated above. If these probabilities were identical across trials (e.g., if the probability of a non-perseverative repetition was always 0.5), it would be possible to use a simple binomial test to determine whether the observed number of repeated responses was significantly greater than predicted under this null hypothesis. However, as discussed above, these probabilities vary considerably from trial to trial depending on the number of times the response is produced during the session, the number of preceding trials, etc. It is thus necessary to use a Poisson binomial test instead. This is a generalization of the binomial test in which the probability of success (in this case, a repetition) can vary from trial to trial. We used this test, as implemented by the `ppoibin` function in the poibin package for R (version 1.5; Hong & R Core Team, 2020), to determine whether or not the observed number of repetitions was greater than predicted under the null hypothesis—in other words, whether a significant number of true perseverations had occurred. Applied to a sufficiently large sample of responses, like the large-set picture naming task, this analysis can detect perseverations over very long distances, since it takes all previous trials in the same session into account.

3.1.2 Detecting perseverations over short and long distances

This technique can be modified to detect only short-lag, or only long-lag, perseverations. In either case, rather than calculating the probability of another instance of the same response appearing in *any* previous trial, we calculate the probability within a specific window. The choice of window involves a tradeoff: a smaller window has a higher power to detect perseverations within it, due to the smaller *n*; however, making the window too small risks missing true perseverations over slightly longer or shorter lags. This problem is not specific to lexical perseverations. In fact, the analysis described above is unique in that the statistical approach does not require an upper bound to the analysis window. In studies of letter perseverations, the relatively small number of letters in the alphabet results in a rapid increase in the probability of repeating a letter by chance as the size of the window increases. These studies have thus tended to use windows of 5 or fewer preceding trials (e.g., McCloskey et al., 2006). However, the goal of these studies was not to distinguish between short- and long-lag perseverations. Due to the design of the current study, only the residual activation hypothesis would predict, or could account for, short-lag perseverations. Since this activation decays exponentially (Cohen & Dehaene, 1998), it specifically predicts that the probability of a perseveration should be highest at lag 1 (i.e., immediately after the previous instance of the response) and drop rapidly as lag increases. A survey of the literature, however, suggests that lag 1 might be special. For example, using a blocked cyclic naming task with a limited set of items, Hsiao et al. (2009) found the chance-adjusted rate of lexical perseverations to be highest at lag 2, and not lag 1. In fact, they found evidence that lag-1 perseverations were lower than expected by chance. The most likely reason for this is the immediate suppression of words after production, which is necessary to avoid constant repetition of the same word in neurotypical systems (Dell, Burger, et al., 1997; Nozari et al., 2015). Even if this mechanism were impaired in damaged systems, the residual tendency to suppress the word immediately after its production might have prevented many of the perseverations that might otherwise have occurred at lag 1. Because words would be released from suppression after lag 1, true perseverations would thus be most likely to manifest at lag 2. Beyond that point, the probability of a perseveration due to residual activation drops steeply, consistent with the steady decrease to chance levels at lag 6 found in Hsiao et al. (2009). We thus opted for a window of 2 previous trials when testing for short-lag perseverations, and if a participant tested positive for perseveration at lag 2, extended the lag to 5 (Hsiao et al., 2009) to test the robustness of the effect in lags 2-5. We considered this window of 5 previous trials to be a reasonable upper bound for the effects of residual activation, based on the finding in Hsiao et al. (2009) that the rate of repeated responses dropped to chance levels by lag 6 in both related and unrelated conditions. In contrast, lags longer than 5 (with no upper bound) were used to test for long-lag perseverations.

To restrict the analysis to a specific window, we adjusted the parameters of the hypergeometric distribution. In the test for short-lag perseverations, n is set to the minimum of two and the number of preceding trials in the current session. K and N are also changed to exclude the two preceding and following trials, consistent with comparable analyses in previous studies (e.g., McCloskey et al., 2006). By excluding responses within this window, which would not be independent from the current trial if there were in fact true perseverations, the restriction on the range of lags can be exploited to increase the sensitivity of the test. The rest of the analysis remains unchanged. For the long-lag test, we excluded lexical substitutions that were repetitions of responses within a window of $w = 5$ previous trials. Following the same logic as the short-lag test, we calculated the chance of a non-perseverative long-lag repetition, given that no short-lag repetitions had occurred, for each of the remaining lexical substitutions. Here, the parameters of the hypergeometric distribution were as follows: for the lexical substitution of response x in i^{th} trial of the session, $n = i - w$ was the number of preceding trials, excluding the current trial and the short-lag window w, N was the total number of trials, excluding the current trial and the short-lag window, and K was the total number of x responses, excluding the current response.

4 Results

Combining the PNT and the large-set picture naming task, each participant completed 1,063 picture naming trials (619 for CE and EG). There were a total of 2,301 lexical substitutions, of which 842 were repetitions of a previous response. Figure 3 shows the distributions of semantically related (N = 525) and unrelated (N = 317) repetitions as a function of lag. There were clear distinctions between the two distributions. The distribution of unrelated was highly skewed, with a peak at very short lags. In contrast, semantically related repetitions were distributed more evenly across lags, with a peak near the mean distance between semantically related items (22). The shapes of the empirical distributions of unrelated and related repetitions thus match the predictions of the residual activation and incremental learning accounts, respectively. Our analyses focus on understanding the distribution of true perseverations of each type (unrelated short-lag vs. related long-lag) across participants.

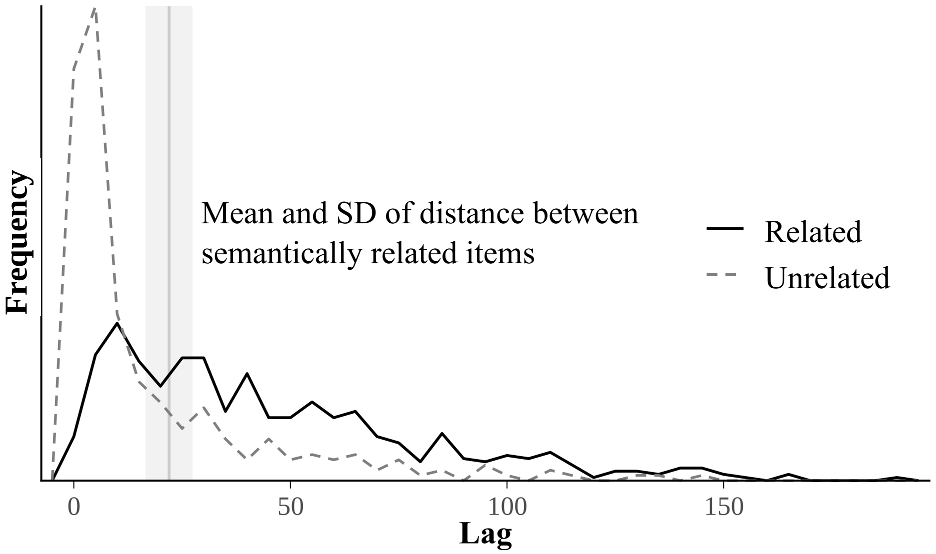

Figure 3. Distributions of semantically related and unrelated perseverations across all participants.

Table 5 summarizes the repeated responses produced by each participant and their properties, and Figure 4 shows the ranges of lags. Rates of repeated responses ranged from 7% to 32% of trials, and median lags ranged from 4 to 40.5, demonstrating considerable variability in both the rate and lag of repeated responses. Note that lags are calculated over all repeated responses, as it is impossible to identify true perseverations at the level of individual trials. As such, the lags shown in Table 5 are only proxies for the lags of true perseverations. They are thus affected by the noise added by the included non-perseverative repeated responses. However, since there are sufficient true perseverations for the statistical tests to produce significant results and the lags for the remaining non-perseverative repetitions should be uniformly distributed, it is safe to assume that the median lags calculated for all repeated responses are in fact reflective of the range over which true perseverations occur. Most of the participants were detected as perseverators by one of the two tests. The only exception was OW, who, despite making many repeated responses (188 in total) did not produce true perseverations according to either the short- or the long-lag tests. We will return to his case in the Discussion.

Table 5. Summary of the repeated responses produced by each participant and their properties. Repetitions are shown as the percentage of all responses. Percent Related is the percentage of repeated responses that are semantically related to their targets. The p-values in the Short-Lag column are from a test for true perseverations with a lag of at most 2 trials. The Long-Lag p-values are from a test for true perseverations with a lag of more than 5 trials (see text for details).

		Percent	**Median**	**P-Value of Test for Perseverations**	
Participant	**Repetitions**	**Related**	**Lag**	*Short-Lag (≤ 2)*	*Long-Lag (> 5)*
QD	7%	95%	40.5	.887	**.010**
GK	13%	91%	39.5	.772	**.024**
FJ	13%	87%	35.0	.992	**.008**
OW	18%	74%	31.5	.605	.099
CE	17%	21%	4.0	< **.001**	.941
EG	32%	24%	9.0	< **.001**	.932

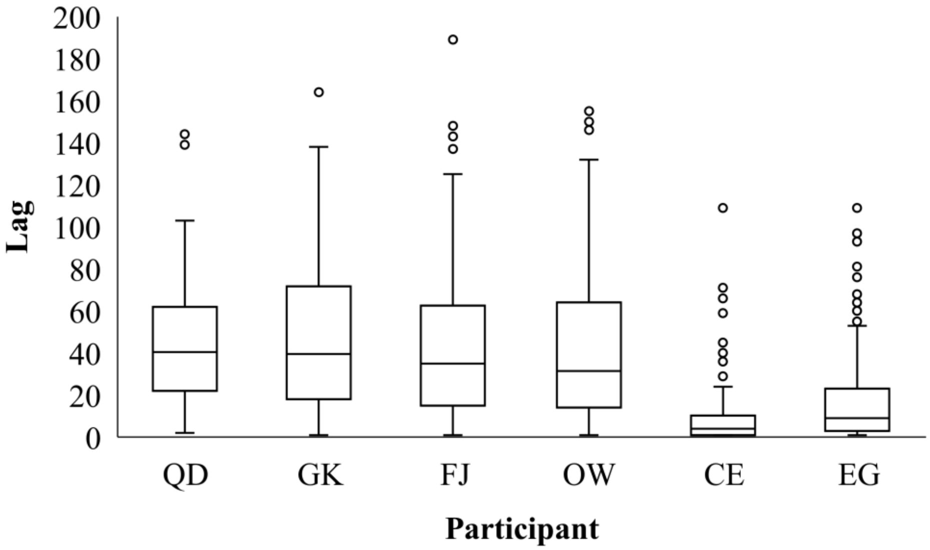

Figure 4. Distribution of lags of the repeated responses produced by each participant.

Three participants (QD, GK, and FJ) tested positive for true long-lag perseverations. They had the longest median lags (40.5, 39.5, and 35 trials, respectively) in the group and produced very few repeated responses (a combined total of just 4) with lags ≤ 2 (out of a total of

353 repeated responses). Moreover, their repeated responses were mostly semantically related perseverations (95%, 91%, and 87% related, respectively).

In contrast, two other participants (CE and EG) made significant numbers of perseverations over short lags (both *p*s < .001 vs. all other *p*s > .5). To verify that this result was not an artifact of the specific window size we had chosen for the test, we repeated the short-lag test with windows of 3, 4, and 5 previous trials. The pattern of results did not change. In all cases, CE and EG were detected as short-lag perseverators (all *p*s < .001) while the others were not (all *p*s > .2). Their repeated responses were mostly semantically unrelated (21% and 24% related, respectively, vs. 74% for the next lowest). Of their semantically unrelated repetitions, most were not phonologically related to the target either (86% and 84%, respectively). Both CE and EG had relatively short median lags (4 and 9 trials, respectively, vs. 31.5 for the next shortest in the group), although these are longer than the upper bound of lag 1-2 typically found for the effects of residual activation (e.g., Joordens & Besner, 1992). As explained earlier, this discrepancy arises from the way in which the median lag is calculated. Since there is no way to determine with certainty whether an individual trial is a true perseveration (see Section 3.1.1), a significant proportion of the repeated responses included in the calculation are likely to be non-perseverative, raising the resulting median.

To summarize, two distinct groups emerged in our analyses: one group produced significant numbers of semantically related perseverations over long lags, and the second produced significant numbers of unrelated perseverations over short lags. While it was theoretically (and statistically) possible to find both types of perseverations within the same participant, none of the participants in our sample showed this pattern. The separation of the two groups is evident in Figure 5, with CE and EG clustering in the lower left corner and the other four participants clustering in the upper right corner. This clustering creates the semblance of a strong linear relationship between participants' median lags and proportions of repeated responses semantically related to the target (Pearson's $r(4) = .996$, $p < .001$). It is, however, noteworthy, that even when the lower cluster is excluded, the relationship between lag and proportion semantically related is still significant in the upper cluster (Pearson's $r(2) = .957$, $p = .043$), suggesting some degree of continuity in the relationship.

To further explore the potential differences between the two groups, we divided commission errors into repetitions (also broken down into short- and long-lag repetitions) and non-repetitions in order to compare the rates of semantically related errors in each category (Table 6). Across the group, there was a strong covariance between the proportions of semantically related errors in repeated and non-repeated errors. For the long-lag perseverators (QD, GK, and FJ) and OW, the proportion of semantically related errors in the two categories was very similar. The short-lag perseverators (CE and EG), on the other hand, produced much larger proportions of semantically unrelated errors in the repetitions than the non-repetitions. We will discuss this finding in more detail in the Discussion.

Table 6 shows that although the task was not designed to elicit them, both semantically related short-lag and semantically unrelated long-lag repetitions were sometimes produced. It is important to explain the origin of these errors: semantically related short-lag repetitions occur if an incorrect response in a preceding trial happened to be semantically related to the current target (e.g., responding “cat” to a picture of a tree, then “cat” to a picture of a dog). This could be caused by either residual activation or incremental learning. As for the origin of unrelated long-lag repetitions, it is important to recall that it is impossible to distinguish between perseverative and non-perseverative repetitions at the level of individual trials. As explained earlier, some of the repeated responses are simply not true perseverations, and thus have nothing to do with a previously produced response, being the result of neither residual activation nor incremental learning mechanisms. Unrelated long-lag repetitions are most likely non-perseverative responses of this kind.

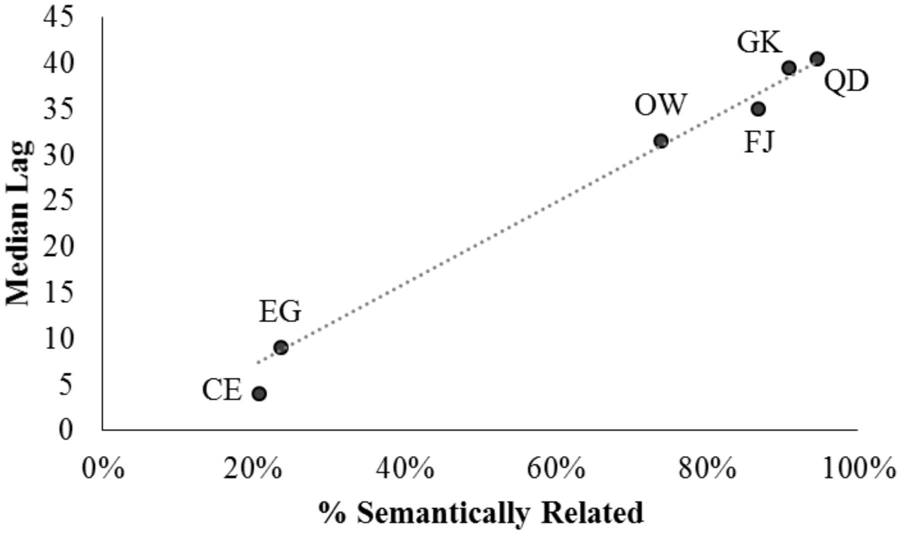

Figure 5. *Relationship between lag and proportion of repeated responses semantically related to the target. The dotted line shows the trend for QD, GK, FJ, and OW.*

Table 6. *Breakdown of commission errors into short- and long-lag repetitions (lags ≤ 2 and lags > 5, respectively) and non-repetitions for each participant, including the rate of semantically related errors in each category.*

	Non-Repetitions		**Short-Lag Repetitions**		**Long-Lag Repetitions**		**All Repetitions**	
Participant	*Total*	*% Related*	*Total*	*% Related*	*Total*	*% Related*	*Total*	*% Related*
QD	186	95%	1	100%	69	97%	74	95%
GK	305	90%	3	67%	136	91%	142	91%
FJ	299	88%	1	0%	124	90%	137	87%
OW	368	83%	8	25%	168	78%	188	74%
CE	113	62%	37	11%	41	37%	106	21%
EG	188	40%	43	12%	117	30%	195	24%

5 Discussion

The aim of this study was to determine which mechanism—residual activation or incremental learning—is responsible for lexical perseverations in aphasia. Previous studies, which have relied on tasks biased towards the production of short-lag, semantically related perseverations that can be explained by either mechanism (e.g., Hsiao et al., 2009). While certain findings (e.g., insensitivity to the time between trials) have been taken as evidence in support of the incremental learning account, other evidence (e.g., the above-chance rate of unrelated perseverations at lags of up to five intervening items) are hard to explain by this account. Importantly, key predictions that differentiate the two accounts have never been tested. These include the existence of long-lag semantically related perseverations, which can only be explained by the incremental learning account, and short-lag, semantically unrelated perseverations, which are only predicted by the residual activation account. Since the two proposed mechanisms are not mutually incompatible, the possibility that both may underlie lexical perseverations in aphasia could not be ruled out. If both mechanisms are in fact at work, it would be necessary to determine whether they coexist in the same system or are tied to specific states.

By administering a picture naming task with no repetitions and long lags between semantically related items and devising a novel method of statistical analysis to detect true perseverations at short and long lags, we were able to answer these questions. First, recall that one individual, OW, did not show evidence of true lexical perseverations at either short or long lags. Although he produced many repeated responses (188 in total), many of them were more consistent with predilections for particular words rather than true perseverations. For example, he incorrectly responded “mosquito” to 11 pictures (mostly insects, but also pliers, a fox, and a plate of spaghetti), “wok” to 11 pictures (mostly kitchen utensils, but also some food items), and “eraser” to 18 pictures (mostly stationery, but also unrelated items like a harmonica and a brush). Importantly, these repeated responses were not all produced within the same sessions. His propensity to produce “mosquito” was evident in three separate sessions, “wok” in two, and “eraser” in four. His profile highlights the importance of carefully determining the baseline probability of non-perseverative repetitions in order to obtain a robust test for true perseverations.

All of the remaining five participants tested positive for true perseverations. Moreover, each participant tested positive for only one type of perseveration, not both. If this reflects a problem in lexical retrieval in all five cases, then this finding answers the first question of the paper unambiguously: both residual activation and incremental learning mechanisms are responsible for lexical perseverations in aphasia, and either mechanism alone is sufficient to cause such errors. The detailed profiles of each participant suggest that the problem is indeed at the level of retrieving lexical items. Serious semantic problems have been ruled out, a majority of the commission errors were lexical, and all participants had better word repetition than naming performance indicative of their stronger phonological encoding compared to lexical retrieval mechanisms. One issue remains to be addressed. Two participants (OW and FJ) had severely impaired phonological working memory. This deficit often results in segmental perseverations (Cohen & Dehaene, 1998; Moses et al., 2004). We thus reasoned that if these participants were to produce short-lag perseverations, it would be necessary to rule out mechanisms that cause perseveration at the level of individual phonemes rather than entire lexical items. Neither participant showed such a pattern. We can thus rule out segment-level mechanisms as the main cause of any participant's perseverations in this study, and conclude that residual activation of lexical representations and incremental learning in mapping semantic features to lexical items constitute two independent mechanisms for generating lexical perseverations in individuals with aphasia.

The second question of the paper was whether the two mechanisms operate in parallel within the same system or whether they are tied to specific states of the system. Our results support the latter view. The next section elaborates on this finding, and its implications for theories of residual activation and incremental learning.

5.1 Implications for the residual activation and incremental learning accounts

Although not to the same extent as individuals with brain damage, neurotypical speakers do sometimes produce perseverative errors resembling those predicted by the residual activation account, especially under demanding situations. For example, tongue-twisters like "lean reed reef leech" elicit errors like "lean ree<u>n</u> reef leech" or "lean reed reef <u>reef</u>" in which segments and/or whole lexical items perseverate soon after a recently produced target (e.g., Nozari & Dell, 2012). A closer look at the literature, however, reveals a more nuanced picture. The ratio of perseverative to anticipatory errors (e.g., "lean reed" → "lean ree<u>n</u>" vs. "lea<u>d</u> reed") decreases as the system's overall performance improves, whether this difference in performance is due to practice (the anticipatory practice effect; Estes, 1972; MacKay, 1987), maturation (children have a higher perseveration to anticipation ratio; Stemberger, 1989), or brain damage (individuals with aphasia make more perseverative errors than neurotypical individuals; Albert & Sandson, 1986; Helmick & Berg, 1976; Shindler et al., 1984). Since a good chunk of these data come from repetitions over very short distances, these findings tie the persistence of residual activation to the strength of the production system by showing that stronger production systems are less likely to repeat the past and more likely to anticipate the future (see Dell, Burger, et al., 1997, for a computational account). There are thus hints in the literature to suggest that residual activation (and its consequences) may be associated with weaker systems. This prediction is borne out by the current data. The two participants whose lexical perseverations were mostly a product of residual activation had the most damaged production systems (both < 35% correct on PNT).

At the same time, neurotypical systems are also subject to incremental learning in various domains, including lexical semantic processing (e.g., Howard et al., 2006; Schnur et al., 2006, 2009), phonological processing (Breining et al., 2019), phonotactic processing (e.g., Warker & Dell, 2006), and higher-level functions like control processes (Freund & Nozari, 2018). As explained in the introduction, one manifestation of incremental learning in language processing is semantic interference. A comparison of the magnitude of semantic interference between individuals with aphasia and neurotypical controls has found larger interference in the former group (Schnur et al., 2006). This finding may suggest that, like residual activation, incremental learning is more pronounced in weaker production systems. However, closer examination of the literature shows that the finding of stronger semantic interference in individuals with brain damage compared to neurotypical participants is not robust (e.g., Piai, Riès, & Swick, 2016; Riès, Karzmark, Navarrete, Knight, & Dronkers, 2015), although the studies that failed to replicate the effect did not have provide enough details in the profiles of individual participants to allow for the development of cognitive hypotheses regarding the reason for the absence of a difference between the two groups.

Our findings can shed light on this issue. First, the hypothesis that incremental learning, like residual activation, is a feature of weaker production systems is not supported by our data. On the contrary, only the three participants with the strongest production systems showed evidence of lexical perseveration due to incremental learning. Second, Table 6 shows a clear dissociation between the rates of semantically related responses in repetitions vs. non-repetitions in the two participants whose perseverations were consistent with residual activation. This finding is important, because it shows that committing a semantic error is not the same thing as repeating a semantically related item. It also addresses another discrepancy in the past literature. On the one hand, N. Martin and Dell (2007) reported similar proportion of semantically related errors in repeated and non-repeated responses (see also N. Martin, Roach, Brecher, & Lowery, 1998). On the other hand, Moses, Sheard, & Nickels (2007) have reported discrepancies in these proportions, with two of their participants producing semantically unrelated errors at a higher rate in repeated responses. Our results indicate that, while there is a close relationship between the dynamics of spreading activation and selection which cause semantically related errors and those which increase the propensity of the system to repeat previous responses in participants with stronger production systems, this relationship breaks down in weaker systems. Participants with weaker production systems are in fact relatively less likely to repeat a semantically related response. Together, these two findings suggest that incremental learning is actually a feature of relatively *stronger* production systems.

Why might this be the case? For incremental learning to have any meaningful effect, semantic-to-lexical mappings must be sufficiently intact to consistently activate lexical items in the correct semantic category. Computationally, this means the connections between semantic features and lexical items must be relatively strong, and there must be relatively low levels of noise in the system. Such a system would be more likely to produce semantically related errors than unrelated ones in non-repeated responses (e.g., Dell, Schwartz, et al., 1997). Compared to a system in which the connections were weaker and noise was higher, incremental learning would be more likely to make meaningful adjustments to connection weights, since adjustments based on noisy patterns of activation would themselves add further noise to the system, rather than systematic biases. This would increase the probability of observing semantically related perseverations arising from these incremental changes to the connections.

To summarize, tying the findings of the current study to the broader literature on language production indicates that the two mechanisms behind lexical perseverations are indeed linked to different states of the production system. Residual activation is most likely to generate perseverations in relatively weak systems, or systems under stress (e.g., demand for processing related strings under time pressure). Incremental learning, on the other hand, requires a certain degree of structural and functional preservation, and is therefore the underlying cause of perseverations only in relatively stronger production systems.

Our results have another critical implication for the theories of incremental learning. One of the key predictions of the incremental learning account, both for perseverations specifically and for semantic interference effects more broadly, is that the cumulative changes to connection weights are long-lasting (Oppenheim et al., 2010). The structural changes to the connections between semantic features and lexical representations induced by a naming attempt should not decay appreciably unless and until semantically related items are named in subsequent trials. In other words, the theory predicts no clear upper bound on the lag at which semantically related errors can be induced by incremental learning mechanisms. However, this prediction has not found clear support in the empirical data. Using a continuous naming task, Schnur (2014) found that semantic interference disappeared after as few as 8 items, unless a “biasing mechanism” that boosted the effects of incremental learning was triggered by a pair of semantically related items appearing close together (separated by only two unrelated items) appearing within the sequence.

Even with such a mechanism, semantic interference was only observed up to 14 items[1]. These findings, while clearly beyond the scope of simple residual activation, suggest that truly long-lasting structural changes in the production system may only occur in certain limited contexts.

The findings of the current study speak directly to the longevity of the structural changes in the language production system induced by incremental learning. Despite the long gaps between the semantically related items in the design (none separated by as few as two unrelated items, which Schnur's findings indicated to be necessary to trigger the "biasing mechanism"), three out of six participants showed true perseverations with median lags ranging from 35 to 40.5 trials and maximum lags. Together with our method for distinguishing true perseverations from non-perseverative repetitions, these data support a fundamental claim of the incremental learning accounts: every act of production changes the production system forever.

5.2 Challenges, solutions, and conclusions

Analyses of perseverative errors have provided a number of important insights into the language production system. For example, in written language, they have been used to determine the components and structures of orthographic representations. Elements that can perseverate independently must be represented independently, so the finding that gemination (i.e., doubling of a letter) can perseverate independently from letter identity (e.g., "tragic" → T<u>RR</u>ACE after EXCE<u>SS</u>) has provided strong evidence in favor of complex, multidimensional orthographic representations (Fischer-Baum & Rapp, 2014). Perseverations have also been used to investigate how position is represented, by exploiting the fact the perseverated letters tend to maintain roughly the same position (Fischer-Baum et al., 2010). In the current study, we used perseverations to test a key prediction of incremental learning models in language production.

---

[1] It must be noted that longer lags were not tested in this study.

Thus, beyond their significance as a symptom of brain damage, perseverative errors are extremely valuable tools for the studying of cognition.

At the same time, the study highlighted some of the common challenges in studying this particular type of error. The finding of a large number of true perseverations at long lags showed that limiting the analysis window to a small number of items is likely to miss a substantial number of perseverative errors. For the reasons discussed earlier, this may be inevitable for segmental perseverations, but it is not so for lexical perseverations. Moreover, half of our participants showed statistically significant evidence of perseverations only when such restrictions were removed. Based on these results, we advise against limiting the study of lexical perseverations to small windows, and instead recommend using the statistical analysis method described here, which can be easily applied to different production tasks with flexibility in the number of administered items. Combining the use of simple tasks (like picture naming) and complex tasks that simultaneously tap multiple cognitive processes (like the Wisconsin Card Sorting task) and using statistical methods that help separate predilections and chance repetitions from true perseverations, provide the best opportunity for learning about the fundamental mechanisms underlying the generation of perseverative errors.

In conclusion, we found clear evidence for both the residual activation and incremental learning accounts, demonstrating that either one is sufficient to cause perseveration in an individual with aphasia. Perseveration due to residual activation was associated with more severely impaired systems than those due to incremental learning, suggesting that a certain degree of structural and functional integrity was necessary for incremental learning to operate. Finally, the results supported a key prediction of the incremental learning account by showing true perseverations over longer lags than have previously been reported.

6 Acknowledgements

We would like to thank QD, GK, FJ, OW, CE, and EG for their participation in many hours of testing, Michael Freund and Alison Trude for collecting much of the data, and Denise McCall at SCALE for her assistance in coordinating the testing sessions. This work was supported in part by the Therapeutic Cognitive Neuroscience Fund at Johns Hopkins University, and a Science of Learning grant from Johns Hopkins University awarded to N.N.

7 Appendix A

Table A1. Additional information on participants' working memory and inhibitory control abilities. The Category and Rhyme Probe tasks (adapted from Freedman & Martin, 2001) assess semantic and phonological working memory, respectively. Scores on these tasks are reported as the length of the longest sequence at which the participant met the criterion for continuing (75% accuracy or higher), followed by a decimal indicating the proportion of the next block answered correctly (e.g., 2.5 means the maximum length at which criterion was met was at length 2, and 50% of responses at length 3 were correct). These scores varied considerably across participants, and, except for QD's Category Probe and GK's Rhyme Probe scores, were lower than control scores. Importantly, the scores showed no relation to the types of perseverations made in picture naming tasks. Scores on the Corsi Block task (Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000), which measures spatial working memory, showed little variation across participants and all fall within the normal range. The Go/NoGo and Simon tasks assess inhibitory control. Participants had varying degrees of difficulty with these tasks, with all but GK scoring outside of normal range, but, as in the working memory tasks, their performance was not predictive of the types of perseverations they made.

	Category Probe	**Rhyme Probe**	**Corsi Blocks**	**Go/NoGo**		**Simon Effect**	
				NoGo Errors	*Go RT*	*Errors*	*RT*
QD	4.60	Unavailable	Unavailable	6	465	5	171
GK	2.67	5.54	17	6	355	2	60
FJ	2.66	0.50	19	10	389	18	47
OW	1.67	0.25	17	10	399	8	175
EG	2.42	2.58	15	12	396	1	91

8 References

Albert, M. L., & Sandson, J. (1986). Perseveration in aphasia. *Cortex*, *22*(1), 103–115. https://doi.org/10.1016/S0010-9452(86)80035-1

Allison, R. S., & Hurwitz, L. J. (1967). On perseveration in aphasics. *Brain*, *90*(2), 429–448. https://doi.org/10.1093/brain/90.2.429

Bayles, K. A., Tomoeda, C. K., Kaszniak, A. W., Stern, L. Z., & Eagans, K. K. (1985). Verbal perseveration of dementia patients. *Brain and Language*, *25*(1), 102–116. https://doi.org/10.1016/0093-934X(85)90123-3

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 Section A*, *58*(4), 667–692. https://doi.org/10.1080/02724980443000142

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *The Quarterly Journal of Experimental Psychology*, *66*(11), 2135–2160. https://doi.org/10.1080/17470218.2013.775318

Boucher, J. (1977). Alternation and sequencing behaviour, and response to novelty in autistic children. *Journal of Child Psychology and Psychiatry*, *18*(1), 67–72. https://doi.org/10.1111/j.1469-7610.1977.tb00417.x

Breining, B. L., Nozari, N., & Rapp, B. (2016). Does segmental overlap help or hurt? Evidence from blocked cyclic naming in spoken and written production. *Psychonomic Bulletin & Review*, *23*(2), 500–506. https://doi.org/10.3758/s13423-015-0900-x

Breining, B. L., Nozari, N., & Rapp, B. (2019). Learning in complex, multi-component cognitive systems: Different learning challenges within the same system. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *45*(6), 1093–1106. https://doi.org/10.1037/xlm0000630

Buckingham, H. W., Avakian-Whitaker, H., & Whitaker, H. A. (1978). Alliteration and assonance in neologistic jargon aphasia. *Cortex*, *14*(3), 365–380. https://doi.org/10.1016/S0010-9452(78)80063-X

Buckingham, H. W., Whitaker, H., & Whitaker, H. A. (1979). On linguistic perseveration. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in Neurolinguistics* (pp. 329–352). https://doi.org/10.1016/B978-0-12-746304-9.50016-1

Cohen, L., & Dehaene, S. (1998). Competition between past and present. Assessment and interpretation of verbal perseverations. *Brain*, *121*(9), 1641–1659. https://doi.org/10.1093/brain/121.9.1641

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

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, *104*(1), 123–147. https://doi.org/10.1037/0033-295X.104.1.123

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: W. H. Winston & Sons.

Fischer-Baum, S., McCloskey, M., & Rapp, B. (2010). Representation of letter position in spelling: Evidence from acquired dysgraphia. *Cognition*, *115*(3), 466–490. https://doi.org/10.1016/j.cognition.2010.03.013

Fischer-Baum, S., Miozzo, M., Laiacona, M., & Capitani, E. (2016). Perseveration during verbal fluency in traumatic brain injury reflects impairments in working memory. *Neuropsychology*, *30*(7), 791–799. https://doi.org/10.1037/neu0000286

Fischer-Baum, S., & Rapp, B. (2012). Underlying cause(s) of letter perseveration errors. *Neuropsychologia*, *50*(2), 305–318. https://doi.org/10.1016/j.neuropsychologia.2011.12.001

Fischer-Baum, S., & Rapp, B. (2014). The analysis of perseverations in acquired dysgraphia reveals the internal structure of orthographic representations. *Cognitive Neuropsychology*, *31*(3), 237–265. https://doi.org/10.1080/02643294.2014.880676

Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology*, *18*(3), 193–226. https://doi.org/10.1080/02643290126002

Freeman, T., & Gathercole, C. E. (1966). Perseveration—The Clinical Symptoms—In Chronic Schizophrenia and Organic Dementia. *The British Journal of Psychiatry*, *112*(482), 27–32. https://doi.org/10.1192/bjp.112.482.27

Freund, M., & Nozari, N. (2018). Is adaptive control in language production mediated by learning? *Cognition*, *176*, 107–130. https://doi.org/10.1016/j.cognition.2018.03.009

Fuld, P. A., Katzman, R., Davies, P., & Terry, R. D. (1982). Intrusions as a sign of Alzheimer dementia chemical and pathological verification. *Annals of Neurology*, *11*(2), 155–159. https://doi.org/10.1002/ana.410110208

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 9, pp. 133–177). https://doi.org/10.1016/S0079-7421(08)60270-4

Gotts, S. J., della Rocchetta, A. I., & Cipolotti, L. (2002). Mechanisms underlying perseveration in aphasia: Evidence from a single case study. *Neuropsychologia*, *40*(12), 1930–1947. https://doi.org/10.1016/S0028-3932(02)00067-2

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404–411. https://doi.org/10.1037/h0059831

Helmick, J. W., & Berg, C. B. (1976). Perseveration in brain-injured adults. *Journal of Communication Disorders*, *9*(2), 143–156. https://doi.org/10.1016/0021-9924(76)90006-X

Hirsh, K. W. (1998). Perseveration and activation in aphasic speech production. *Cognitive Neuropsychology*, *15*(4), 377–388. https://doi.org/10.1080/026432998381140

Hong, Y., & R Core Team. (2020). poibin: The Poisson binomial distribution (Version 1.5) [R]. Retrieved from https://CRAN.R-project.org/package=poibin

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

Hsiao, E. Y., Schwartz, M. F., Schnur, T. T., & Dell, G. S. (2009). Temporal characteristics of semantic perseverations induced by blocked-cyclic picture naming. *Brain and Language*, *108*(3), 133–144. https://doi.org/10.1016/j.bandl.2008.11.003

Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 483–491. https://doi.org/10.1037/0278-7393.18.3.483

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258. https://doi.org/10.1207/S15324826AN0704_8

Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson’s disease. *Brain*, *106*(2), 257–270. https://doi.org/10.1093/brain/106.2.257

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. https://doi.org/10.1007/978-1-4612-4754-8

Martin, N., & Dell, G. S. (2007). Common mechanisms underlying perseverative and non-perseverative sound and word substitutions. *Aphasiology*, *21*(10–11), 1002–1017. https://doi.org/10.1080/02687030701198346

Martin, N., Roach, A., Brecher, A., & Lowery, J. (1998). Lexical retrieval mechanisms underlying whole-word perseveration errors in anomic aphasia. *Aphasiology*, *12*(4–5), 319–333. https://doi.org/10.1080/02687039808249536

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*(1), 83–111. https://doi.org/10.1006/jmla.1994.1005

McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, *23*(2), 278–307. https://doi.org/10.1080/02643290442000518

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, *9*(1), 90–100. https://doi.org/10.1001/archneur.1963.00460070100010

Miozzo, M., Fischer-Baum, S., & Caccappolo-van Vliet, E. (2013). Perseverations in Alzheimer's disease: Memory slips? *Cortex*, *49*(8), 2028–2039. https://doi.org/10.1016/j.cortex.2012.10.016

Moses, M. S., Nickels, L. A., & Sheard, C. (2004). Disentangling the web: Neologistic perseverative errors in jargon aphasia. *Neurocase*, *10*(6), 452–461. https://doi.org/10.1080/13554790490894057

Moses, M. S., Sheard, C., & Nickels, L. A. (2007). Insights into recurrent perseverative errors in aphasia: A case series approach. *Aphasiology*, *21*(10–11), 975–1001. https://doi.org/10.1080/02687030701198312

Nozari, N. (2019). The dual origin of semantic errors in access deficit: Activation vs. inhibition deficit. *Cognitive Neuropsychology*, *36*(1–2), 31–53. https://doi.org/10.1080/02643294.2019.1587397

Nozari, N., & Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1084–1090. https://doi.org/10.1037/a0026933

Nozari, N., & Dell, G. S. (2013). How damaged brains repeat words: A computational approach. *Brain and Language*, *126*(3), 327–337. https://doi.org/10.1016/j.bandl.2013.07.005

Nozari, N., Dell, G. S., Schneck, K., & Gordon, B. (2015). Implementation of selective attention in sequential word production. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual*

Meeting of the Cognitive Science Society (pp. 1745–1750). Retrieved from https://mindmodeling.org/cogsci2015/papers/0303/index.html

Nozari, N., Freund, M., Breining, B. L., Rapp, B., & Gordon, B. (2016). Cognitive control during selection and repair in word production. *Language, Cognition and Neuroscience*, *31*(7), 886–903. https://doi.org/10.1080/23273798.2016.1157194

Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language*, *63*(4), 541–559. https://doi.org/10.1016/j.jml.2010.08.001

Nozari, N., & Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics*, *53*, 100875. https://doi.org/10.1016/j.jneuroling.2019.100875

Oppenheim, G. M. (2018). The paca that roared: Immediate cumulative semantic interference among newly acquired words. *Cognition*, *177*, 21–29. https://doi.org/10.1016/j.cognition.2018.02.014

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*(3), 249–262. https://doi.org/10.1016/0028-3932(82)90100-2

Piai, V., Riès, S. K., & Swick, D. (2016). Lesions to lateral prefrontal cortex impair lexical interference control in word production. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00721

R Core Team. (2020). R: A language and environment for statistical computing (Version 3.6.3). Retrieved from https://www.R-project.org/

Riès, S. K., Karzmark, C. R., Navarrete, E., Knight, R. T., & Dronkers, N. F. (2015). Specifying the role of the left prefrontal cortex in word selection. *Brain and Language*, *149*, 135–147. https://doi.org/10.1016/j.bandl.2015.07.007

Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia naming test: Scoring and rationale. In M. L. Lemme (Ed.), *Clinical aphasiology* (Vol. 24, pp. 121–133). Retrieved from http://aphasiology.pitt.edu/215/

Sandson, J., & Albert, M. L. (1984). Varieties of perseveration. *Neuropsychologia*, *22*(6), 715–732. https://doi.org/10.1016/0028-3932(84)90098-8

Schnur, T. T. (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language*, *75*, 27–44. https://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. https://doi.org/10.1016/j.jml.2005.10.002

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. https://doi.org/10.1073/pnas.0805874106

Shindler, A. G., Caplan, L. R., & Hier, D. B. (1984). Intrusions and perseverations. *Brain and Language*, *23*(1), 148–158. https://doi.org/10.1016/0093-934X(84)90013-0

Stemberger, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language*, *28*(2), 164–188. https://doi.org/10.1016/0749-596X(89)90042-9

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

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

Warker, J. A., & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 387–398. https://doi.org/10.1037/0278-7393.32.2.387

Yamadori, A. (1981). Verbal perseveration in aphasia. *Neuropsychologia*, *19*(4), 591–594. https://doi.org/10.1016/0028-3932(81)90026-9