



Published in final edited form as:

Cogn Neuropsychol. 2008 June ; 25(4): 463–492. doi:10.1080/02643290701674851.

Where is the effect of frequency in word production? Insights from aphasic picture naming errors

Audrey K. Kittredge^a, Gary S. Dell^a, Jay Verkuilen^b, and Myrna F. Schwartz^c

^a Beckman Institute, University of Illinois, 405 N. Matthews Ave., Urbana, IL 61801, USA

^b Department of Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820, USA

^c Moss Rehabilitation Research Institute, 1200 West Tabor Road, Philadelphia, PA 19141, USA

Abstract

Some theories of lexical access in production locate the effect of lexical frequency at the retrieval of a word's phonological characteristics, as opposed to the prior retrieval of a holistic representation of the word from its meaning. Yet there is evidence from both normal and aphasic individuals that frequency may influence both of these retrieval processes. This inconsistency is especially relevant in light of recent attempts to determine the representation of another lexical property, age of acquisition or AoA, whose effect is similar to that of frequency. To further explore the representations of these lexical variables in the word retrieval system, we performed hierarchical, multinomial logistic regression analyses of 50 aphasic patients' picture-naming responses. While both log frequency and AoA had a significant influence on patient accuracy and led to fewer phonologically related errors and omissions, only log frequency had an effect on semantically related errors. These results provide evidence for a lexical access process sensitive to frequency at all stages, but with AoA having a more limited effect.

Introduction

A speaker's success in retrieving words when producing sentences depends on a daunting variety of syntactic, semantic and pragmatic factors. Those who study lexical access often use single-word production tasks and narrow the scope of investigation to word-level properties, such as the frequency with which a word occurs. Beginning with Oldfield and Wingfield's picture-naming study in 1965, the facilitative effect of high-frequency words like "ball" has been documented in a multitude of production tasks (and comprehension tasks, as well; see e.g. Dahan, Magnuson, & Tanenhaus, 2001). The robustness of this effect, combined with lexical frequency's psychological importance as a measure of linguistic experience, has given it a prominent role in many theories of lexical access (e.g. Levelt, Roelofs, & Meyer, 1999; Jescheniak & Levelt, 1994).

However, there is still disagreement about the representation of frequency within the word production system. This debate has also become more important in the face of the rising theoretical importance of another experience-based variable, the age at which a word is acquired (age-of-acquisition, or AoA). "Ball" is also easier to produce because it was acquired early in life. Which variable matters more, and at which point during the retrieval of a word? Answering these questions is important for a better understanding of the language production system as a whole. To this end we investigate the impact of frequency and AoA on the picture-naming errors of patients with aphasia.

A large body of research indicates that lexical access can be separated into two distinct steps (Harley & Bown, 1998). Assuming that a lexical concept has already been identified, the first step is the retrieval of a holistic lexical representation corresponding to that concept. This has alternately been called “lemma access” (Levelt et al., 1999), access of “word nodes” during “step one” (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), access of “phonological and orthographic lexemes” (Caramazza, 1997), and “L-retrieval” (Rapp & Goldrick, 2000); we will use this last, more neutral label for the remainder of the paper. The second step is the retrieval of the word’s phonological characteristics, or *phonological retrieval*. In all of these theories, the two steps are, to a greater or lesser extent, associated with different kinds of speech errors. Importantly, the first step, the access of a holistic lexical unit from meaning, is a major locus for semantic substitutions, such as saying “knee” for “elbow” or “drive” for “fly.” In contrast, all theories assume that the phonological retrieval step is the locus of most errors that are phonologically related to the target, such as when a target is mispronounced as a nonword (e.g. “heminopter” for “helicopter”). We will be particularly concerned with the effects of frequency (and AoA) on semantic and phonological errors.

A prominent two-step theory of lexical access maintains that frequency affects phonological retrieval, but not L-retrieval (Levelt et al., 1999). Its main impetus is a study by Jescheniak and Levelt (1994), who showed that low- and high-frequency homophones, which are thought to share phonological but not L-level representations, were translated from English to Dutch with equal speed. High-frequency control words, though, were translated faster than low-frequency controls. In contrast, there was no effect of frequency on an object recognition task (using pictures that had elicited a frequency effect in a naming task), or on delayed word naming of the picture names, which suggested that the frequency effect found in the translation task did not originate at the pre-lexical (conceptual) or post-lexical (articulatory) levels of processing. Jescheniak and Levelt (1994) also showed that when subjects made a decision about the grammatical gender of the same picture names (a task that is thought to reflect L-retrieval), there was no robust effect of frequency. Taken together, these results suggested that frequency operates exclusively at the level of phonological retrieval.

A phonological-level locus of the lexical frequency effect is also supported by the speech error literature: High- and low-frequency homophones are equally prone to experimentally elicited phonological errors (Dell, 1990), and low-frequency words are more likely than high-frequency words to elicit errors that are phonologically related to the target, in both normal participants (Dell, 1990; Laubstein, 1999; Vitevitch, 1997) and aphasic patients (Gagnon, Schwartz, Martin, Dell, & Saffran, 1997; Gordon, 2002; Schwartz, Wilshire, Gagnon, & Polansky, 2004).

The hypothesis that frequency only affects phonological retrieval has weighed heavily on the interpretation of experimental findings (Garrett, 2001; Harley & MacAndrew, 2001), at times to the point that frequency is used as a “litmus test” of phonological retrieval (e.g. Graves, Grabowski, Mehta, & Gordon, 2007; Griffin & Bock, 1998; Ferreira & Pashler, 2002). Yet there is evidence suggesting that frequency affects L-retrieval as well as phonological retrieval, implying that frequency is represented at multiple stages. A number of studies on homophones have challenged the findings of Jescheniak and Levelt (1994): Caramazza, Costa, Miozzo, and Bi (2001) failed to find any evidence that pictures of low-frequency homophones were named faster than other low-frequency words. Gahl’s (2006) speech corpus analysis exploited the known tendency of speakers to produce high-frequency words with shorter acoustic durations to investigate the production of homophones, and found that those high in frequency had reliably shorter durations than their low-frequency counterparts. Finally, Jescheniak, Meyer, and Levelt (2003) found that low-frequency homophones were translated from English to German faster than low-frequency controls, but not as fast as high-frequency homophones (see also, Caramazza, Bi, Costa & Miozzo, 2004). Aside from the homophone evidence, recent

studies have found robust effects of frequency on noun phrase production and gender decisions to pictures, which were assumed to reflect L-retrieval because grammatical information was relevant (Alario, Costa, & Caramazza, 2002; Navarette, Basagni, Alario, & Costa, 2006).

The normal speech error literature is likewise suggestive of frequency-sensitivity at L-retrieval, although it may not be as robust as the effect on phonological retrieval. Harley and MacAndrew (2001) found that low-frequency words were more likely to elicit errors semantically related to the target, and Vitkovtich and Humphreys (1991) found a greater incidence of (mostly semantic) errors to speeded naming of pictures with low-frequency names. Gollan and Brown (2006) found that word “difficulty” (correlated with frequency) influenced the probability of retrieving words from pictures in normal individuals, both when the nature of the error reflected problems with word-form access (a tip-of-the-tongue, or TOT state) and word-meaning access (retrieving an incorrect name or failure to retrieve any name). Effects of frequency on semantic errors have also been found in individuals with aphasia. Feyereisen, Van der Borgh and Seron (1988) found significant effects of frequency on “semantico/visual”, phonological, and no response errors made by 18 aphasics in a picture-naming study that orthogonally manipulated frequency and operativity (a variable that reflects the extent to which the named object can be “handled or used in daily life situations”, p. 401). Studies in which more psycholinguistic variables were included have found isolated effects for individual aphasics: Nickels and Howard (1994) note that two of their 15 patients made more semantic errors to low-frequency, low-imageability targets, while one patient in Cuetos, Aguado, Izura and Ellis (2002), whose errors were mostly semantic, showed a significant effect of frequency on accuracy. Our analysis will allow us to see whether such effects might emerge in a larger sample of aphasics.

In addition to clarifying the representation of frequency, our study has the potential to answer questions about AoA. Early AoA (measured by subjective adult ratings or child performance data) tends to imply high frequency, and the two variables have a facilitative effect similar in pattern and size for most production tasks (although the size of the AoA effect is much larger in some, such as picture naming; Brysbaert & Ghyselinck, 2006; Morrison, Chappell, & Ellis, 1997). Several earlier studies found no effect of frequency when AoA was controlled, leading to speculations that only the latter has a real effect on processing, but subsequent experiments have shown this to be unlikely (Barry, Hirsh, Johnston, & Williams, 2001; Johnston & Barry, 2006). One goal of our study is thus to test for distinct effects of AoA and frequency, which is especially crucial if we wish to draw conclusions about frequency’s representation.

As is the case with frequency, claims have been made about the locus of effects of AoA on word retrieval. Some theories propose that AoA affects phonological retrieval, with early-acquired word forms being easier to access (Barry et al., 2001; Brown & Watson, 1987). Support for these theories comes from the effect of AoA on word naming (see Johnston & Barry, 2006 for a review of the literature). Others propose that while both AoA and frequency may affect phonological retrieval, only AoA affects tasks sensitive to L-retrieval, such as picture naming (Brysbaert & Ghyselinck, 2006). The effect of AoA on picture naming, combined with the influence of AoA on word-associate generation and word categorization, have also generated hypotheses of a pure conceptual locus (Ghyselinck, Lewis, & Brysbaert, 2004; Steyvers & Tenenbaum, 2005). Steyvers and Tenenbaum (2005) offer a mechanism whereby this might arise: In a localist semantic network, new conceptual nodes are more likely to attach to nodes with many other connections, resulting in clustering of nodes around early-acquired concepts and allowing them to be accessed more easily.

Ellis and Lambon Ralph (2000) present a similar hypothesis in a distributed connectionist network, but propose that AoA (and frequency) is represented at all levels of the network. Early-acquired words have a greater role in shaping the network’s weights, so they will be accessed more easily. The effects of any lexical variable such as AoA, however, will be most

pronounced in a distributed network when the mapping between the network's input and output is arbitrary, as is the case for the mapping from semantics to phonology in lexical access. This is because a nonarbitrary, or consistent, mapping (such as accessing phonemes from graphemes in a regularly-spelled word) is, by definition, one that draws on the network's existing knowledge. Thus, the properties of that specific mapping (such as its frequency or AoA) will not matter as much because knowledge from previous mappings can be generalized (see Lambon Ralph & Ehsan, 2006 for a more detailed discussion). Ellis and Lambon Ralph's (2000) hypothesis is supported by the presence of AoA effects both in tasks that involve semantic and phonological processing, as well as the larger effect of AoA in picture naming (Johnston & Barry, 2006). As can be seen from this brief review, there is even greater disagreement about the representation of AoA than that of frequency.

How are we to tease apart these highly correlated variables, frequency and AoA, and test the hypotheses concerning their representation? We addressed this question by performing multiple regression analyses of aphasic picture naming errors, to determine which variables are most influential for predicting patient performance. AoA effects on picture naming speed are typically as large or larger than frequency effects, suggesting that picture naming is sensitive to both variables. We used data from aphasics because they produce a much greater number and variety of errors in picture naming than normal participants. As it has been hypothesized that both normal and aphasic speech errors can be accommodated in a unified model of lexical access (e.g. Dell et al., 1997; Rapp & Goldrick, 2000), we use our results to make inferences about normal lexical access (see Rapp, 2001). Finally, we chose to analyze errors because, as we mentioned earlier, certain error types can be associated with distinct stages of lexical access. We performed two sets of statistical analyses: The first to determine which lexical properties of the target make it more or less susceptible to errors (analysis I), and the second to explore the relationship between the target's lexical properties and those of the error it elicits (analysis II).

Analysis I – Target Lexical Properties

Before proceeding to the analysis, we must clarify and formalize our assumptions regarding the association between error types and stages. Semantic errors, if not originating from central conceptual damage, are thought to arise during the first step of lexical access. As discussed previously, current theories of lexical access instantiate this as a mapping from conceptual representation(s) to an intermediate, whole-word unit (although the particular nature of this word unit is debated). However, there is less agreement regarding the number and type of representations accessed from the whole-word unit during the second step of lexical access, at which phonological errors arise: Are they whole lexemes (Levelt, 1989), morphemes (Levelt et al., 1999), or individual phonemes (Caramazza, 1997; Foygel & Dell, 2000; Rapp & Goldrick, 2000)? Here, we use the model of Foygel and Dell (2000) in several computational simulations that allow us to formalize what we mean by the two 'steps' of access, and to identify predicted effects of lexical variables on error production.

Foygel and Dell's model, like those of Rapp and Goldrick (2000) and Caramazza (1997), assumes the existence of semantic features, lexical units (i.e. words), and phonological units (i.e. phonemes) (Figure 1). L-level selection entails the selection of a lexical unit, given semantic input, and phonological retrieval consists in the subsequent selection of phonological units. In this manner, the three layers of this model map onto the two lexical selection steps. Although it is difficult to fit the more complex model of Levelt et al. (1999) into Foygel and Dell's network, it can be done provided that we recognize that Levelt et al.'s lemma-access stage is the stage at which conceptually-driven lexical access occurs. This is where semantic interference can slow naming times (e.g. Schriefers, Meyer, & Levelt, 1990) and cause semantic substitution errors. It is clearly analogous to the L-level selection in Foygel and Dell's

architecture. Subsequent processes in Levelt et al.'s model are responsible for retrieving a word's morphological and phonological units, adjusting those units to fit the context, and retrieving articulatory/motor units. Leaving aside the articulatory/motor aspect of the model, we can loosely identify these post-lemma-access components of Levelt et al.'s model with the phonological retrieval step of Foygel and Dell's model. The key is that, for both models, they occur *after* semantically driven access and *after* the step responsible for semantic errors. Thus, Levelt et al.'s hypothesis that frequency affects "word form" or "lexeme" retrieval corresponds to a locus for frequency in the phonological retrieval step in Foygel and Dell's model (e.g. the homophone inheritance effect, if true, would then presumably be due to influences arising from phonological representations that are important for this step).

In Foygel and Dell's (2000) model, retrieval is effected by spreading activation along bidirectional connections linking adjacent network layers. The semantic-word connections have a strength, or weight, of s ("lexical-semantic weight"), and the word-phoneme connections have a weight of p ("lexical-phonological weight"). As instantiated in Foygel and Dell (2000), the model does not represent variations in lexical properties (e.g. frequency, AoA); all connections have the same weight and all nodes the same resting level.

During the first step (L-retrieval), a jolt of activation is sent to the semantic features of the target word (e.g. "cat"), and this activation spreads to the target word as well as to other words sharing the target's semantic features. Because the model is interactive, the activation at the word level spreads back up to the semantic feature level, as well as downward to phonemes. L-retrieval ends with the selection of the most activated word. Semantic errors (e.g. "dog" for "cat") can occur at this step because semantically-related words receive some activation from semantic features shared with the target. Random noise (i.e. values selected at random from a normal distribution with a mean of 0) is continually being added to the activation of all nodes in the model, and thus a related word may end up being more activated than the target.

The second step, phonological retrieval, begins with a jolt of activation to the word selected at L-retrieval. This activation spreads to the word's phonemes, which send activation back to the word level (both to the target word and to other words in which they appear, e.g. "cat", "mat", "rat"). Phonological retrieval ends with the selection of the most active set of phonemes. Errors at this step are form-related or "phonological" errors. Errors may be "nonwords" (e.g. "gat" for "cat") or real word "formal" errors such as "mat" for "cat" (although it is hypothesized that formal errors can also occur during L-retrieval, due to interactivity; see Schwartz, Dell, Martin, Gahl, & Sobel, 2006).

The model can be lesioned to simulate aphasic performance by reducing the value of s or p (Foygel & Dell, 2000; Schwartz et al., 2006). The reduced weights lead to more semantic and phonological errors because the random noise then has more of an impact. Patients also sometimes fail to respond or make "omissions", but it is unclear exactly what causes these errors. The existing models of omissions have represented them as the failure of a word node to reach an activation threshold or the failure of phonological retrieval, tentatively associating them with L-retrieval (Dell, Lawler, Harris, & Gordon, 2004; Laine, Tikkala, & Juhola, 1998) or phonological retrieval (Harley & Bown, 1998), respectively.

As stated above, the Foygel and Dell model associates semantic errors with L-retrieval and phonological errors, particularly nonwords, with phonological retrieval. For example, if a lexical variable such as frequency affects phonological retrieval, its influence should be seen largely on the rate of nonword errors. We confirmed this prediction in a simulation described in Appendix A. In the simulation, we varied the "frequency" (or "AoA") of the target word by changing the weights of either its lexical-semantic connections or its lexical-phonological connections, while leaving the connections of all other words at their baseline levels. Like in

any connectionist model, the weights are the natural repository for the effects of learning or experience. The weight manipulations allowed us to determine that as expected, putting e.g. frequency in the processes responsible for L-selection (i.e. in the lexical-semantic weights) affects semantic errors but not phonological errors, while associating frequency with phonological retrieval (i.e. in the lexical-phonological weights) affects phonological rather than semantic errors.

This simulation is particularly important because of the model's interactive properties; activation flows upward as well as downward. Intuitively, one might have expected manipulations of lexical-phonological weights to affect semantic errors, and lexical-semantic manipulations to affect phonological errors. This is not the case, however; the influences are quite selective (see Appendix A). To use the terms of Rapp and Goldrick (2000), the model is not "very highly interactive"; instead it is "globally modular" (Dell & O'Seaghdha, 1992) because the jolt of activation to the selected word imparts a stage-like character to the model's processes. The results of our simulation demonstrate that, even in this interactive model, a variable must directly influence a particular lexical access step to have any substantial effect on errors arising there. This allows us to make clear predictions for the regression analyses of patients' picture naming responses. If lexical frequency only affects phonological retrieval, as suggested by some of the literature, it should predict only the incidence of phonological errors. Alternatively, if frequency also affects L-retrieval, it should additionally predict the incidence of semantic errors. Similarly, if AoA primarily affects phonological or L-retrieval, then it should predict phonological or semantic errors, respectively. If it affects both steps of lexical access, it should predict both error types. We refrain from making firm predictions about omissions, given the large uncertainty surrounding this error type (see Dell et al., 2004).

Previous multiple regression analyses of semantic, phonological and omission picture naming errors, for which both AoA and frequency were entered as predictors, have either found effects of AoA (but no effect of frequency) on semantic errors (15 aphasics in Nickels & Howard, 1995), semantic and phonological errors (16 aphasics in Cuetos et al., 2002), or an influence of frequency and AoA on omissions but not on semantic or phonological errors (11 aphasics in Kremin et al., 2003). Crucially, our study, compared to the previous ones, incorporates an increase in the number and variety of patients and will be more likely to reveal small effects and less likely to produce spurious ones. Moreover, because the model's predictions for frequency (or AoA) effects on particular error types are largely independent of specific deficits (see Appendix A), our analysis seeks patterns that apply generally to this large, diverse sample. Nevertheless, the emergence of effects at the group level does depend on the specific deficits of each patient, insofar as patients with damage to a particular level of lexical access produce more picture naming errors originating at that level. Thus, the distribution of deficits is worth noting: Eighty percent of the patients in the sample have below-normal integrity of both L-retrieval and phonological retrieval, as assessed by the values of the *s* and *p* weights obtained from fitting the Foygel and Dell (2000) model to their data (see Schwartz et al., 2006). Based on this assessment, we may expect errors in both L-retrieval and phonological retrieval from most patients, hence providing a good opportunity to discover any general effects of frequency and AoA on these error types.

We also adopt a more appropriate statistical analysis than in previous studies. Instead of using arc-sine transformed proportions of error types amalgamated across patients as the dependent variable, we used multinomial logistic regression to accommodate the categorical nature of the data and the fact that there are multiple categories arising from a single multinomial process (e.g. 1 = Semantic error, 2 = Phonological error, etc.). Importantly, we also accommodate the nesting of responses within patients by using hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002). This technique allows us to predict trends across patients while taking into account patient-specific variability (instead of grouping all patients together or performing a

separate regression analysis for each patient). This type of analysis has been used to study vowel quality in aphasia (Haley, Ohde, & Wertz, 2001), but not picture naming to our knowledge.

Methods

Participants—The 50 aphasic participants were an unsystematic subset of those studied in Schwartz et al. (2006). The inclusion criteria were set broadly, with the aim of maximizing the diversity of the sample: Individuals were sought who had clinically significant chronic aphasia as a result of a left hemisphere cerebrovascular accident (CVA). Patients whose clinical or CT/MRI records revealed evidence of prior left hemisphere stroke(s) were included unless there was multi-focal damage throughout the left hemisphere (see Schwartz et al., 2006 for a full list of exclusion criteria).

Background information for the 50 patients is presented in Table 1 (codes correspond to those used in Schwartz et al., 2006). The mean (and range) for age in years at first testing is 61.6 (22–86), for known years of education, 14 (8–20), for months post-onset at first testing, 33.3 (1–155), and for accuracy on the Philadelphia Naming Test (PNT), 59% (3% – 98%). As a group, the patients are clearly less accurate than a group of 30 age-matched, normal control participants tested on the PNT (average accuracy 97%, range 82%–100%). Forty percent of the aphasic participants were female, and 34% were African American. According to the WAB classification scheme (Western Aphasia Battery; Kertesz, 1982), 34% had Broca’s aphasia, 32% had anomic aphasia, 16% had conduction aphasia, 16% had Wernicke’s aphasia, and 2% had transcortical sensory aphasia. All patients were right-handed with the exception of one ambidexterous patient (patient EAC).

Using such a heterogeneous group of patients (including ones who made very many or very few errors) is consistent with our theoretical perspective and our research goals. As stated previously, we ascribe to a framework in which both normal and aphasic speech errors originate from similar mechanisms and, furthermore, that a model of these mechanisms predicts associations between error types and lexical variables that are largely independent of degree or nature of deficit. We are using aphasic data not to investigate aphasia per se, but rather the nature of the lexical access system. Moreover, any effects that are significant when the heterogeneous patients are treated as random effects must necessarily be quite robust in the face of between-patient noise. However, semantic errors in an unselected patient group could arise either from impaired conceptual processing or from impaired L-retrieval. We thus identified a “Non-semantic” subset of our patients whose conceptual processing was relatively intact and, hence, whose semantic errors were more likely to have occurred during L-retrieval (see Caramazza & Hillis, 1990; Rapp & Goldrick, 2000). This patient subset will be particularly important for our analysis of semantic errors, whereas the full set will be the focus for phonological errors (although all error categories will be analyzed for both the large group and the Non-semantic group). Following Schwartz et al. (2006), exclusion from the Non-semantic group resulted from two or more negative z-scores on three semantic tests (procedures described in Martin, Schwartz, & Kohen, 2005): The Philadelphia Naming Verification Test, or PNV (picture-name verification with semantic foils), Pyramids and Palm Trees, or PPT (picture match-to-sample test based on categorical or associative relationship), and Synonymy Judgments with Nouns and Verbs (closest meaning match on written word triads). This excluded 14 patients. To be conservative regarding whom we placed in the Non-semantic group, we also excluded 11 patients with two or more unavailable z-scores, leaving a total of 25 Non-semantic patients (indicated in Table 1).

Stimuli—The stimuli were the 175 single-word pictures of the PNT, a set of pictures with very high name agreement, but whose names had variable lexical properties (Dell et al.,

1997; Roach, Schwartz, Martin, Grewal, & Brecher, 1996). In addition to obtaining the pictures' frequency and AoA values, we also sought phonological density, length, imageability, and name agreement measures for our regression analyses. These variables are correlated with frequency and AoA, and have been shown to affect aphasic picture naming (Gordon, 2002; Kremin et al., 2003; Nickels & Howard, 1994). Each word's base 10 log frequency per million was obtained from the lemma corpus of the on-line and published CELEX databases (see Baayen, Piepenbrock, & van Rijn, 1993). Raw density values for each word (number of phonologically similar words or "neighbors") were obtained from the "Hoosier Mental Lexicon", an online lexicon of 20,000 words (see Luce & Pisoni, 1998) and converted to base 10 log density values. Log-transformed frequency and density values were used as in previous studies, because log frequency is better correlated with many measures of performance (Vitevitch, 1997) and because the distributions of item frequencies and densities tend to be positively skewed (see Gordon, 2002). AoA values based on the published norms for the American version of the MacArthur Communicative Development Inventories, or CDI (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994) were obtained from the on-line International Picture-Naming Project database (Szekely et al., 2004). Words were divided into three categories: Acquired on average between 8 and 16 months (1), 17 and 30 months (2), and above 30 months (3). Imageability ratings (on a 100–700 scale) were obtained from an online version of the MRC Psycholinguistic Database (Wilson, 1988). Length values represent the number of phonemes in a word. Name agreement values represent the percentage of 60 age-matched control participants who named the item correctly.

Log Frequency, AoA, Log Density, Length, Imageability, and Name Agreement values were available for 124 (71%) of the targets; Tables 2 and 3 display the descriptive statistics and correlation matrix for these items. Log Frequency is mildly associated with Log Density, AoA, Length, and Imageability in the expected directions, and Length and Log Density show a strong negative correlation¹, confirming the need for a regression analysis to disentangle the effects of these variables.

Procedure—The PNT stimuli were presented one at time on a computer, and each trial was ended after 30 seconds if the participant had not responded. At the end of each trial the experimenter said the target name. Sessions were scored online by an experienced speech-language pathologist and were also audiotaped. Only the first complete response produced on each trial was scored, and only exact matches to the target were counted as Correct except when patients had clinically obvious articulatory-motor impairments (for further details see Schwartz et al., 2006). The scoring procedure yielded the following error categories: Semantic (synonym of the target or coordinate, superordinate or subordinate member of its category, e.g. "apricot" for the target "pineapple"), Formal (any word response that meets the PNT's phonological similarity criterion², e.g. "pillow" for "pineapple"), Mixed (response that meets both semantic and phonological similarity criteria, e.g. "banana" for "pineapple"), Unrelated (response that meets neither semantic nor phonological similarity criteria and is not visually related to the target, e.g. "gun" for "pineapple"), and Nonword (neologism that is not also a blend, which either did or did not meet the PNT's phonological similarity criterion, e.g.

¹Even just two highly correlated variables can lead to multicollinearity in a dataset, which can drastically affect the estimation of model parameters. To check for multicollinearity we calculated Variance Inflation Factors (VIFs) for all six independent variables, none of which were over 10 (a standard threshold for multicollinearity; Kutner, Nachtsheim, Neter, & Li, 2004). We also fit the data to a multinomial, logistic regression model (with non-parametric clustering by subject) using Stata 8.2 (Stata Corporation, 2003) with five independent variables, combining the highly correlated variables of Log Density and Length, and obtained the same pattern of results as found in the hierarchical multinomial logistic regression. Together, these diagnostic measures suggest that multicollinearity is not a problem in our dataset.

²To meet the PNT's phonological similarity criterion, the error must bear one of the following similarities to the target: (1) Start or end with the same phoneme, (2) share a common phoneme at another corresponding syllable or word position (aligning words left to right), or (3) share more than one phoneme in any position (excluding unstressed vowels).

“pineme” or “fepe” for “pineapple”). The rest of the errors were classified with one of three additional codes: “description/circumlocution”, henceforth “description” omission (description of the target, often with semantic or phonologically relevant information, e.g. “you eat it” for “pineapple”), “no response” omission (null or semantically empty response e.g. “I don’t know”) and “miscellaneous error” (including visual errors).

Results – Main analysis

In our analyses we focus on three main error types: Semantic, Phonological (collapsing across Formal and Nonword responses), and Omission (collapsing across “no response” and “description” omissions). This is because Semantic and Phonological errors have been linked to distinct lexical access stages (e.g. Dell et al., 1997). Also, these errors and Omissions are the most common aphasic picture naming errors (see Cuetos et al., 2002). The percentage of each of these response types (relative to the total number of responses analyzed, which was constant across patients) is presented for all 50 patients and for Non-semantic patients in Table 4, and for individual patients in Table 5. The remaining response types (Mixed, Unrelated, and miscellaneous errors) were collapsed to form an “Other” category.

First, the patients’ responses to the 124 lexical property-bearing PNT targets were coded as either Correct or Incorrect (collapsing across all the error categories), and submitted to hierarchical, binomial, multiple logistic regression analysis. Hierarchical modeling can be applied to both continuous and categorical data, and is used when observations are not independent. In this case, each patient provides more than one response, which induces correlations among the individual observations. In this situation it is necessary not only to model the effect of lexical variables at the individual level, but also the contribution of patient-specific variability. We used the HLM 6 software (SSI, Inc., 2004) to compute parameter estimates by the penalized quasi-likelihood method. HLM models the log odds that the dependent variable takes on a value of 1 (i.e. a Correct response is made) as a linear function of the six lexical variables (Log Frequency, AoA, Log Density, Length, Imageability, and Name Agreement, centered around their group means) plus a subject-specific random intercept to account for heterogeneity between subjects, who may have different propensities to be accurate not accounted for by the six lexical variables. Thus, the coefficient estimates from this analysis represent the change in the log odds of making a Correct response, *across patients*, due to a one-unit increase in the independent variable of interest. The model is specified as:

$$\log(P_c/P_1) = \beta_0 + \mathbf{x}'_{ijc} \boldsymbol{\beta}_c + u_{ic},$$

where $i = 1, \dots, 50$ indexes subjects, $j = 1, \dots, 124$ indexes trials within subject, and $c = 1, 2$ indexes choices ($c = 1$ being Incorrect, 2 being Correct). β_0 is the intercept, \mathbf{x}'_{ijc} is the vector of lexical variables for a given choice, including a category-specific intercept which ensures that the margins are fit properly, while $\boldsymbol{\beta}_c$ gives the category-specific regression coefficients for the log-odds. The u_{ic} terms are assumed to be normally distributed from a multivariate normal with mean 0 and an unstructured covariance matrix. In addition to fitting this model to the data of the entire group of 50 patients, we also did so for the Non-semantic group separately.

We then evaluated the effect of the independent variables on different types of errors by submitting the 50 patients’ response data to hierarchical, *multinomial*, multiple logistic regression. This model is an extension of the binomial model described above, for data that fall into more than 2 choice categories. As there is dependency between these choices (i.e. if a patient makes a Semantic error, he or she is necessarily NOT making a Correct response, a

Phonological error, etc.), it is more appropriate to simultaneously estimate the effects of lexical variables for each choice, as opposed to conducting separate binomial regressions for each.

In this model, the log-odds of each choice category (Semantic, Phonological, Omission, and Other) relative to the baseline category (Correct) is modeled as a linear function of the six lexical variables and a subject- and category-specific random intercept. The coefficient estimates for this model represent the change in the *relative* log odds (across patients) of making a particular type of error, due to a one-unit increase in a particular lexical variable. For example, a positive AoA coefficient significantly different from zero for Phonological errors would imply the following: A one-unit increase in AoA is associated with a significant increase in the log odds of making a Phonological error, relative to making a Correct response. This model is specified exactly as in the binomial case above, except that the index c represents the multiple choices ($c = 1$ being Correct, 2 being Semantic, 3 being Phonological, 4 being Omission, and 5 being Other). This analysis was also performed for the Non-semantic patient group.

We present the results of significance tests with robust standard errors³. Robust standard errors, also known as heteroscedasticity-consistent standard errors, are less sensitive to violation of model assumptions than are the usual model-based standard errors generated by maximum likelihood estimation. So long as the fixed effects (e.g. coefficient estimates) of the model are correctly specified, robust standard errors are statistically consistent. They are typically somewhat wider than ordinary model-based standard errors when heteroscedasticity is present. For instance, if there is a small amount of dependence among the observations left after including the regressors and random effects that would induce heteroscedasticity, robust standard errors will be adjusted appropriately. However, the regular and robust standard errors estimated by our models are similar, which suggests that there is no notable heteroscedasticity (see Long & Ervin, 2000 for a short introduction and e.g. Cameron & Trivedi, 2005 for a full discussion of the theory).

Furthermore, we focus only on significant effects in the expected directions. A substantial body of literature has shown that words high in frequency, acquired early in life, short, highly imageable, and with high name agreement are more easily produced. The direction of the effect of density, however, depends on the situation (see simulations performed by Dell & Gordon, 2003). Thus for Log Frequency, AoA, Length, Imageability, and Name Agreement, we consider only results in the expected direction to be significant. We also do not present the results for Other errors from the multinomial analyses, as effects for this ad-hoc category are not relevant to the goals of this investigation.

Overall analysis of 50 patients—The effects of Log Frequency and AoA are discussed separately below. A summary of the direction of these effects can be found in Table 6.

Log Frequency—The binomial logistic regression revealed a significant increase in the log odds of making a Correct response per unit increase in Log Frequency (regression coefficient = .414 (standard error = .078); $t = 5.290$, $p < .001$, i.e. the p-value was less than the smallest non-zero p-value reported by the software). This effect of frequency on accuracy has been seen in previous aphasic picture naming studies (Cuetos et al., 2002; Gordon, 2002; Santo Pietro & Rigrodsky, 1982), as well as in speech error data from normal participants (e.g. Dell, 1990). In the multinomial logistic regression analysis, high Log Frequency significantly decreased the log odds of making a Phonological error relative to making a Correct response⁴ (–.588 (.

³Robust standard errors were not used for the multinomial regression with Non-semantic patients, as HLM was only able to estimate regular standard errors due to the smaller number of patients. In general, robust standard errors are not advisable in small samples.

⁴A separate multinomial, hierarchical logistic regression analysis showed that Log Frequency had a significant effect on both phonological word (–.704 (.122); $p < .001$) and nonword (–.484 (.083); $p < .001$) errors.

072); $t = -8.219$, $p < .001$). This is consistent with the effects of frequency found on form-related errors in normal individuals (Dell, 1990; Harley & MacAndrew, 2001) and form-related picture naming errors of aphasics (Gordon, 2002). The effects of frequency on the probability of making Correct and Phonological errors are thus entirely expected and uncontroversial.

High Log Frequency also significantly decreased the log odds of making a Semantic error relative to making a Correct response ($-.377$ (.095); $t = -3.964$, $p < .001$). This result is considerably more noteworthy, as it may be counter to the claim by Jescheniak and Levelt (1994) and Levelt et al. (1999) that frequency only affects retrieval processes subsequent to semantically-driven lexical access. Although sometimes found in normal participants (e.g. Harley & MacAndrew, 2001), only Feyereisen et al. (1988) have found frequency effects on semantic errors in a patient group study, while other patient studies have not found this effect at a group level (Cuetos et al., 2002; Kremin et al., 2003; Nickels & Howard, 1994; Nickels & Howard, 1995). We attribute our positive finding largely to the increased power of our study. It is also important to note that the Semantic-error effect cannot be attributed to Imageability, which was included as a predictor variable (cf. Harley & MacAndrew, 2001).

There was also a significant effect of Log Frequency on Omissions, like in the aphasic picture-naming studies of Kremin et al. (2003) and Feyereisen et al. (1988): An increase in Log Frequency predicted a decrease in making an Omission relative to a Correct response ($-.411$ (.133); $t = -3.095$, $p = .002$). We also ran a multinomial, hierarchical logistic regression analysis in which description omissions were removed from the Omission category and added to Other errors, and found that the Log Frequency effect was significant for “no response” omissions alone ($-.446$ (.173); $t = -2.573$, $p = .010$).

AoA—There was a significant decrease in the log odds of a Correct response per unit increase in AoA ($-.256$ (.043); $t = -5.916$, $p < .001$). This is consistent with the facilitative effect of early-acquired words on normal and aphasic individual and group picture naming accuracy (see Johnston & Barry, 2006; Bradley, Davies, Parris, Su, & Weekes, 2006; Cuetos et al., 2002, 2005; Hirsch & Ellis, 1994; Nickels & Howard, 1995).

Late-acquired words increased the log odds of making a Phonological error relative to a Correct response⁵ (.242 (.055); $t = 4.398$, $p < .001$), but interestingly did not increase the log odds of making a Semantic error relative to a Correct response (.054 (.089); $t = .614$, $p = .539$). By contrast, Cuetos et al. (2002) found an effect of AoA on both phonologically and semantically related errors in a group of 17 aphasics, while in a study of 15 aphasics, Nickels and Howard (1995) found an effect of AoA on semantically related, but not phonologically related picture naming errors.

Late-acquired words also increased the log odds of making an Omission relative to a Correct response (.404 (.064); $t = 6.345$, $p < .001$), including for “no response” omissions alone (.500 (.072); $t = 6.896$, $p < .001$). This is reminiscent of a deep dyslexic’s tendency to omit late-acquired words (Gerhand & Barry, 2000). There is one direct precedent for this finding in the picture naming studies with groups of aphasics: Kremin et al. (2003) found that AoA predicted the absence of a response in 11 fluent/posterior aphasics. Other studies provide indirect support for this result: AoA has been found to predict the naming performance (Cuetos, Asuncion, & Perez, 2005) and recovery (Laganaro, Di Pietro, & Schnider, 2006) of anomic patients, the majority of whose errors are typically omissions.

⁵In a separate multinomial, hierarchical logistic regression analysis, we confirmed that the effect of AoA was significant for both phonological word (.182 (.058); $p = .002$) and nonword (.286 (.078); $p < .001$) errors.

Phonological Neighborhood Density, Length, Imageability, and Name

Agreement—Table 7 presents the coefficients, standard errors, and p-values for effects of Length, Log Density, Imageability, and Name Agreement. As these variables were mainly included to avoid confounds with Log Frequency and AoA, we only provide a brief discussion of results in the expected direction.

Short Length increased the log odds of making a Correct response, as well as decreasing the log odds of making a Phonological or Omission error relative to a Correct response. These results support the general finding that short length (in particular fewer phonemes; see Nickels & Howard, 2004) protects against error production in normal and aphasic individuals (Gordon, 2002; Nickels & Howard, 1994; Nickels & Howard, 1995).

High Log Density decreased the log odds of making a Semantic and Omission error relative to a Correct response. These results are consistent with the simulations of Gordon and Dell (2001) and Dell and Gordon (2003) in which semantic errors occurred less frequently to high-density targets. Also, these results are consistent with density effects on tip-of-the-tongue states in normal individuals (Harley & Bown, 1998), given that these have often been compared to aphasic omissions (Laine et al., 1998).

Although facilitative effects of high imageability have been found in other aphasics (Nickels & Howard, 1994) and deep dyslexics (Plaut & Shallice, 1993), Imageability was not a significant predictor for any of the response types of our 50 aphasics. There was also no significant effect of Name Agreement, although this variable has been found to affect aphasic picture naming in other studies (e.g. Kremin et al., 2003). These results are perhaps unsurprising, given the restricted range of Imageability and Name Agreement values of the PNT target pictures analyzed (see Table 2).

Subgroup analysis of Non-semantic patients

Log Frequency: How are we to be sure that the effect of Log Frequency on Semantic errors can be localized to L-retrieval? Our conservative classification of Non-semantic patients allows us to assume that their Semantic errors most likely arose during L-retrieval. We performed separate hierarchical, multiple logistic regression analyses identical to the overall analyses within this patient group and found that the smaller group mirrored the findings of the overall group (see Table 6 for a summary of the results). Most importantly, Log Frequency significantly affected Non-semantic patients' propensity for making Semantic errors relative to Correct responses, and the coefficient was about the same size as that for 50-patient group ($-.365 (.182)$; $t = -2.005$, $p = .045$). Caramazza and Hillis (1990) report a similar effect of frequency in their patient RGB, whose semantic errors are thought to involve L-retrieval. Also, Non-semantic patients exhibited a significant Log Frequency effect on their likelihood of making a Correct response ($.441 (.085)$; $t = 5.178$, $p < .001$), as well as on making a Phonological ($-.546 (.114)$; $t = -4.773$, $p < .001$) and Omission ($-.586 (.148)$; $t = -3.958$, $p < .001$; for "no response" omissions only, $-.602 (.177)$; $t = -3.396$, $p = .001$) error relative to a Correct response.

AoA: Again, the Non-semantic patients matched the overall pattern of results. They showed a significant effect of AoA on Correct responses ($-.291 (.059)$; $t = -4.909$, $p < .001$). AoA also significantly affected their tendency to make a Phonological ($.310 (.066)$; $t = 4.695$, $p < .001$) and Omission ($.460 (.086)$; $t = 5.358$, $p < .001$; for "no response" omissions only, $.530 (.103)$; $t = 5.136$, $p < .001$) error relative to a Correct response, and, as before, did *not* significantly affect their propensity to make a Semantic error relative to a Correct response ($-.084 (.113)$; $t = -.742$, $p = .458$). Given that the Semantic error coefficient's sign is actually in the wrong direction, this is not a case of an effect in the expected direction that simply failed to reach significance.

Results – contrast analysis

Logistic regression does not allow for the calculation of R^2 values, which provide important information about effect size. To compensate for this, we conducted a series of contrast analyses to ask which variable (Log Frequency or AoA) had a *stronger* effect for a particular response type, and for a given variable, whether the strength of its effect differed for a given pair of response types. We first conducted separate binomial and multinomial logistic regression analyses for each patient (using Stata 8.2; Stata Corporation, 2003), to predict the change in the log odds of a Correct response or the change in the log odds of a Semantic, Phonological, Omission or Other error relative to a Correct response, respectively. We then performed Wilcoxon signed rank tests on the resulting Z test statistics for patient-specific pairs of Log Frequency and AoA coefficients (excluding regression coefficients that could not be estimated due to insufficient data or from patients who made 2 or fewer errors of a given type). We performed one-tailed tests when the direction of the effect was implied from the significance testing in the main analysis (e.g. comparing the effect of Log Frequency, which was significant, to that of AoA, which was not, on Semantic errors), and two-tailed tests otherwise (e.g. comparing the effects of Log Frequency and AoA on Phonological errors).

Consistent with the main analysis, Log Frequency had a stronger effect than AoA on Semantic errors, $n = 36$, W (sum of ranks) = -243 , $p = .028$. There was no significant difference between the strength of the Log Frequency and AoA effects on Phonological errors, $n = 39$, $W = -260$, $p = .069$, Omission errors, $n = 34$, $W = 161$, $p = .170$, or Correct responses, $n = 50$, $W = -12$, $p = .952$.

We also tested contrasts of coefficients for the same lexical variable, from different error types. The effect of Log Frequency was significantly stronger for Phonological than Semantic errors, $n = 32$, $W = -304$, $p = .004$. This last result reinforces the many other studies that have found robust frequency effects on errors and other measures that implicate phonological retrieval. The Log Frequency effect was also stronger for Phonological errors compared to Omissions, $n = 33$, $W = -222$, $p = .047$, but not for Semantic errors compared to Omissions, $n = 32$, $W = 152$, $p = .157$.

Consistent with the main analysis, the effect of AoA was significantly stronger for Phonological than Semantic errors, $n = 32$, $W = 235$, $p = .014$, and for Omissions compared to Semantic errors, $n = 31$, $W = -331$, $p = .001$. Although AoA significantly affected both Omissions and Phonological errors in the main analysis, the AoA effect on Omissions was stronger than the effect of AoA on Phonological errors, but only marginally so, $n = 32$, $W = -208$, $p = .051$.

Summary – analysis I

In both the whole group of 50 patients and within a subset of Non-semantic aphasics, regression analyses revealed that an increase in Log Frequency predicted more Correct responses, as well as fewer Semantic errors, Phonological errors, and Omissions. A decrease in AoA also predicted more Correct responses and fewer Phonological errors and Omissions, but did not influence patients' likelihood of making a Semantic error. The main finding of a supporting contrast analysis was that the effect of Log Frequency was stronger for Phonological than for Semantic errors. Although more limited in scope than Log Frequency, high Log Density and short Length had similarly facilitative effects, with Log Density predicting Semantic errors and Omissions and Length predicting Correct responses, as well as Phonological errors and Omissions. Imageability and Name Agreement failed to significantly predict the incidence of any response type.

Analysis II – Error Lexical Properties

The results of analysis I demonstrate that low-frequency targets are more prone to both Semantic and Phonological errors. Given that the 2-step interactive activation model (as well as most other models of lexical access) envisions lexical retrieval as a competitive process in which targets and potential errors vie for production, it follows that errors, like targets, might also show effects of frequency.

Many authors have reasoned that if production is competitive and frequency-sensitive, errors should be more frequent than targets on average because they are ultimately produced instead of the target (see Garrett, 2001). The traditional way to test this hypothesis has been to compare the average frequency of errors to the average frequency of the targets that elicit them. However, this approach is problematic, as noted by Nickels (1997), if there are many low-frequency targets in the test set: Because low-frequency targets are more error-prone, even errors generated by a random process will be more frequent on average than their targets, creating the false impression of a frequency-sensitive competition. The flip-side also applies: If the targets are all very high in frequency, as is sometimes the case in aphasic picture-naming studies, then even errors generated by a frequency-sensitive process may not be more frequent on average than their targets. The ambiguous results in the literature comparing the mean frequencies of targets and errors are likely to reflect these problems (del Viso, Igoa, & Garcia-Albea, 1991; Gerhand & Barry, 2000; German & Newman, 2004; Gordon, 2002; Harley & MacAndrew, 1995; Harley & MacAndrew, 2001; Vitevitch, 1997; see Nickels, 1997). Our study is no exception: While the median-frequency error produced by our patients has a larger log frequency value (1.36) than the median-frequency PNT target (1.28), this difference is small and would be regarded as equivocal. We submit, though, that the ambiguity of our values is not due to the small size of the difference, but rather because it is inappropriate to use equality of target and error frequencies as the null hypothesis for examining whether frequency matters to the target-error competition.

Given the problems with the method outlined above, how are we to look for frequency-sensitive competition in word production? An error will not necessarily be more frequent than its target on average, because frequency is not the only variable that determines whether or not an error is produced. However, a potential error must have the *chance* to compete with the target, and being higher in frequency increases this chance. For instance, given a high-frequency target, a potential error has a higher probability of being produced if it has a similarly high frequency. In the case of a low-frequency target, the frequency of the error does not have to be as high in order to compete with the target. The logical consequence of these observations is a positive correlation between target and error frequencies, that is, an increase in the frequency of error-eliciting targets should predict an increase in the frequency of the errors. Indeed, studies have found this positive correlation more reliably than a difference in the mean frequencies between targets and errors (Garrett, 2001). Although past researchers have not interpreted it as such (e.g. Garrett, 2001; Harley & McAndrew, 1995), we argue that for the reasons detailed above, the positive correlation is indicative of a competitive, frequency-sensitive lexical access process.

Because this correlation has not been seen as an intuitive byproduct of competition in the lexical access system, we demonstrate that our prediction is borne out in a second simulation study using the Foygel and Dell (2000) model (details in Appendix B). This simulation specifically shows that manipulation of lexical properties (such as frequency) in the lexical-semantic weights for both the target and for a potential semantic error leads to two main effects and an interaction. The main effects are, unsurprisingly, that semantic errors are more likely when the target is lower in frequency, and that they are less likely when the potential semantic error is lower in frequency. The important part is the interaction. On top of these main effects there is

a positive association between target and error frequency, such that with high frequency targets it is all the more important that the potential error be high in frequency in order for the error to happen. Although the simulation, like the one presented in Appendix A, used the Foygel and Dell model, its specific properties (e.g. interactivity) are not important here. What matters is that the most active word is chosen at L-retrieval, that is, that there is a competition. We propose that when this is the case, target and error frequencies (or any other lexical variable) will tend to be positively correlated, *provided that the variable in question truly has an effect on lexical activation at this model stage*. If the variable does not influence the units in competition for production, then we would not expect targets and errors to show a positive correlation of that variable. In analysis II, we test this prediction experimentally by asking whether lexical properties of the targets predict lexical properties of the errors.

Results

In keeping with the focus of analysis I, we limit our investigation to the influences of Log Frequency and AoA, and to the lexical properties of Semantic and Phonological word errors (by definition we cannot examine the lexical properties of Omissions or Phonological nonword errors). We identified 102 target-Phonological word error pairs and 175 target-Semantic error pairs for which there were Log Frequency, AoA, Imageability, and Length values, for a total of 277 target-Overall word error pairs. We obtained lexical property values for the errors from the same databases as for the targets. In order to isolate a particular error property, we first regressed the property of interest (e.g. Log Frequency) on the other error properties (e.g. AoA, Imageability and Length), yielding residuals that represent, for example, the part of Log Frequency that is not related to AoA, Imageability, or Length. In a second set of analyses, we used hierarchical, multiple regression to predict the error lexical property residuals of interest from target Log Frequency, AoA, Imageability, and Length. We performed a separate regression for each of the 4 error property residuals (Log Frequency, AoA, Length, and Imageability residuals), within Semantic and Phonological word errors, as well as across these word error types (Overall word errors). Here we discuss the statistical estimation process and results of this second set of analyses.

As in analysis I, the independent variables were group-mean centered and HLM was used to fit models to the data. However, in these analyses the dependent variable (error lexical property residuals) was continuous, and so HLM estimates the regression coefficients using generalized least squares estimation, instead of the penalized quasi-likelihood method used in analysis I. The model is specified as:

$$Y = \beta_0 + \mathbf{x}'_{ij}\beta + u_i + r_{ij},$$

where $i = 1, \dots, 50$ indexes subjects, and j indexes trials within subject (which varied across subjects). Y represents the relevant set of lexical property residuals, β_0 is the intercept, \mathbf{x}'_{ij} is the vector of lexical variables, and β gives the regression coefficients. The u_i and r_{ij} terms are each assumed to be normally distributed from a multivariate normal with mean 0 and an unstructured covariance matrix.

We present the results of significance tests (with robust standard errors) for the coefficient estimates, which in this case reflect the average increase in, for example, the Log Frequency residuals of a Semantic error due to a one-unit increase in the Log Frequency of the target. As in analysis I, we only discuss coefficients that are significantly different from 0 with p less than .05 in the expected direction. Additionally, we only consider results for which the predictor and predicted variables are the same (e.g. an increase in target Log Frequency predicting an

increase in error Log Frequency), due to the difficulty of interpreting effects between different variables. Coefficients that did not meet these criteria appear in Tables 8 and 9.

Log Frequency—In the analysis of Overall word errors (.266 (.057); $t = 4.631$, $p < .001$) and of Semantic errors (.391 (.076); $t = 5.124$, $p < .001$), an increase in target Log Frequency predicted an increase in error Log Frequency. This result corresponds to the positive correlation observed in Appendix B, as well as with several reports of positive target-error frequency correlations in normal and aphasic participants (see Garrett, 2001; Hotopf, 1980; Raymond & Bell, 1996). It also complements the finding from analysis I, that Semantic errors are more likely to low-frequency targets. Together, the results suggest a lexical access process in which high frequency increases the chances for potential semantic errors to compete with their targets. We found no significant relationship between target and Phonological error (only including phonological errors making words) Log Frequency (.109 (.088); $t = 1.241$, $p = .218$). In contrast, Gordon (2002) found a significant correlation between the log frequency of PNT targets and the form-related word errors of 43 aphasics. Our null result may be due to the fact that we were only able to obtain error lexical property values for 29% of the Phonological (word) errors made to the targets analyzed in analysis I, as compared to 58% of the Semantic errors made to those targets.

AoA—Target AoA did not significantly predict error AoA for Overall word errors ($-.010$ (.055); $t = -.191$, $p = .849$), Semantic errors (.036 (.068); $t = .531$, $p = .596$), or Phonological (word) errors ($-.034$ (.086); $t = -.398$, $p = .691$). The negative finding with Semantic errors is consistent with the lack of an AoA effect on Semantic errors in analysis I. Gerhand and Barry (2000) found that their deep dyslexic patient's semantic errors were more early-acquired than their targets, but this effect could not be deconfounded with length, and more importantly they did not examine the correlation between target and error AoA. Also, although AoA did have an effect on Phonological errors in analysis I, the small percentage of Phonological *word* errors examined in this analysis may be responsible for the absence of such an effect here.

Imageability and Length—Although the results for Imageability and Length (Tables 8 and 9) were not the focus of our analyses, it is worth noting that there was a significant tendency for target Length to predict Overall word error Length. This is consistent with the positive correlation between target and form-related error lengths found by Gordon (2002) in a group of aphasics. This suggests that Length also influences the competition between targets and potential errors. However, the p-value is rather large ($p = .021$) and thus less convincing that those for Log Frequency, especially given the number of different regressions being performed. Other effects of Length and Imageability were not significant.

Summary – analysis II

For Overall word errors as well as within Semantic errors (but not within Phonological word errors), an increase in the Log Frequency of the target significantly predicted an increase in error Log Frequency. However, no such relationship between targets and errors was found for AoA, for any of the error types. There was also a significant effect of target Length on Overall word error Length, but no other effects of Length or Imageability were significant.

General Discussion

In this study we investigated the influence of target Log Frequency and AoA on aphasic picture naming in a series of hierarchical, multiple logistic regressions, including multinomial regressions of particular error types. We found that high Log Frequency and early AoA predicted greater accuracy, independently of other correlated variables (Log Density, Length, Imageability, and Name Agreement). However, Log Frequency and AoA differed both in their

impact on the probability of error types in analysis I, and in their ability to predict lexical properties of the errors in analysis II. Frequency had extensive effects in both analyses: High target Log Frequency predicted fewer Semantic, Phonological, and Omission errors (analysis I), and high error Log Frequency within Overall word errors and Semantic errors alone (analysis II). The hypothesized L-retrieval locus of the Semantic error effect was supported by the presence of the effect in the Non-semantic patient group. Also, the contrast analysis revealed that the effect of Log Frequency was stronger for Phonological than Semantic errors, consistent with past literature. The effects of AoA were similar to those of frequency in some respects, and different in others. Early AoA, like high Log Frequency, predicted fewer Phonological and Omission errors (analysis I). AoA did not, however, have a noticeable influence on Semantic errors in terms of either target susceptibility to error (analysis I) or target-error correlation (analysis II). The presence of these influences of frequency and AoA on errors from a diverse sample of aphasics attests to their robustness, while the additional support afforded by the Non-semantic group analysis highlights the importance of analyzing subsets of patients with more circumscribed impairments.

Theoretical Implications

The results make it clear that the effect of frequency in picture naming cannot be attributed to AoA (contrary to Bonin, Chalard, Meot, & Fayol, 2002). Also, the finding that target Log Frequency predicts fewer Semantic as well as fewer Phonological errors is striking. Our first computational simulation allows us to interpret this as an effect of frequency on both steps of word retrieval, revealing a lexical access process that adapts to frequency at all stages (e.g. Knobel, Finkbeiner, & Caramazza, in press; Seidenberg & McClelland, 1989). In addition to the behavioral studies cited earlier, this view is also consistent with neuroimaging evidence showing that word frequency, concept familiarity, and length modulate the activity of overlapping brain areas (Graves et al., 2007). The target-error frequency correlation, predicted by our second simulation and subsequently found for Semantic errors and Overall word errors, further supports this interpretation and suggests a competitive, frequency-sensitive error production process.

Our interpretation of the effects of frequency on Semantic errors as evidence for its role in the mapping from meaning to holistic word representations was based in part on simulations of the 2-step interactive model of lexical access (Appendices A and B). It should be noted, though, that these simulations are not the ideal vehicle for exploring the development of frequency-sensitive representations. Frequency differences were imposed upon the model, and not allowed to develop with training as in other connectionist networks (e.g. Ellis & Lambon Ralph, 2000; Seidenberg & McClelland, 1989). More generally, the 2-step interactive activation model (and for that matter, *any* model) cannot capture the full complexity of the real-life lexical access system. For example, it is not inconceivable that the interactive feedback from phonological to lexical units is so strong in a real lexicon that the effect of frequency on semantic errors is entirely due to frequency-sensitive p weights, whose influence is fed back to higher levels. However, other prominent models of speech production allow for *less* feedback from word nodes to semantic features than the 2-step interactive model (e.g. Caramazza, 1997; Rapp & Goldrick, 2000; Levelt et al., 1999). Our results are thus difficult to accommodate within a feed-forward or feedback-restricted model in which frequency only affects production processes that occur after semantically driven lexical access (e.g. Levelt et al., 1999). Aside from this important conclusion, our data do not speak directly to questions concerning the architecture of the language production system, such as whether there are one or two holistic lexical representations (Caramazza, 1997; Miozzo & Caramazza, 1997), or whether there is feedback or cascading of activation (Rapp & Goldrick, 2000).

Turning to our results concerning AoA, a key finding was that late AoA, like low Log Frequency, predicted more Phonological errors. This is consistent with effects of AoA on *word* naming, which have been interpreted as evidence that AoA is important for phonological retrieval. For example, one widely cited paper proposes that early-acquired words are represented in a relatively complete form in the phonological output lexicon (e.g. children do not make use of redundancy in word endings), while the phonological representations of late-acquired words store less information explicitly, forcing their pronunciation to be generated anew upon each retrieval (Brown & Watson, 1987). Although this particular account is now largely out of favor, the evidence against it (e.g. Monaghan & Ellis, 2002) does not exclude other mechanisms whereby AoA might influence phonological retrieval. Barry et al. (2001) discuss the possibility that phonological representations of early-acquired words may have higher resting activation levels than those acquired later in life, making them easier to retrieve (see also Hernandez & Fiebach, 2006).

AoA, as well as Log Frequency, additionally had an effect on Omissions: These errors were more likely to late-acquired, low-frequency targets. How can we interpret these effects? Unfortunately, omissions are poorly understood and thus theories of their origin are at best incomplete. They have traditionally been related to tip-of-the-tongue (TOT) states in normal participants, which are generally thought to result from failed phonological retrieval (Harley & Bown, 1998). Associating our Omissions with this second step of lexical access is quite consistent with the AoA and Log Frequency effects on Phonological errors. However, omissions, like TOTs, could result from faulty processing at a higher level. Gollan and Brown (2006) demonstrated that it is possible to separate omissions (made by unimpaired speakers providing names for pictures) into 2 categories: Failures to complete phonological retrieval and failures to complete L-retrieval. It is also conceivable that impaired pre-lexical, conceptual processing might result in an omission. The Omissions analyzed in this study may have been generated at more than one, or perhaps all of these loci, depending on the impairment of the patient producing them. Thus, our results may also be consistent with theories that locate AoA at a pre-lexical, conceptual level (e.g. Steyvers & Tenenbaum, 2005).

In contrast to Log Frequency, AoA did not appear to affect the first step of lexical access: It did not influence patients' likelihood of making a Semantic error, a finding that also held true for the Non-semantic group. This is consistent with the fact that there is not much direct evidence for effects of AoA on L-retrieval (although see Bates, Devescovi, Pizzamiglio, D'Amico, & Hernandez, 1995). For instance, Brysbaert and Ghyselinck (2006) suggest that the literature is consistent with either a lexical-semantic *or* conceptual locus of AoA. However, in the influential model of Ellis and Lambon Ralph (2000), both frequency and AoA affect the extent to which the weights between inputs and outputs change in response to training. As these effects are seen in simulations with a variety of frequency levels and weight decay strengths, the authors make the prediction that any task affected by AoA should also be affected by frequency. This proposal would seem to be at odds with our results: AoA and frequency influenced patients' overall accuracy, but only frequency affected both phonological *and* semantic errors. However, as noted by Cuetos et al. (2002), Ellis and Lambon Ralph's (2000) network is not a model of word retrieval with distinct L-retrieval and phonological retrieval steps. Consequently it is unclear (although implied) that Ellis and Lambon Ralph's model necessarily predicts effects of both variables on errors arising at each stage of lexical access.

In fact, our results are consistent with the model of Ellis and Lambon Ralph (2000) if we consider Dell and Kim's (2005) suggestion that in a model with distributed L-representations, L-retrieval is a less arbitrary mapping than phonological retrieval. L-level representations are thought to reflect word-specific grammatical information, and words that are similar in meaning tend to have similar grammatical properties, e.g. verbs with similar meanings ("love" and "like") will have more similar subcategorization constraints. By contrast, there is no such

tendency for words from similar grammatical classes to have similar sounds. If L-retrieval can be seen as a less arbitrary mapping than phonological retrieval, then according to Ellis and Lambon Ralph (2000), AoA should have a larger effect on phonological than semantic errors. This is the case in our data, both for AoA (which had a significant effect on Phonological, but not Semantic errors in the main analysis) and Log Frequency (which had a stronger effect on Phonological than Semantic errors, as revealed in the contrast analysis).

There is also some support for a separation of Log Frequency and AoA effects from recent neuroimaging studies using the lexical decision task. In an fMRI study, Fiebach, Friederici, Muller, von Cramon, & Hernandez (2003) found that differential brain areas were activated by late- and early-learned words, and that these areas were not modulated by word frequency. However, Tainturier, Tamminen, and Thierry (2005) recorded Event-Related Potentials (ERPs) and found that AoA modulated the amplitude of the P300 in a way very similar to that of frequency in previous studies. This is reminiscent of the similar effect size of AoA and frequency on lexical decision reaction times (Morrison & Ellis, 1995). It remains to be seen whether differential effects of AoA and frequency on ERPs can be found with more meaning-based tasks.

Integrating our findings with those of past aphasic picture naming studies

How do our results and conclusions about frequency and AoA mesh with the findings of other, similar studies? Consistent with our results, Feyereisen et al. (1988) found that low-frequency pictures elicited more semantic and phonological errors (as well as omissions) when AoA was held constant. However, three other aphasic picture naming studies that controlled for more confounding variables have yielded different outcomes. For example, while the 11 aphasics in Kremin et al. (2003) exhibited effects of frequency and AoA on their omissions, the omissions of the 16 aphasics in Cuetos et al. (2002) only showed effects of visual complexity and familiarity. However, it is unlikely that familiarity and visual complexity were confounded with the effects of Log Frequency and AoA on Omissions in our study, because Kremin et al. (2003) controlled for these variables and still found effects similar to ours.

In these previous studies, phonological errors are predicted either by visual complexity (Kremin et al., 2003), familiarity and length (15 aphasics in Nickels & Howard, 1995), or AoA alone (Cuetos et al., 2002). Given that Cuetos et al. (2002) controlled for familiarity and visual complexity, it is unlikely that these variables are confounded with the AoA effect on Phonological errors in our study. While it is possible that the effect of Log Frequency on the Phonological errors of our patients could be confounded with visual complexity and familiarity, this is highly unlikely given past research. Frequency has an influence on normal participants' picture naming reaction times when visual complexity and familiarity are controlled (Jescheniak & Levelt, 1994; Lambon Ralph & Ehsan, 2006), not to mention a well-documented effect on tasks not involving pictorial stimuli (e.g. Jescheniak & Levelt, 1994).

As for semantic errors, Kremin et al. (2003) failed to find any variables that significantly predicted them, while both Cuetos et al. (2002) and Nickels and Howard (1995) found significant effects of AoA. By contrast, Semantic errors in our study did not exhibit an effect of AoA, suggesting that AoA does not influence L-retrieval. One might argue that the effects of AoA found for each of phonological, semantic, and omission errors in two of these four studies are suggestive of a more distributed effect of AoA. However, given the effect of AoA on Phonological errors in our study, it is unlikely that insensitivity of our AoA measure is responsible for the absence of an effect on Semantic errors. Also, if a significant proportion of the patients in Cuetos et al. (2002) and Nickels and Howard (1995) had central semantic damage, it is possible that the effects of AoA on their semantic errors could reflect a pre-lexical, conceptual influence (Steyvers & Tenenbaum, 2005).

Differences in the measures of both frequency and AoA, the languages spoken by the patients, the stimuli, and the particular groups of aphasics used are no doubt responsible for some of the discrepancies among the studies reviewed above. Amidst this variability, we suggest that our findings are particularly worthy of consideration. In addition to including many different lexical variables, we examined a much larger sample of patients with a more appropriate and sophisticated statistical analysis. Multinomial, hierarchical multiple regression is an effective tool for analyzing error data from more than one patient when there are multiple error categories, as well as several independent variables. Importantly, our analysis of a Non-semantic group of patients, as well as the results of analysis II, confirm and thus lend credibility to the findings of analysis I.

Conclusion

In sum, this large-scale regression analysis of aphasic picture naming errors has revealed a number of effects highly relevant to theories of lexical access. Age of acquisition and lexical frequency both emerged as significant predictors of aphasic picture naming accuracy and seemed to be important for multiple (and possibly, nonidentical) stages of lexical access. Most strikingly, lexical frequency not only affected phonological retrieval but also L-retrieval, contrary to one prominent view. This is compatible with suggestions in the language comprehension literature that frequency of occurrence affects all levels of processing: Increased experience with a word boosts the retrievability of its abstract and phonological representations, making it more likely to be produced. AoA influenced phonological retrieval but not L-retrieval, and both frequency and AoA affected the likelihood of making an omission, a little understood but common error type in aphasia and normal speech (in the form of transient “tip-of-the-tongue” states). These results are consistent with theories that locate the effect of AoA at the retrieval of a word’s form, but may also be compatible with theories that place AoA in the conceptual system. Further empirical and computational work is needed to determine whether this pattern of results is truly due to the lack of an AoA effect on the first step of lexical access, or whether it, like frequency, is represented at multiple stages.

In closing, we recommend that the further investigation needed to resolve these questions consider, to a greater extent, the interrelationships that exist among multiple error categories in neuropsychological data. If a patient makes error type A, not only is the response incorrect, but it also necessarily is not error type B. While computational models of aphasic error patterns have respected the multinomial nature of the process that generates responses, empirical investigations of aphasic errors typically have not. Analytic procedures based on regression usually predict a single response type at a time, largely because of the unavailability of suitable multinomial software. It is now possible to do multinomial logistic regression in a hierarchical fashion, as we have done, and we therefore recommend adoption of this tool in analyses of multiple response types made by neuropsychological patients.

Acknowledgments

This research was supported by the National Institutes of Health (DC000191, HD44458, and MH1819990). We would like to thank Paula Sobel, Gemma Baldon, and Grant Walker at the Moss Rehabilitation Institute. We would also like to thank Carolyn Anderson for help with the statistical analyses. Finally, we would like to thank Andrew Ellis, Matthew Lambon Ralph, Michele Miozzo, and two anonymous reviewers for their suggestions and comments, as well as Kathryn Bock for reading an earlier version of this manuscript.

References

- Alario FX, Costa A, Caramazza A. Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes* 2002;17:299–319.
- Baayen, RH.; Piepenbrock, R.; van Rijn, H. The CELEX lexical database (CD-ROM). Pennsylvania: Linguistic Data Consortium, University of Pennsylvania, PA; 1993.

- Barry C, Hirsh KW, Johnston RA, Williams CL. Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of Memory and Language* 2001;44:350–375.
- Bates E, Devescovi A, Pizzamiglio L, D'Amico S, Hernandez A. Gender and lexical access in Italian. *Perception and Psychophysics* 1995;57:847–862. [PubMed: 7651809]
- Bonin P, Chalard M, Meot A, Fayol M. The determinants of spoken and written picture naming latencies. *British Journal of Psychology* 2002;93:89–114. [PubMed: 11839103]
- Bradley V, Davies R, Parris B, Su IF, Weekes BS. Age of acquisition effects on action naming in progressive fluent aphasia. *Brain and Language* 2006;99:117–118.
- Brown GD, Watson FL. First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition* 1987;15:208–216.
- Brysbaert M, Ghyselinck M. The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition* 2006;13:992–1011.
- Cameron, AC.; Trivedi, PK. *Microeconometrics: Methods and Applications*. New York: Cambridge; 2005.
- Caramazza A. How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 1997;14:177–208.
- Caramazza A, Bi Y, Costa A, Miozzo M. What determines the speed of lexical access: Homophone or specific-word frequency? A reply to Jescheniak et al. (2003). *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2004;30:278–282.
- Caramazza A, Costa A, Miozzo M, Bi Y. The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2001;27:1430–1450.
- Caramazza A, Hillis AE. Where do semantic errors come from? *Cortex* 1990;26:95–122. [PubMed: 2354648]
- Cuetos F, Aguado G, Izura C, Ellis AW. Aphasic naming in Spanish: Predictors and errors. *Brain and Language* 2002;82:344–365. [PubMed: 12160529]
- Cuetos F, Asuncion M, Perez A. Determinants of lexical access in pure anomia. *Journal of Neurolinguistics* 2005;18:383–399.
- Dahan D, Magnuson JS, Tanenhaus MK. Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology* 2001;42:317–367. [PubMed: 11368527]
- del Viso S, Igoa JM, Garcia-Albea JE. On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research* 1991;20:161–185.
- Dell GS. Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes* 1990;5:313–349.
- Dell, GS.; Kim, AE. Speech errors and word form encoding. In: Hartsuiker, RJ.; Bastiaanse, R.; Postma, A.; Wijnen, F., editors. *Phonological encoding and monitoring in normal and pathological speech*. Hove, U.K.: Psychology Press; 2005.
- Dell, GS.; Gordon, JK. Neighbors in the lexicon: Friends or foes?. In: Schiller, NO.; Meyer, AS., editors. *Phonetics and phonology in language comprehension and production: Differences and similarities*. New York: Mouton de Gruyter; 2003. p. 9–37.
- Dell GS, O'Seaghdha PG. Stages of lexical access in language production. *Cognition* 1992;42:287–314. [PubMed: 1582160]
- Dell GS, Lawler EN, Harris HD, Gordon JK. Models of errors of omission in aphasic naming. *Cognitive Neuropsychology* 2004;21:125–145.
- Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and non aphasic speakers. *Psychological Review* 1997;104:801–838. [PubMed: 9337631]
- Ellis AW, Lambon Ralph MA. Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2000;26:1103–1123.
- Fiebach CJ, Friederici AD, Muller K, von Cramon DY, Hernandez AE. Distinct brain representations for early and late learned words. *Neuroimage* 2003;19:1627–1637. [PubMed: 12948717]

- Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. *Monographs of the Society for Research in Child Development* 1994;59:1–173. [PubMed: 7845413]
- Ferreira VS, Pashler H. Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2002;28:1187–1199.
- Feyereisen P, Van der Borgh F, Seron X. The operativity effect in naming: A re-analysis. *Neuropsychologia* 1988;26:401–415. [PubMed: 2453813]
- Foygel D, Dell GS. Models of impaired lexical access in speech production. *Journal of Memory and Language* 2000;43:182–216.
- Gagnon DA, Schwartz MF, Martin N, Dell GS, Saffran EM. The origins of formal paraphasias in aphasics' picture naming. *Brain and Language* 1997;59:450–472. [PubMed: 9299072]
- Gahl, S. Is frequency a property of phonological forms? Evidence from spontaneous speech. Paper presented at the CUNY Sentence Processing Conference; New York, NY. 2006.
- Garrett, M. Now you see it, now you don't: Frequency effects in language production. In: Dupoux, E., editor. *Language, brain, and cognitive development: Essays in honor of Jacques Mehler*. Cambridge, MA: MIT Press; 2001. p. 227-240.
- Gerhand S, Barry C. When does a deep dyslexic make a semantic error? The roles of age-of-acquisition, concreteness, and frequency. *Brain and Language* 2000;74:26–47. [PubMed: 10924215]
- German DJ, Newman RS. The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research* 2004;47:624–636.
- Ghyselinck M, Lewis MB, Brysbaert M. Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica* 2004;115:43–67. [PubMed: 14734241]
- Gollan TH, Brown AS. From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General* 2006;135:462–483. [PubMed: 16846276]
- Gordon JK. Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language* 2002;82:113–45. [PubMed: 12096871]
- Gordon JK, Dell GS. Phonological neighborhood effects: Evidence from aphasia and connectionist modeling. *Brain and Language* 2001;79:21–23.
- Graves WW, Grabowski TJ, Mehta S, Gordon JK. A neural signature of phonological access: Distinguishing the effects of word frequency from familiarity and length in overt picture naming. *Journal of Cognitive Neuroscience* 2007;19:617–631. [PubMed: 17381253]
- Griffin ZM, Bock K. Constraint, word frequency, and the relationship between lexical processing levels in spoken word recognition. *Journal of Memory and Language* 1998;38:313–338.
- Haley KL, Ohde RN, Wertz RT. Vowel quality in aphasia and apraxia of speech: Phonetic transcription and formant analyses. *Aphasiology* 2001;15:1107–1124.
- Harley TA, Bown HE. What causes a tip-of-the-tongue state? Evidence for lexical neighborhood effects in speech production. *British Journal of Psychology* 1998;89:151–174.
- Harley, TA.; MacAndrew, SBG. Interactive models of lexicalization: Some constraints from speech error, picture naming, and neuropsychological data. In: Levy, JP.; Bairaktaris, D.; Bullinaria, JA.; Cairns, P., editors. *Connectionist models of memory and language*. London: UCL Press; 1995. p. 311-331.
- Harley TA, MacAndrew SBG. Constraints upon word substitution speech errors. *Journal of Psycholinguistic Research* 2001;30:395–418. [PubMed: 11529522]
- Hernandez AE, Fiebach CJ. The brain bases of reading late learned words: Evidence from functional MRI. *Visual Cognition* 2006;13:1027–1043.
- Hirsh KW, Ellis EW. Age of acquisition and lexical processing in aphasia: A case study. *Cognitive Neuropsychology* 1994;11:435–458.
- Hotopf, WHN. Semantic similarity as a factor in whole-word slips of the tongue. In: Fromkin, V., editor. *Errors in linguistic performance: Slips of the tongue, ear, pen and hand*. New York: Academic Press; 1980. p. 97-109.

- Jescheniak JD, Levelt WJM. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1994;20:824–43.
- Jescheniak JD, Meyer AS, Levelt WJM. Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2003;29:432–438.
- Johnston RA, Barry C. Age of acquisition and lexical processing. *Visual Cognition* 2006;13:789–845.
- Kertesz, A. *Western Aphasia Battery*. New York: Grune & Stratton; 1982.
- Knobel M, Finkbeiner M, Caramazza A. The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production. *Cognitive Neuropsychology*. in press.
- Kremin H, Lorenz A, de Wilde M, Perrier D, Arabia C, Labonde E, Buitoni CL. The relative effects of imageability and age-of-acquisition on aphasic misnaming. *Brain and Language* 2003;87:33–34.
- Kutner, MH.; Nachtsheim, CJ.; Neter, J.; Li, W. *Applied Linear Statistical Models*. New York, NY: McGraw-Hill/Irwin; 2004.
- Laganaro M, Di Pietro M, Schnider A. What does recovery from anomia tell us about the underlying impairment: The case of similar anomic patterns and different recovery. *Neuropsychologia* 2006;44:534–545. [PubMed: 16112147]
- Laine M, Tikkala A, Juhola M. Modelling anomia by the discrete two-stage word production architecture. *Journal of Neurolinguistics* 1998;11:275–294.
- Lambon Ralph MA, Ehsan S. Age of acquisition effects depend on the Mapping between representations and the frequency of occurrence: Empirical and Computational evidence. *Visual Cognition* 2006;13:928–948.
- Laubstein AS. Lemmas and lexemes: The evidence from blends. *Brain and Language* 1999;68:135–143. [PubMed: 10433750]
- Levelt, WJM. *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press; 1989.
- Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 1999;22:1–75. [PubMed: 11301520]
- Long JS, Ervin LH. Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *American Statistician* 2000;54:217–224.
- Luce PA, Pisoni DB. Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing* 1998;19:1–36. [PubMed: 9504270]
- Martin N, Schwartz MF, Kohen FP. Assessment of the ability to process semantic and phonological aspects of words in aphasia: A multi measurement approach. *Aphasiology* 2005;20:154–166.
- Miozzo M, Caramazza A. Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1997;23:1410–1423.
- Monaghan J, Ellis AW. Age of acquisition and the completeness of phonological representations. *Reading and Writing* 2002;15:759–788.
- Morrison CM, Chappell TD, Ellis AW. Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology* 1997;50:528–559.
- Morrison CM, Ellis AW. Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1995;21:116–133.
- Navarette E, Basagni B, Alario FX, Costa A. Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology* 2006;59:1681–1690. [PubMed: 16945853]
- Nickels, L. *Spoken word production and its breakdown in aphasia*. East Sussex, UK: Psychology Press; 1997.
- Nickels L, Howard D. A frequent occurrence? Factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology* 1994;11:289–320.
- Nickels L, Howard D. Aphasic naming: What matters? *Neuropsychologia* 1995;33:1281–1303. [PubMed: 8552229]

- Nickels L, Howard D. Dissociating effects of number of phonemes, number of syllables, and syllabic complexity on word production in aphasia: It's the number of phonemes that counts. *Cognitive Neuropsychology* 2004;21:57–78.
- Oldfield RC, Wingfield A. Response latencies in naming objects. *The Quarterly Journal of Experimental Psychology* 1965;17:273–281. [PubMed: 5852918]
- Plaut DC, Shallice T. Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology* 1993;10:377–500.
- Rapp, B., editor. *The handbook of cognitive neuropsychology: What deficits reveal about the human mind*. New York: Psychology Press; 2001.
- Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. *Psychological Review* 2000;107:460–499. [PubMed: 10941277]
- Raudenbush, SW.; Bryk, AS. *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage; 2002.
- Raudenbush, SW.; Bryk, AS.; Cheong, YF.; Congdon, RT. *HLM 6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International; 2004.
- Raymond W, Bell A. Frequency effects in malapropisms. *The Journal of the Acoustical Society of America* 1996;100:2572.
- Roach A, Schwartz MF, Martin N, Grewal RS, Brecher A. The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology* 1996;24:121–133.
- Santo Pietro MJ, Rigoordsky S. The effects of temporal and semantic conditions on the occurrence of the error response of perseveration in adult aphasics. *Journal of Speech and Hearing Research* 1982;25:184–192. [PubMed: 7120957]
- Schriefers H, Meyer AS, Levelt WJ. Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language* 1990;29:86–102.
- Schwartz MF, Wilshire CE, Gagnon DA, Polansky M. Origins of nonword phonological errors in aphasic picture naming. *Cognitive Neuropsychology* 2004;21:159–186.
- Schwartz MF, Dell GS, Martin N, Gahl S, Sobel P. A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language* 2006;54:228–264.
- Seidenberg MS, McClelland JL. A distributed, developmental model of word recognition and naming. *Psychological Review* 1989;96:523–568. [PubMed: 2798649]
- Stata Corporation. *Stata 8.2 for Windows*. Texas: Stata Corporation; 2003.
- Steyvers M, Tenenbaum JB. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science* 2005;29:41–78.
- Szekely A, Jacobsen T, D'Amico S, Devescovi A, Andonova E, Herron D, Lu CC, Pechmann T, Pleh C, Wicha N, Federmeier K, Gerdjikova I, Gutierrez G, Hung D, Hsu J, Iyer G, Kohnert K, Mehotcheva T, Orozco-Figueroa A, Tzeng A, Tzeng O, Arevalo A, Vargha A, Butler AC, Buffington R, Bates E. A new on-line resource for psycholinguistic studies. *Journal of Memory and Language* 2004;51:247–250.
- Tainturier MJ, Tamminen J, Thierry G. Age of acquisition modulates the amplitude of the P300 component in spoken word recognition. *Neuroscience Letters* 2005;379:17–22. [PubMed: 15814191]
- Vitevitch MS. The neighborhood characteristics of malapropisms. *Language and Speech* 1997;40:211–228. [PubMed: 9509578]
- Vitkovitch M, Humphreys GW. Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1991;17:664–680.
- Wilson MD. *The MRC psycholinguistic database: Machine readable dictionary, version 2*. Behavioural Research Methods, Instruments and Computers 1988;20:6–11.
- Yule, GU.; Kendall, MG. *An Introduction to the Theory of Statistics*. 14. London: Griffin; 1950.

Appendix A – Simulating effects of lexical variables at different lexical access stages

We compared the performance of the original Foygel and Dell (2000) model with 2 modified models: The “s-manipulated” model featured a reduction of the target (“cat”)’s bidirectional semantic-word weights⁶ by 80% to simulate an effect of a lexical variable on L-retrieval, while in the “p-manipulated” model the target’s word-phonological weights were reduced by 80% to simulate an effect on phonological retrieval. The weights of all other words in the lexicon were not modified. We created s- and p-manipulated models for each of 8 baseline *s* and *p* weight values, to simulate a continuum from normal to severely impaired performance and varying levels of relative *s* and *p* weight damage⁷.

Table 10 presents the baseline *s* and *p* values for the 8 models and, for each model, presents the increase in the proportion of semantic and nonword errors induced by the reduction of the target word’s semantic and phonological weights. Across severities and levels of damage, the s-manipulated model shows a large increase in semantic errors (except for the “severe” model, in which the lack of semantic errors does not allow the effect to emerge), while there is little or no increase in nonword errors. In the p-manipulated model, there is a medium to large increase in nonword errors across severities (except for the “normal” model, which hardly produces any nonword errors to begin with), but there is little effect on semantic errors. This overall pattern of results is demonstrated in Figure 2, which averages the increases in error proportions across the different *s* and *p* baseline weight values. Clearly, there is a very large degree of selective influence—phonological manipulations affect phonological (nonword) errors, and semantic manipulations affect semantic errors. The only deviation from selective influence occurs when the *s* weight is much larger than the *p* weight (as in “mod-mild” and “sev-mod” models with greater *p* damage). Here, manipulation of the target’s *s* weights exerts a small influence on nonword errors (e.g. an increase of .031 in the $s=.03, p=.02$ model, and an increase of .024 in the $s=.025, p=.015$ model). However, because our main theoretical concern is with implications of (e.g. frequency) effects on *semantic* errors for the locus of such effects in the production system, this not a concern. What is key is that semantic errors are, in the model, relatively unaffected by manipulation of the target’s phonological weights. Hence, effects on semantic errors cannot be attributed to variations in phonological weights. For more extensive investigations of the possible loci of frequency in models similar to the Foygel and Dell model, see Knobel et al. (in press).

Appendix B – Simulations manipulating lexical properties of a target and a potential semantic error

We created 4 versions of the Foygel and Dell (2000) model, orthogonally varying the *s* weights (e.g. representing frequency or AoA in the semantic-lexical mapping) of the target “cat” and its semantic competitor “dog”. The weights of other words in the model’s lexicon were not modified. The baseline (e.g. high-frequency) value for all *p* weights in the model was set to .04, while the baseline value for all *s* weights ranged from .04 to .01, to simulate a continuum

⁶Notice, that, logically, one can also place frequency in hypothesized resting levels of activation for lexical units (e.g. Dell, 1990). This is, for our purposes, equivalent to putting frequency in the lexical-semantic weights because the manipulation has its impact *prior* to the selection of the word unit. As Knobel et al. (in press) have shown using variants of the Foygel and Dell model, manipulation of a target lexical resting level is largely equivalent to manipulation of lexical-semantic weights to and from that unit. There is only a miniscule effect of variations in resting levels and lexical-semantic weights on the model’s phonological decisions because the word selected in the L-retrieval stage is given a large jolt of activation and this diminishes influences of lexical-semantic weights and lexical resting levels on phonological retrieval.

⁷We also conducted these simulations with an intermediate low-frequency weight value of 90% of baseline, and found that the results were approximately halfway between those of the baseline models and the 80% weight-manipulated models presented here. Thus, variations in weights in the model create gradient effects on error probability.

from normal to severely impaired L-retrieval. Only the s weights were varied to simulate severity because, as demonstrated in Appendix A, these exert the most influence on semantic errors. Moreover, it is important to use models with a high p weight so that semantic errors created during the model's L-retrieval are not "lost" by distortion due to poor phonological retrieval. That is, we want errors such as CAT→DOG, not CAT→MOG. To simulate a e.g. low frequency word, s weights of the target and/or competitor were set to 60% of their baseline value.

Tables 11a–d each present results from simulations with a different baseline s weight value. Each table shows the number of semantic errors generated over 10,000 runs by the 4 models with the specified target and competitor s weights. The tables exhibit two main effects, which we will describe in terms of frequency for the sake of simplicity. Across the range of severities, when the target is low-frequency ("Reduced" in the tables), there are more semantic errors than when the target is high-frequency ("Baseline" in the tables), regardless of whether the competitor is high- or low-frequency. Likewise, when the competitor is low-frequency, there are fewer semantic errors overall compared to when it is high-frequency. To test whether there is also an interaction between target and error frequency (that is, a lack of statistical independence), we calculated Yule's Q (a statistic ranging from -1 to 1 , for which an absolute value of 1 implies a perfect association and 0 no association; Yule & Kendall, 1957) for each table. On average, we found a mild positive association between target and error frequency, although this dropped to near zero for the most severe models, in which lexical access was dominated by noise (Table 11a, $Q = .61$; Table 11b, $Q = .23$; Table 11c, $Q = .07$; Table 11d, $Q = .02$). Thus, these simulations show that when frequency is manipulated in the s weights, there is a positive correlation between target and semantic error frequency. One can imagine an analogous result if we were to simulate a frequency effect on phonological errors by manipulating the p weights. Moreover, if AoA inhabits s and/or p weights in the same way, positive correlations on this variable are also expected for semantic and phonological errors, respectively.

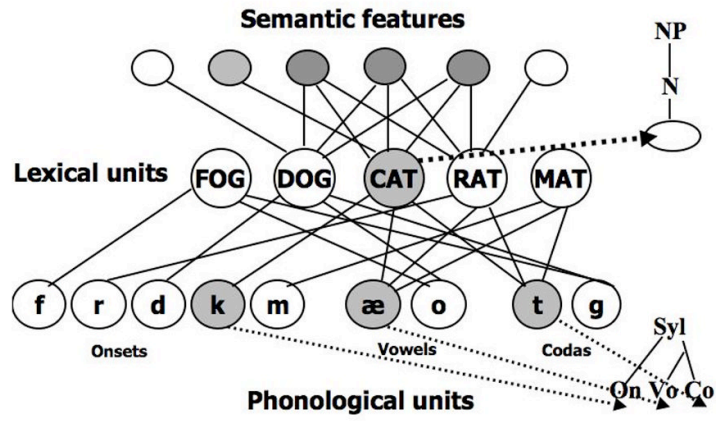


Figure 1.
The two-step interactive activation model

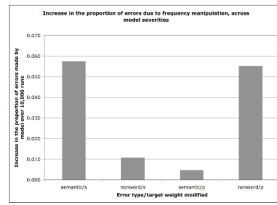


Figure 2. Increase in the proportion of semantic and nonword errors made due to a decrease in the target's *s* or *p* weights, across the range of severities

Table 1

Background information for the 50 participants

Patient code	Sex	Age at first testing	Post CVA (months) at first testing	WAB category	Non-semantic group
BAC	M	52	55	Broca	yes
BAL	F	77	15	Anomic	yes
BBC	F	78	2	Wernicke	no
BI	M	63	7	Wernicke	no
BL	M	86	21	Conduction	no
CAC	F	43	2	Anomic	yes
CE	F	59	12	Broca	no
CK	F	79	4	Anomic	yes
DD	M	54	18	Broca	no
EAC	F	47	19	Anomic	no
EAL	F	77	12	Wernicke	no
EBC	M	62	120	Anomic	yes
EC	F	47	76	Broca	yes
EE	F	55	93	Broca	yes
EL	F	38	51	Broca	yes
FAG	M	72	31	Broca	yes
FAH	M	60	8	Conduction	yes
FBH	M	69	5	Wernicke	yes
FG	F	68	8	Conduction	yes
FJ	M	76	35	Wernicke	no
FL	M	77	133	Wernicke	no
FM	F	57	14	Conduction	yes
GAL	M	70	38	Anomic	yes
HH	M	53	151	Broca	no
KAC	M	60	2	Conduction	yes
KCC	M	65	82	Anomic	no
KD	F	49	6	Conduction	yes
KE	M	65	5	Conduction	no
KI	M	48	23	Broca	no

Patient code	Sex	Age at first testing	Post CVA (months) at first testing	WAB category	Non-semantic group
KK	M	58	1	Wernicke	no
MAC	M	22	6	Anomic	no
MBC	F	66	12	Anomic	no
MD	M	48	28	Broca	yes
ME	F	74	5	Broca	no
MG	M	81	51	Anomic	no
NAC	F	74	4	Transcortical Sensory	no
ND	F	76	39	Broca	no
NH	M	59	19	Broca	yes
NI	M	53	128	Broca	no
OE	M	59	9	Anomic	yes
QB	F	37	1	Anomic	no
SAM	M	76	55	Anomic	no
SE	M	74	43	Conduction	yes
SI	M	61	39	Anomic	no
SL	M	48	16	Anomic	yes
TE	F	54	38	Broca	yes
TG	F	66	5	Anomic	yes
UL	F	77	1	Wernicke	yes
WAD	M	55	32	Broca	no
XD	M	54	86	Broca	yes

Table 2

Descriptive statistics of 124 PNT target word properties

	Mean	Range
Log Frequency	1.3760	0 – 3.2119
Log Density	0.8514	0 – 1.6021
AoA	2.0161	1 – 3
Length	4.2742	2 – 10
Imagability	591.8871	369 – 644
Name Agreement	.9676	.82 – 1.00

Table 3

Correlation matrix for 124 PNT target word properties

	Log Frequency	Log Density	AoA	Length	Imageability	Name Agreement
Log Frequency	1					
Log Density	0.38	1				
AoA	-0.36	-0.09	1			
Length	-0.35	-0.84	0.10	1		
Imageability	0.25	-0.02	-0.11	0.01	1	
Name Agreement	.07	-.09	-.04	.07	-.01	1

Table 4

Percentage of each response type (relative to the total number of analyzed responses) made by patient groups

Response Type	All patients	Non-semantic patients
Correct	59%	63%
Semantic	5%	4%
Phonological	15%	17%
Omission	12%	9%

Table 5

Percentages of Correct, Semantic, Phonological, and Omission responses (relative to the total number of analyzed responses) made by each patient on the PNT

Patient code	Correct	Semantic	Phonological	Omission
BAC	89%	0%	11%	0%
BAL	45%	2%	4%	45%
BBC	28%	1%	61%	6%
BI	8%	6%	40%	19%
BL	58%	7%	11%	6%
CAC	68%	2%	23%	5%
CE	27%	12%	20%	21%
CK	77%	4%	10%	2%
DD	40%	7%	16%	26%
EAC	54%	10%	10%	22%
EAL	17%	13%	9%	39%
EBC	84%	2%	0%	8%
EC	62%	3%	1%	31%
EE	66%	12%	2%	9%
EL	85%	2%	11%	1%
FAG	44%	6%	27%	13%
FAH	3%	4%	43%	23%
FBH	49%	3%	39%	1%
FG	75%	4%	14%	3%
FJ	54%	6%	15%	15%
FL	40%	9%	16%	15%
FM	42%	2%	36%	12%
GAL	94%	2%	2%	0%
HH	47%	13%	2%	31%
KAC	41%	1%	56%	1%
KCC	95%	1%	1%	0%
KD	78%	1%	20%	1%
KE	89%	2%	5%	0%
KI	15%	10%	21%	40%
KK	10%	2%	27%	28%
MAC	79%	6%	0%	11%
MBC	93%	2%	2%	2%
MD	91%	4%	1%	1%
ME	95%	1%	0%	0%
MG	70%	10%	4%	8%
NAC	73%	4%	6%	2%
ND	61%	7%	3%	14%
NH	88%	1%	6%	2%
NI	20%	6%	17%	43%

Patient code	Correct	Semantic	Phonological	Omission
OE	78%	7%	1%	10%
QB	94%	0%	2%	1%
SAM	82%	6%	3%	4%
SE	57%	4%	34%	1%
SI	46%	6%	2%	34%
SL	86%	6%	2%	2%
TE	48%	7%	19%	17%
TG	70%	6%	17%	1%
UL	48%	6%	18%	19%
WAD	98%	2%	0%	0%
XD	8%	4%	31%	29%

Table 6

Direction of Log Frequency and AoA effects, from analyses of all patients and Non-semantic patients. Arrows represent a significant increase or decrease in the log odds (due to increasing Log Frequency or AoA) of making a Correct response, or of a particular error (Semantic, Phonological, Omission) relative to being Correct. "N.S." denotes a non-significant effect.

	All patients		Non-semantic patients	
	Log Frequency	AoA	Log Frequency	AoA
Correct	↑	↓	↑	↓
Semantic	↓	N.S.	↓	N.S.
Phonological	↓	↑	↓	↑
Omission	↓	↑	↓	↑

Length, Log Density, Imageability, and Name Agreement regression coefficients (with standard error) from the binomial and multinomial regression analyses for all patients.

Table 7

	Log Density		Length		Imageability		Name Agreement	
	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Correct	.104	.125	-.166*	.040	-.000	.001	-.907	.845
Semantic	-.569*	.221	-.240◆	.086	-.001	.002	-.212	1.08
Phonological	.057	.173	.252*	.055	.001	.001	1.29	1.03
Omission	-.296*	.141	.159*	.048	.001	.001	1.13	1.16

* Indicates that the coefficient was significantly different from zero with p less than .05 in the expected direction, and
 ◆ in the unexpected direction.

Table 8

Regression coefficients (with standard errors) for Overall word errors not discussed in the text.

Overall word error residuals (lexical variable)	Target Log Frequency		Target AoA		Target Length		Target Imageability	
	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Log Frequency	.266*	.057	.111 ◆	.041	.035	.021	-.000	.001
AoA	-.077	.081	-.010	.055	.038	.042	-.000	.001
Length	-.024	.175	.140	.096	.091*	.039	-.001	.002
Imageability	-2.23	2.88	-3.14	1.93	4.71 ◆	1.50	.011	.032

* Indicates that the coefficient was significantly different from zero with p less than or equal to .05 in the expected direction, and
 ◆ in the unexpected direction.

Table 9

Regression coefficients (with standard error) for Phonological word and Semantic errors not discussed in the text.

	Target Log Frequency		Target AoA		Target Length		Target Imageability	
	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Phonological word Log Frequency	.109	.088	.106	.061	-.005	.042	.001	.001
Phonological word AoA	-.166	.191	-.034	.086	-.162	.085	.006	.002
Phonological word Length	.066	.418	.155	.159	.073	.123	-.000	.002
Phonological word Imageability	.083	5.02	-.991	3.19	4.47	2.78	-.019	.071
Semantic Log Frequency	.391*	.076	.155	.074	.060	.033	-.001	.001
Semantic AoA	.070	.101	.036	.068	.078	.062	-.002	.002
Semantic Length	-.152	.164	.085	.137	-.002	.082	-.000	.002
Semantic Imageability	-1.26	4.04	-5.60*	2.34	5.54	2.39	.030	.039


* Indicates that the coefficient was significantly different from zero with p less than or equal to .05 in the expected direction, and  in the unexpected direction.

Table 10

Increase in the proportion of errors made by each modified model over 10,000 runs, compared to the baseline (unmodified target weights) model for that severity

Baseline weights (severity)	<i>s</i> -manipulated model		<i>p</i> -manipulated model	
	semantic	nonword	semantic	nonword
normal (98% correct) <i>s</i> = <i>p</i> = .04	.074	.002	.006	.006
mild (93% correct) <i>s</i> = <i>p</i> = .03	.075	.010	.009	.034
moderate (65% correct) <i>s</i> = <i>p</i> = .02	.055	.005	.004	.081
severe (18% correct) <i>s</i> = <i>p</i> = .01	.009	.005	-.003	.029
mod-mild, greater <i>s</i> damage (79% correct) <i>s</i> = .02, <i>p</i> = .03	.068	.004	.007	.041
mod-mild, greater <i>p</i> damage (82% correct) <i>s</i> = .03, <i>p</i> = .02	.080	.031	.003	.092
sev-mod, greater <i>s</i> damage (60% correct) <i>s</i> = .015, <i>p</i> = .025	.052	.006	.011	.054
sev-mod, greater <i>p</i> damage (60% correct) <i>s</i> = .025, <i>p</i> = .015	.047	.024	.001	.106

Table 11a

Number of semantic errors made by each modified model over 10,000 runs, with the baseline of the s weights set to .04. Reduced s weights are .024 (60% of the baseline value).

		Model: competitor s weights		Mean number of semantic errors
		Baseline	Reduced	
Model: target s weights	Baseline	200	10	105
	Reduced	2854	593	1723.5
Mean number of semantic errors		1527	301.5	914.25

Table 11b

Number of semantic errors made by each modified model over 10,000 runs, with the baseline of the s weights set to .03. Reduced s weights are .018 (60% of the baseline value).

		Model: competitor s weights		Mean number of semantic errors
		Baseline	Reduced	
Model: target s weights	Baseline	354	80	217
	Reduced	2398	870	1634
Mean number of semantic errors		1376	475	925.5

Table 11c

Number of semantic errors made by each modified model over 10,000 runs, with the baseline of the s weights set to .02. Reduced s weights are .012 (60% of the baseline value).

		Model: competitor s weights		Mean number of semantic errors
		Baseline	Reduced	
Model: target s weights	Baseline	766	373	569.5
	Reduced	2151	1199	1675
Mean number of semantic errors		1458.5	786	1122.25

Table 11d

Number of semantic errors made by each modified model over 10,000 runs, with the baseline of the s weights set to .01. Reduced s weights are .006 (60% of the baseline value).

		Model: competitor s weights		Mean number of semantic errors
		Baseline	Reduced	
Model: target s weights	Baseline	1371	1039	1205
	Reduced	1915	1501	1708
Mean number of semantic errors		1643	1270	1456.5