

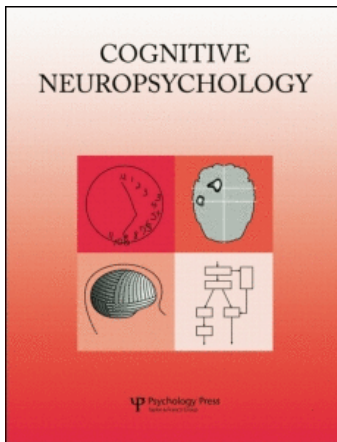
This article was downloaded by: [Macquarie University]

On: 20 August 2008

Access details: Access Details: [subscription number 788668924]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Cognitive Neuropsychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713659042>

### The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production

Mark Knobel <sup>a</sup>; Matthew Finkbeiner <sup>b</sup>; Alfonso Caramazza <sup>ac</sup>

<sup>a</sup> Cognitive Neuropsychology Laboratory, Harvard University, Cambridge, MA, USA <sup>b</sup> Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia <sup>c</sup> Center for Mind/Brain Sciences, University of Trento, Trento, Italy

First Published on: 09 September 2007

**To cite this Article** Knobel, Mark, Finkbeiner, Matthew and Caramazza, Alfonso(2007)'The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production', *Cognitive Neuropsychology*, 25:2, 256 — 286

**To link to this Article:** DOI: 10.1080/02643290701502425

**URL:** <http://dx.doi.org/10.1080/02643290701502425>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production

---

**Mark Knobel**

*Cognitive Neuropsychology Laboratory, Harvard University, Cambridge, MA, USA*

**Matthew Finkbeiner**

*Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia*

**Alfonso Caramazza**

*Cognitive Neuropsychology Laboratory, Harvard University, Cambridge, MA, USA,  
and Center for Mind/Brain Sciences, University of Trento, Trento, Italy*

The effect of lexical frequency on language-processing tasks is exceptionally reliable. For example, pictures with higher frequency names are named faster and more accurately than those with lower frequency names. Experiments with normal participants and patients strongly suggest that this production effect arises at the level of lexical access. Further work has suggested that within lexical access this effect arises at the level of lexical representations. Here we present patient E.C. who shows an effect of lexical frequency on his nonword error rate. The best explanation of his performance is that there is an additional locus of frequency at the interface of lexical and segmental representational levels. We confirm this hypothesis by showing that only computational models with frequency at this new locus can produce a similar error pattern to that of patient E.C. Finally, in an analysis of a large group of Italian patients, we show that there exist patients who replicate E.C.'s pattern of results and others who show the complementary pattern of frequency effects on semantic error rates. Our results combined with previous findings suggest that frequency plays a role throughout the process of lexical access.

Perhaps the most robust effect in all of psychology is the effect of frequency. Frequency effects have been observed in a wide range of behaviour, from

complex sequence prediction (e.g., Stadler, 1992) to face recognition (e.g., Lewis, 1999). Essentially, the more often one encounters a stimulus the

---

Correspondence should be addressed to Mark Knobel, Cognitive Neuropsychology Laboratory, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA (E-mail: knobel@wjh.harvard.edu).

We are especially grateful to E.C. for his participation in this study. We would like to thank Wheeler Ruml, Gary Dell, Matthew Goldrick, and an anonymous reviewer for their helpful comments and suggestions. We would also like to thank Jay Fournier and Janna Fucillo for help in testing E.C., Laruen Moo for interpreting E.C.'s MRI scans, and Gabriele Miceli and Rita Capasso for collecting the data on the Italian patients. In addition, we are grateful to Gary Dell for providing us the source code for his implementation of his model, and to Brenda Rapp and Matthew Goldrick for sharing a detailed specification of their modelling neighbourhood. The research reported here was supported in part by NIH Grant DC04542 to Alfonso Caramazza. Mark Knobel was supported by the Sackler Scholars Programme in Psychobiology.

more quickly and easily one is able to process it. With respect to the cognitive system, frequency effects are presumed to reflect the increase in processing fluency that comes with greater practice or learning.

Within the domain of lexical access, the more frequently a particular lexical item occurs in spoken and written language (measured by how often a word occurs in a large corpus of written text and/or speech), the more quickly and accurately that item is able to be processed. This so called "lexical frequency effect" has been observed in a wide variety of tasks, including picture naming (e.g., Gilhooly & Gilhooly, 1979; Oldfield & Wingfield, 1965), written picture naming (e.g., Bonin & Fayol, 2002), word naming (e.g., Ellis & Morrison, 1998; Forster & Chambers, 1973; Gerhand & Barry, 1999a), and lexical decision (e.g., Brysbaert, Lange, & Van Wijndaele, 2000; Gerhand & Barry, 1999b; Morrison & Ellis, 1995). In this article, we focus on the lexical frequency effect in speech production tasks and, specifically, in picture naming.

Many of the above studies and others have also looked at the contribution of age of acquisition (AoA), a variable correlated with frequency, to performance on various lexical tasks. It remains an open question, particularly in picture naming, whether the lexical frequency effect is really an AoA effect, or whether both factors contribute to performance on these tasks (see Johnston & Barry, 2006, for a recent review). We do not wish to prejudge the issue and use the term frequency only to reflect the fact that we have used frequency norms in our study.

Evidence for frequency effects in picture naming has come from the contrasting response latencies observed in picture naming and picture recognition tasks. While picture naming regularly produces frequency effects, no effects of lexical frequency are found when participants simply need to recognize the pictures (Jescheniak & Levelt, 1994; Wingfield, 1967, 1968; but see Bartram, 1976; Kroll & Potter, 1984). These results have been interpreted as showing that frequency effects arise in word production processes subsequent to picture recognition processes.

The effects of word frequency are perhaps most pronounced in the performance of neuropsychological patients, including Alzheimer's disease patients (e.g., Sailor, Antoine, Diaz, Kuslansky, & Kluger, 2004; Silveri, Cappa, Mariotti, & Puopolo, 2002; Thompson-Schill, Gabrieli, & Fleischman, 1999), semantic dementia patients (e.g., Bird, Lambon Ralph, Patterson, & Hodges, 2000), and aphasic patients. With respect to this latter group, it has been widely demonstrated that patients name pictures with high-frequency names more accurately than they name pictures with low-frequency names (e.g., Cuetos, Aguado, Izura, & Ellis, 2002; Feyereisen, Van der Borgh, & Seron, 1988; Gordon, 2002; Howard, Patterson, Franklin, Morton, & Orchard-Lisle, 1984; Nickels & Howard, 1994, 1995; Schwartz, Wilshire, Gagnon, & Polansky, 2004; Wilshire, 2002).

Given the ubiquity of the frequency effect in lexical access, it is not surprising that models of lexical access have been developed to explain first and foremost the frequency effect. For example, the search model (Forster, 1976), which assumes localist representations, stipulates that lexical representations are positioned within the search space on the basis of frequency, with high-frequency (HF) words at the beginning of the search and low-frequency (LF) words at the end. In contrast, activation models that assume localist representations implement frequency either by varying selection thresholds, with HF nodes having a lower selection threshold than LF nodes, or by increasing resting activations for HF nodes relative to LF nodes. Distributed connectionist (henceforth connectionist) models, on the other hand, which do not posit localist lexical representations, implement frequency effects through the strengthening of connections between semantic, phonological, and intermediary (i.e., hidden) nodes. As an emergent property of their learning rules, connectionist models come to represent lexical frequency through training (Monsell, 1991). Although the search model and connectionist models have enjoyed some success in the word recognition literature, most models of speech production, while based on

the activation framework like connectionist models, assume localist lexical representations and are not trained. Since frequency is not the guiding principle in this modelling framework, as it is in the search model, for example, there are more options for how frequency can be implemented in models of speech production. Before we discuss implementing lexical frequency in these models, we briefly introduce their basic architecture.

The current models of lexical access in speech production (e.g., Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999) agree on three stages in word production: semantic, lexical, and segmental. The semantic stage encompasses selecting the concepts that one wishes to express verbally. The lexical stage includes selecting the lexical nodes that correspond to the selected concepts. Finally, the segmental stage involves selecting the phonological segments (i.e., phonemes) specified by the lexical node. The output from this final stage is eventually translated into the movements of the articulators. This depiction is, of course, a simplification that ignores the many differences between these models, including differences in processing dynamics and in the number of lexical stages. We return to these issues and the specifics of each model in our General Discussion.

Word frequency effects can be implemented in this simple three-stage model at five possible loci. These are: (a) semantics, (b) the interface between semantics and the lexical level, (c) the lexical level, (d) the interface between the lexical and segmental levels, and (e) the segmental level (see Figure 1). Placing lexical frequency effects at the level of semantics (Option 1), though logically possible, is unlikely since words are not represented at this level. If lexical frequency effects are found to originate from the level of semantics, it would most likely indicate that a semantic variable that is correlated with lexical frequency, such as familiarity, concreteness, or imageability, underlies these effects (e.g., Bates et al., 2003). Likewise, placing lexical frequency at the segmental level (Option 5) also seems unlikely, given that only sublexical phonological

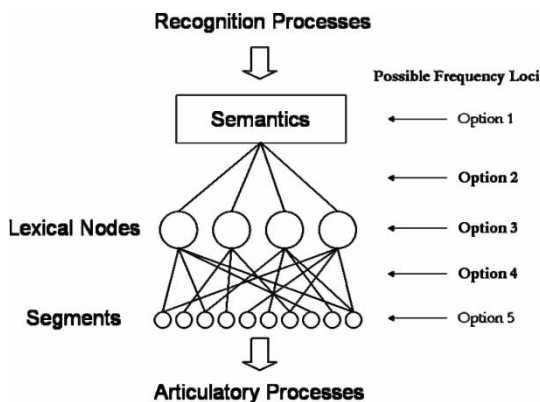


Figure 1. A basic model of word production. The five logically possible loci for the frequency effect are shown, with the three likely loci highlighted.

information is represented at this level. Here again, factors correlated with lexical frequency, such as number of phonemes or syllables, phoneme or syllable frequency, or transitional probabilities between phonemes, would be the more likely cause, not lexical frequency (e.g., Vitevitch, 2002). By process of elimination, it seems reasonable to suggest that lexical frequency effects arise either as a function of the connections that interface with the lexical level (Options 2 and 4) or as a function of the lexical representations themselves (Option 3).

We point out that these three possible loci of the lexical frequency effect do not necessarily constitute mutually exclusive hypotheses. It is not only logically possible that frequency may affect processing at several stages of lexical access, but we suggest that this should be taken as the default assumption. Since lexical frequency effects in speech production are said to reflect the frequency with which different lexical items are retrieved from memory and encoded phonologically for production, it would be surprising if the advantage for higher frequency words were not reflected in all stages of lexical access.

Previous work on isolating the locus of the frequency effect in speech production has led to the conclusion that frequency effects arise at the lexical level. For example, Dell (1990) found that

neurologically intact participants produced fewer errors on high-frequency word pairs (e.g., “vote pass”) than on low-frequency word pairs (e.g., “vogue pang”) in a phonological error elicitation task. Dell successfully simulated these effects by representing lexical frequency in the resting activation levels of lexical nodes in a localist spreading-activation network. Though modelling frequency in this way led to results that were consistent with the results of the behavioural experiments, implementing frequency at other loci may have led to similarly successful results.

In another line of work with normal individuals, Jescheniak and Levelt (1994) determined that lexical frequency effects arise at the lexical level by ruling out the semantic and postlexical levels as possible loci. They reasoned that if frequency occurred at the level of lexical selection, then the time difference between accessing HF and LF words would disappear with a delay that was long enough to allow for the access of even the slowest LF words. On the other hand, if frequency effects arose at postlexical levels in production, then frequency effects should still be present after a delay. Jescheniak and Levelt had participants name pictures and found that the frequency effects obtained in picture naming disappeared in a delayed cued-naming paradigm. Thus, Jescheniak and Levelt interpreted the disappearance of frequency effects after a delay as indicating that the frequency effects were not coming from postlexical processes.

Furthermore, Jescheniak and Levelt (1994) also found that the frequency effects observed in picture naming disappeared in a word–picture matching task. Since participants were still required to recognize the picture in order to reply if it matched the previously presented word, Jescheniak and Levelt argued that frequency effects could not be coming from recognition processes. Having eliminated both pre- and postlexical loci for frequency effects, Jescheniak and

Levelt concluded that the frequency effects observed in the no-delay condition arose from the lexical level of word production. Their modelling proposal was quite similar to that of Dell (1990).<sup>1</sup> They suggested that these effects could be modelled as a difference in resting activation levels between HF and LF words, with HF words having a higher resting activation. This increase in resting activation would allow HF words to be available for lexical selection (and, hence, produced) more quickly than LF words in localist spreading-activation type models (see also Levelt et al., 1999).

A particularly clear demonstration of the lexical locus of the frequency effect comes from the study by Finocchiaro and Caramazza (2006). They found that production latencies of pronominal clitics (i.e., pronouns) were sensitive to the lexical frequency of the replaced noun. They ruled out the possibility that participants were covertly producing the noun’s phonology by showing that phonologically related distractor words did not produce facilitation, an effect that is typically observed when participants are asked to name the pictures (e.g., Damian & Martin, 1999; Meyer & Schriefers, 1991). Combined, these two findings suggest that the lexical frequency effect that they observed arose at a level of processing that preceded access to the segmental information. Finocchiaro and Caramazza (2006) concluded that the lexical frequency effect arose at the point of lexical selection.

While the studies that we have reviewed here have concluded that lexical selection is the locus of the frequency effect, frequency effects, as we pointed out above, could also arise in the interface between semantics and the lexicon and/or in the interface between the lexicon and phonology (for this latter possibility see Barry, Morrison, & Ellis, 1997; Brown & Watson, 1987; Caramazza & Miozzo, 1997; MacKay, 1987). In the present study, we pursue these possibilities by

<sup>1</sup> At a finer grain, Dell (1990) and Jescheniak and Levelt (1994) came to different conclusions about the locus of frequency effects within word production. Dell put the frequency effect at the lemma, or syntactic form, level, while Jescheniak and Levelt put it at the lexeme, or phonological form, level. These distinctions within the lexical level are not universally accepted (see Caramazza, 1997; Caramazza & Miozzo, 1997). We discuss the issue of multiple lexical levels in our General Discussion.

investigating the distributions of specific error types as a function of target lexical frequency in an aphasic patient. For example, we ask whether or not semantic errors (e.g., patient responds “tiger” when presented with a picture of a giraffe) or phonological nonword errors (e.g., “piaffe” as an approximation of “giraffe”) occur more often when the to-be-named picture has a low-frequency name than when it has a high-frequency name. On the assumption that the origin of specific error types (e.g., semantic, phonological) can be isolated in the speech production system, this strategy allows for the opportunity to determine which levels of processing are sensitive to lexical frequency. Below we discuss the rationale used to determine the source of specific error types within the speech production system.

Semantic errors can have either a lexical/phonological source or a semantic source. In the case of a semantic source, semantic errors may be due to a damaged semantic system not sending enough activation down to the lexical level to guarantee the selection of the target lexical representation. Alternatively, damage at the semantic level may be so severe that the wrong concept is selected, and lexical access continues accurately, just for the wrong target. In the case of a lexical or phonological source, semantic errors may arise as a result of an impairment in selecting the correct lexical representation or in selecting its corresponding phonology. For example, if the lexical selection mechanism is damaged, it may be difficult to select the target word from its semantically related neighbours that are also highly activated (e.g., Caramazza & Hillis, 1990; Nickels, 1995). Alternatively, if the target’s phonology is unavailable, the patient may reselect a related (semantically or possibly phonologically) lexical item and proceed with producing this new item to show that they understand the task. This latter possibility differs from the other explanations in that it is a conscious strategy on the part of the patient and requires an additional selection step.

Fortunately, it is possible to distinguish patients who make semantic errors as a result of semantic level damage from those with lexical/phonological level damage. Individuals with semantic level

damage typically have deficits in a wide variety of tasks, including those that require the patient to retrieve factual information about the target item (e.g., “Does a bird have four legs?”). In the case of patients who make semantic errors as a result of lexical/phonological (but not semantic) level damage, these patients perform normally in semantic judgement or sentence comprehension tasks.

Phonological errors (both word and nonword) arise during lexical access, or at a later postlexical stage (Kay & Ellis, 1987; Nickels, 1995). The explanation for the former case is that there is not enough activation coming from the lexical level to select all the constituent phonemes correctly, leading to phonological errors. Additionally, phonological word errors may arise by misselection of a word phonologically related to the target (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Gagnon, Schwartz, Martin, Dell, & Saffran, 1997). In the postlexical case, there are problems with peripheral output mechanisms, such as articulatory control.

Once again, it is generally possible to distinguish patients who make phonological errors as a result of lexical level damage from those who make errors as a result of postlexical level damage. Patients with postlexical level damage have similar problems across all production tasks, including naming, reading, and repetition. In contrast, patients with lexical damage should perform worse in naming than in reading and repetition, since these latter tasks can be accomplished using nonlexical processes (Goldrick & Rapp, 2007).

Given that different error types occur as a result of damage to different stages of processing, we can look at the effect of lexical frequency on each error type to determine the range of processing stages that are sensitive to lexical frequency. If lexical frequency, and not a related variable, affects the likelihood of observing a particular error type, then we have support for the notion that lexical frequency is represented at the level of the system from which those errors arise. If we find that semantic errors are affected by frequency and, importantly, that these errors are not due to semantic level damage, then it is reasonable to consider that

lexical frequency is represented either in the interface between semantics and the lexicon or at the point of lexical selection. If we find that lexical frequency affects the likelihood of observing phonological approximations of target names, then we have evidence for lexical frequency being represented in the interface between the lexicon and segmental representations.

Several studies have looked at frequency effects on semantic errors (Caramazza & Hillis, 1990; Nickels & Howard, 1994), phonological errors (Gagnon et al., 1997; Goldrick & Rapp, 2007; Schwartz et al., 2004; Wilshire & Fisher, 2004), or both (Cuetos et al., 2002; Feyereisen et al., 1988; Nickels, 1995). For example, two group studies report frequency effects on semantic errors in picture naming. In the first study, an investigation of operativity (i.e., how manipulable an object is) on aphasic errors, Feyereisen et al. (1988) report an effect of frequency on semantic errors, such that there were fewer semantic errors on higher frequency targets. In the second study, Cuetos et al. (2002) also found a similar pattern, whereby semantic errors decreased as target AoA decreased. Taken together, these findings suggest that frequency may affect semantic error rates; this is especially true for the Cuetos et al. study, which took into account other correlated factors such as imageability and word length. Unfortunately, these studies do not report the nature of the patients' deficits, and so it is unclear whether the lexical frequency effect should be attributed to prelexical, lexical, or post-lexical levels of processing.

The few studies that have investigated the level of damage seem to agree that lexical frequency affects the rate of semantic errors, but only when errors arise as a result of lexical-level damage. For example, Caramazza and Hillis (1990) reported two patients who produced mostly semantic errors but who had intact semantics. Interestingly, though both patients showed frequency effects on their overall reading performance, only one, R.G.B., showed a frequency effect on his rate of semantic errors in picture naming and reading (more errors on categories with lower frequency words than on those with

higher frequency words). Though this pattern of performance has rarely been reported in the literature (see also Hirsh & Ellis, 1994; Zingeser & Berndt, 1988, for potentially similar cases), the case of R.G.B. demonstrates that lexical frequency can have an effect on semantic error rates. In contrast to patient R.G.B., Nickels and Howard (1994) looked at familiarity effects on semantic errors for 8 patients, whom they determined had semantic damage, and found no effect of familiarity (as a proxy for frequency) on their semantic errors (see also Hillis, Rapp, Romani, & Caramazza, 1990).

Studies that have looked at phonological errors have also found that the likelihood of observing these types of error is sometimes sensitive to lexical frequency. In their group of patients, Cuetos et al. (2002) found that phonological errors decreased with target AoA overall. Feyereisen et al. (1988) also found that phonological errors decreased as a function of lexical frequency. Schwartz et al. (2004) also looked at lexical frequency effects on phonological errors in their sample of 18 patients. For their analysis, they split the phonological errors (based on a very liberal criterion) into proximate errors (those that preserved >50% of the target's phonemes) and remote errors (those with <50% of the target's phonemes). They found that log frequency only had a significant effect on proximate errors, not on remote errors. This frequency effect was in the intuitive direction—higher frequency targets led to fewer proximate phonological errors. Again, the findings reported in all three studies suggest that target lexical frequency may affect the likelihood of observing a phonological error. These findings could be taken as initial support for the hypothesis that frequency effects arise at the interface between lexical and segmental levels, but there are two difficulties that prevent making a strong conclusion about the level at which this effect arises: (a) heterogeneous groups of patients were used without specification of the individual deficits, and (b) word and nonword errors were collapsed into one category.

The study by Gagnon et al. (1997) overcomes our second objection by looking only at the

phonological word errors of a group of 9 patients. They found that the frequency of such errors was higher than a conservative chance estimate of such a frequency, suggesting that frequency plays a role in such errors. Based on this and several other findings, they argued that at least some phonological word errors arise at the level of lexical selection. However, by looking only at phonological word errors, they cannot strongly constrain where to place frequency in the lexical system, especially if one believes their arguments for feedback of activation between levels of the lexical system.

Nickels (1995) also found effects of familiarity (as a proxy of frequency) on phonological error rates in her patients. A total of 3 out of the 15 patients showed familiarity effects individually. For all three patients a reduction in phonological errors led to a reduction in overall error. These were the same three patients who showed an overall familiarity effect upon naming; however, none of these patients or any of the others showed a significant effect of familiarity on semantic errors, though this is not surprising given that none of the other patients exhibited a familiarity effect on their *correct* responses. This study goes beyond the previous studies in showing that familiarity (as a proxy of frequency) affects phonological error rates in individual patients.

Similarly, in a study of the effect of phonological neighbourhood size on aphasic speech production, Gordon (2002) reports an effect of lexical frequency on patients' phonological error rates in picture naming and picture description tasks. The frequency effect held for 27 out of 32 individual patients in naming and 26 out of 34 in picture description. Overall, patients found it easier to name words with higher frequency and more dense neighbourhoods.

Mirroring the findings for semantic errors, a recent study by Goldrick and Rapp (2007) found that frequency affects phonological errors (word and nonword) only when such errors arise from a lexical deficit. They contrasted two patients, one with a lexical phonological deficit, and the other with a postlexical phonological deficit. Though both made similar kinds of errors, only the one

with a lexical deficit showed any effect of frequency on response accuracy.

Taken together, these studies show that target lexical frequency affects the likelihood of patients making semantic and phonological errors, so long as their deficit is lexical in nature. Though these studies have revealed a robust effect of frequency on individual error types, they have not revealed the locus of this frequency effect. To do so, it is important to include an in-depth analysis of a patient's deficits and a systematic analysis of all error types as a function of target frequency. One purpose of the present article is to provide such an analysis.

The rest of this article is organized into three sections. In the first section, we present E.C., an aphasic patient who produces many different types of errors and whose overall picture-naming performance is affected by lexical frequency. These features make him an ideal candidate for localizing the effect of lexical frequency. To anticipate our results, we found that target lexical frequency did not modulate the likelihood of E.C. making a semantic error but did modulate his phonological nonword error rate. We take these findings to suggest that one locus of the lexical frequency effect is in the connections between the lexicon and segmental phonology.

In the second section of the paper we test the lexical-segmental hypothesis of lexical frequency by implementing frequency into a range of computational word production models. We directly compare models in which lexical frequency is implemented in the lexical-segmental connections with models in which frequency is implemented in the semantic-lexical connections and in the lexical nodes themselves. To anticipate our findings once again, placing frequency in the lexical-segmental connections led to the best fit of E.C.'s pattern among the three alternative hypotheses. Thus, the modelling work provides support for the suggestion that lexical frequency may be represented in the lexical-segmental connections.

In the third section of the article, we reanalyse the picture-naming data of a group of Italian aphasic patients both to look for a confirmation of E.C.'s pattern of performance and to see

whether this technique of analysing specific error types as a function of lexical frequency might not reveal complementary patterns of performance. In this section of the paper, we report both patients like E.C., who show an effect of lexical frequency on phonological nonword errors, and those who, unlike E.C., exhibit an effect of lexical frequency on semantic errors. Taken together, these patterns of performance challenge the idea of a single locus of the lexical frequency effect in speech production. This work has strong implications for how we should think about lexical frequency and for how lexical frequency should be implemented in computational models of speech production.

## PATIENT E.C.

### *Case history*

E.C. is a 52-year-old, right-handed man with a bachelor's degree. He worked with computers before having a left-hemisphere stroke that resulted in language and attention difficulties. A magnetic resonance imaging (MRI) scan revealed a large perisylvian infarction, which included the superior temporal gyrus, extending into the supra-marginal gyrus, the frontal operculum and inferior frontal gyrus, and part of the insula. E.C.'s language deficit is best classified as conduction aphasia. We started testing the patient 5 years and 8 months after the cerebral vascular accident (CVA). E.C. continues to use his dominant right hand after his stroke.

### *Nonlanguage testing*

E.C. showed no evidence of bucco-facial apraxia. He showed a reduced digit span of 4 forwards and 2 backwards. His performance on Raven's progressive matrices was 52/60 (86.7%), which places him in the 98th percentile for his age. E.C. also correctly copied several pictures that we presented him, including a clock, a vase with flower, and the "Ogden scene" (a house with a fence and two trees). He could easily discern the difference between real and chimeric pictures (e.g., a picture of half of a penguin joined to half of a

traffic light), scoring 20/20 (100%), though he had difficulties naming some of the objects involved.

### *Language evaluation*

E.C. made frequent hesitations in his spontaneous speech, had word-finding difficulties, and made several semantic errors and phonological nonword errors—phonological approximations of the target. On single-word production tasks, he performed poorly across the board. His oral reading was impaired for words (55/75, 73%) and nonwords (7/10, 70%), showing problems with both verbs and nouns. His written spelling was also impaired, when elicited both by a picture (5/10, 50%) and from dictation (3/10, 30%); his oral spelling was similarly impaired (1/5, 20%). Across all spelling tasks, his responses would often share letters with the target, but they were rarely phonologically plausible. E.C. was also markedly impaired in simple repetition tasks for words (21/35, 60%) and nonwords (2/5, 40%). There was a clear length effect on his error rate—logistic regression,  $\chi^2(1) = 6.43$ ,  $p = .011$ , Nagelkerke  $R^2 = .277$ —for word repetition, and errors tended to be phonologically or morphologically related to the target. Picture naming was the most impaired (3/20, 15%), with errors being mostly semantic and phonological in nature.

In contrast to his single-word production, E.C. did quite well in auditory word–picture matching (49/50, 98%), lexical discrimination (40/40, 100%), lexical decision (18/20, 90%), and auditory–visual word matching (19/20, 95%). Together these findings suggest that audition and comprehension for single words was relatively intact.

In sentence processing, E.C. showed a similar disparity between his comprehension and production. His auditory sentence to picture matching was mildly impaired (31/36 correct, 86%). E.C. performed worse on passive (12/16 correct, 75%) than on active sentences (19/20 correct, 95%). He did not make a single mistake (12/12 correct, 100%) in comprehending plural morphology in sentences (e.g., "The horse was

followed by the dogs"). E.C.'s grammaticality judgements were also fairly good (9/10 correct, 90%). His only mistake was in accepting "the boy was followed the girl" as grammatical.

On the production side, E.C. seemed capable of producing short adjective-noun phrases (8/10 correct, 80%). However, his production in unconstrained (1/3 correct, 33%) and constrained sentences (0/3 correct, 0%) was quite low, which matched his performance for picture descriptions. E.C.'s sentence completion was also quite impaired (11/21, 52%; chance = 33%). Similarly, he also had trouble with sentence repetition (2/5 correct, 40%). His difficulty was proportional to the complexity of the sentences.

Further investigation of E.C.'s reading suggested damage to both his sublexical and lexical routes. Evidence of damage to his sublexical route was seen in his impaired nonword reading (34/80; 43%), while evidence for lexical-route damage came from significant effects of frequency and imageability on his word-reading performance (155/240; 65%).<sup>2</sup> He was better at reading more frequent and more imageable words. The imageability effect suggested problems with semantic processing, while the frequency effect provided evidence of lexical damage. The same list of words was also read to E.C. for a repetition task. His performance was slightly better than the reading (173/240; 72%). Word length was the only factor that affected his repetition performance. E.C. was better at reading and repeating shorter words. This effect indicated damage to accessing the segments and/or damage to postlexical processes. The fact that E.C. made similar errors (mostly substitution, addition, deletion, and transposition errors<sup>3</sup>) in reading words and nonwords and in repeating words suggested that there is a common locus of damage where the

lexical and sublexical streams meet, in segmental processing, or in later postlexical stages.

To investigate further the extent of E.C.'s semantic deficit, we used a revised version of the original central attributes test (Caramazza & Shelton, 1998); this revised version of the test consisted of 305 "true/false" statements probing particular attributes of common living and nonliving objects, such as "A shirt is made of cloth" and "A squirrel eats nuts". E.C.'s performance on this task was better than chance (50%), but impaired (219/305 correct, 72%). There was no difference between his knowledge of visual and nonvisual attributes. His attribute knowledge was best for fruits and vegetables (21/24 correct, 88%), followed by animate living things (115/154 correct, 71%), with inanimate objects being worst (83/127 correct, 65%). The fact that that E.C. had trouble with this task indicated that his semantic system is also damaged.

### *Experimental investigation*

We were particularly interested in E.C.'s picture-naming ability because E.C.'s functional deficit appears to span all three levels of the speech production system: semantic, lexical, and postlexical. We reasoned that this particular case afforded a unique opportunity to assess the effects of frequency on a variety of different error types within a single patient.

Over the course of a year E.C. was asked to name all 260 items from the Snodgrass and Vanderwart (1980) picture set twice. E.C. also named all 175 items of the Philadelphia Naming Test (Roach, Schwartz, Martin, Grewal, & Brecher, 1996) once. The stimuli were printed out on separate sheets of paper and were presented one at a time to E.C., who had unlimited time to name them.

<sup>2</sup> A logistic regression was conducted with log frequency, length in letters, imageability, and noun-verb status. The overall model was significant,  $\chi^2(4) = 37.55$ ,  $p < .001$ , Nagelkerke  $R^2 = .199$ . Log frequency (Wald = 7.27,  $p = .007$ ), length in letters (Wald = 13.84,  $p < .001$ ), and imageability (Wald = 6.85,  $p = .009$ ) were the significant individual factors. His repetition data were analysed using a logistic regression with the same four factors. This time the overall model was also significant,  $\chi^2(4) = 9.54$ ,  $p = .049$ , Nagelkerke  $R^2 = .056$ . The only significant factor was length in letters (Wald = 6.28,  $p = .012$ ).

<sup>3</sup> Regularization errors (e.g., [herθ] for "hearth") and morphological errors were also common in word reading, while lexicalization errors were common in nonword reading.

The results from both tests were scored according to the criteria in Dell et al. (1997). We scored the first whole response. The major error categories were semantic, phonological (or formal), mixed, unrelated, or nonword and were classified according to the following rules:

*Correct.* The response must contain all the phonemes of the target. The only exceptions were for plural/singular morphology and for synonyms.

*Phonological (formal).* Unlike the general category of phonological errors that we discussed in the Introduction, this category here (and in all further usage) refers to only phonological *word* errors; phonological nonword errors are included in the nonword category. The criteria were designed to be fairly broad (Dell et al., 1997). In order to qualify, an error had to share a single phoneme with the target at the beginning, end, or any other position when both were lined up left to right. Errors that also shared two phonemes (excluding unstressed vowels and plural morphemes) with the target in any position were also counted as phonological errors.

*Mixed.* An error that meets the criteria for both semantic and phonological errors (see above).

*Unrelated.* All single-word errors that are not related to the target (see *Other* errors).

*Nonword.* Any response that is not in the English lexicon. Both phonologically related and unrelated responses were scored as nonwords, including those with a potential semantic relationship to the target.

*Other.* This is the category of errors that are not explained by the original model and includes: no responses, definitions and circumlocutions, parts of a picture, single morphemes, and visually related responses with no clear semantic relationship.

### *Analyses*

We were most interested in the possible effect of lexical frequency on each of E.C.'s error types. To investigate this possibility we used several binary logistic regressions on the same dataset, one for correct responses and one each of the following error types: semantic, phonological (formal), mixed, and nonword. "Other" errors were not analysed since this was a heterogeneous category and since errors of this type have not been considered in previous modelling work (Dell et al., 1997). We also did not analyse unrelated errors since E.C. produced so few of them.

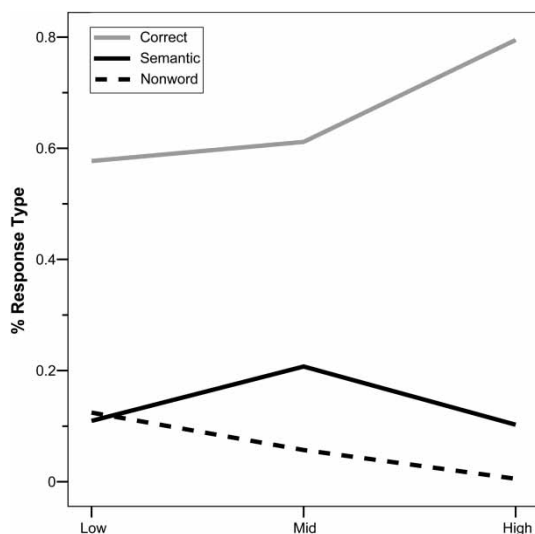
Each binary logistic regression analysis included log frequency as an independent variable:  $\log_{10}(\text{frequency} + 1)$ , where frequency is the CELEX lemma frequency; Baayen, Piepenbrock, & van Rijn, 1993. Independent variables also included imageability (MRC Psycholinguistic Database; Wilson, 1988), number of phonemes and number of phonological neighbours. Phonological neighbours were calculated using those words that appeared in both the Carnegie Mellon Pronouncing Dictionary (Version 0.6; Weide, 1994) and the CELEX lemma frequency dictionary (for a total of 25,815 words). Neighbours were considered to be all words in the intersection dictionary that were a single phoneme addition, deletion, or substitution away from the target word (following Vitevitch, 2002). We ran a second set of analyses without imageability to confirm the other factors since we could not obtain imageability ratings for all our items. Compound, plural, and multiple word targets were not included in the dataset.

### **Results**

Out of the 589 items that we analysed, E.C. made 200 errors (66.0% correct). The breakdown of the errors by types is as follows: 82 semantic (13.9%), 14 phonological (2.4%), 15 mixed (2.5%), 3 unrelated (0.5%), 37 nonword (6.3%), and 49 other (8.3%) errors. Most of E.C.'s nonword errors were clear phonological approximations of the target (e.g., "strew" for the target picture of a screw and "toa, toach, tosh . . ." for the target

picture of a toaster). When looking at the error distribution by frequency, we found a clear pattern for both semantic and nonword errors, the two main categories of E.C. errors. Semantic errors were not affected by frequency, while nonword errors decreased as target frequency increased (see Figure 2). Since we conducted five analyses on the same dataset (the correct response analysis and the analyses of semantic, phonological, mixed, and nonword errors), we adjusted our critical  $\alpha$  level to .01 using Bonferroni correction. We report significant,  $\alpha \leq .01$ , and marginal values,  $.01 < \alpha \leq .05$ . Confirmatory analyses of the other factors using the larger dataset without imageability ratings were held to the same critical alpha level.

For the overall analysis of errors, there were 589 cases, of which 200 were errors. The overall model was significant,  $\chi^2(3) = 31.68$ ,  $p < .001$ , Nagelkerke  $R^2 = .072$ , and the factors of log frequency (Wald = 10.79,  $p < .001$ ) and number of phonemes (Wald = 7.11,  $p = .008$ ) were significant. When the analysis was repeated with imageability, 481 cases were included, of which



**Figure 2.** A breakdown of E.C.'s major response types by target frequency. Only correct response and semantic and nonword errors are shown. Low-frequency items ( $n = 201$ ) had a frequency  $\leq 12$ . Mid-frequency items ( $N = 193$ ) had a frequency  $> 12$  and  $\leq 40$ . High-frequency items ( $N = 195$ ) had a frequency  $> 40$ .

157 were errors. The overall model was again significant,  $\chi^2(4) = 56.7$ ,  $p < .001$ , with a Nagelkerke  $R^2$  of .155. The only significant factor was imageability (Wald = 31.67,  $p < .001$ ), while log frequency (Wald = 5.24,  $p = .022$ ) and number of phonemes (Wald = 5.43,  $p = .020$ ) were marginal. Errors tended to occur on less frequent, longer, and less imageable items.

In the analysis of semantic errors relative to the other error types (including correct responses), the overall model was not significant,  $\chi^2(4) = 8.2$ ,  $p = .085$ , Nagelkerke  $R^2 = .030$ . However, the individual factor imageability was significant (Wald = 7.18,  $p = .007$ ). Semantic errors were potentially more likely for low-imageability words. When the analysis was repeated without imageability, the model was again not significant,  $\chi^2(3) = 3.13$ ,  $ns$ , Nagelkerke  $R^2 = .010$ . None of the factors frequency, length, or number of phonological neighbours was significant in this analysis (all  $ps > .1$ ).

The analysis of nonword errors was significant overall,  $\chi^2(4) = 44.24$ ,  $p < .001$ , Nagelkerke  $R^2 = .250$ , with log frequency being significant (Wald = 11.64,  $p = .001$ ) and length in phonemes being marginal (Wald = 9.66,  $p = .02$ ). Both these factors were significant in the analysis without imageability (Wald = 12.15,  $p < .001$ , for frequency; Wald = 14.06,  $p < .001$ , for phonemes), as was the overall analysis,  $\chi^2(3) = 46.73$ ,  $p < .001$ , Nagelkerke  $R^2 = .204$ . Longer and lower frequency targets were more likely to lead to nonword errors.

Though E.C. made few phonological and mixed errors, the analyses of these errors were both significant: phonological,  $\chi^2(4) = 16.14$ ,  $p = .003$ , Nagelkerke  $R^2 = .234$ ; mixed,  $\chi^2(4) = 16.32$ ,  $p = .003$ , Nagelkerke  $R^2 = .170$ . Imageability (Wald = 14.02,  $p < .001$ ) was the only significant factor, while number of phonological neighbours (Wald = 3.86,  $p = .049$ ) was marginal. Phonological errors were more likely to occur on targets that were lower in imageability and had more phonological neighbours. The analysis without imageability was marginal,  $\chi^2(3) = 9.77$ ,  $p = .021$ , Nagelkerke  $R^2 = .082$ , and showed number of phonological neighbours also to be marginal (Wald = 4.65,  $p = .031$ ). In this analysis,

number of phonemes turned out to be significant (Wald = 11.35,  $p = .001$ ). In the case of mixed errors, lower frequency targets (Wald = 8.48,  $p = .004$ ) and possibly targets in more dense phonological neighbourhoods (Wald = 5.59,  $p = .018$ ) led to more such errors. When the analyses of mixed errors were repeated without imageability, frequency was again a significant factor (Wald = 6.96,  $p = .008$ ), and phonological neighbourhood size was marginal (Wald = 5.35,  $p = .023$ ). The overall model was marginal,  $\chi^2(3) = 10.70$ ,  $p = .013$ , Nagelkerke  $R^2 = .085$ .

## Discussion

The results of the overall naming performance support the idea that E.C. has several impairments spanning the production system. The fact that he produced fewer errors when naming pictures with names that have been rated as being more imageable (just as he did in the reading task) and that this imageability effect was mostly carried by his semantic errors supports the idea that he has damage to the semantic system. The effect of frequency supports the idea that his difficulty in production tasks is also partly due to damage to the lexical system. His similar performance on naming, reading, and repetition suggests that he may additionally have damage to his post-lexical system. This is further supported by his difficulties in reading words and nonwords and with the similar kinds of errors that he produced in all production tasks.

The particularly intriguing aspect of E.C.'s performance comes from the analysis of his specific error types and the effect of lexical frequency on the distributions of those errors. Though there were fewer errors overall as a function of frequency, the effect of lexical frequency on the individual error types varied greatly. The striking finding is that phonological nonword errors were strongly affected by lexical frequency, while semantic errors

were not.<sup>4</sup> Crucially, the effect of frequency on E.C.'s performance held even after the effects of imageability, length, and number of phonological neighbours were factored out. Thus, the effect of lexical frequency on E.C.'s phonological nonword errors cannot be reduced to a combination of variables that are correlated with lexical frequency. Of the three likely implementations of lexical frequency in the basic model (see Figure 1), only a locus at the lexical-segmental interface can readily account for how lexical frequency could affect the likelihood of observing a phonological nonword error.

In the next section, we use a computational modelling approach to test the hypothesis that E.C.'s picture naming and, specifically, the interaction between lexical frequency and phonological nonword error rates, are best accounted for by a model in which lexical frequency is implemented in the lexical-segmental connections. This work reveals that only by implementing lexical frequency in the lexical-segmental connections is one able to successfully model E.C.'s pattern of performance.

## FREQUENCY MODELLING

Several studies have attempted to simulate patterns of errors produced by aphasic patients in a simple picture-naming task (Dell et al., 1997; Foygel & Dell, 2000; Ruml & Caramazza, 2000; Ruml, Caramazza, Capasso, & Miceli, 2005; Ruml, Caramazza, Shelton, & Chialant, 2000; Schwartz, Dell, Martin, Gahl, & Sobel, 2006). The architecture and dynamics of the underlying (normal) model of production in all these studies have been similar. In modelling the frequency pattern exhibited by patient E.C., we started with the same basic architecture (Dell et al., 1997) and varied several core assumptions. We then implemented lexical frequency in three different loci: (a) the interface between semantic and

<sup>4</sup> Frequency was also found to affect mixed errors, but not phonological (word) errors. Since there were few errors in each of these categories (2.4% for phonological; 2.5% for mixed) we do not discuss these findings further.

lexical levels, (b) the lexical level, and (c) the interface between lexical and segmental levels. Our goal was not only to find the best fit to E.C.'s pattern, but also to see how robust our findings would be across several core assumptions. We describe the basic model, our variations, and our extensions of the model to incorporate frequency in turn.

### *Architecture and dynamics*

The architecture of the general lexical model (see Introduction) consists of three levels of representations: semantic, lexical, and segmental. Semantic nodes represent conceptual features, lexical nodes represent abstract word forms, and segmental nodes represent abstract phonemes. The following equations govern the spread of activation for a node  $m$  at any level of the system for the next time step ( $t + 1$ ):

$$\begin{aligned} a_{t+1}(m) &= \textit{old} + \textit{incoming} + \textit{noise} \\ \textit{old} &= (1 - \textit{decay}) \times a_t(m) + \textit{decay} \times \textit{resting}(m) \\ \textit{incoming} &= \sum_{n \in N} \max[0, \textit{connection}(n, m) \times a_t(n)] \end{aligned}$$

where  $a(m)$  represents the activation of node  $m$  at a certain time step (either at the current time step or the previous).  $N$  is the set of all other nodes that are connected to node  $m$ . This architecture and these dynamics are general to a localist connectionist implementation of the general model of lexical access. *Decay*, *resting*, and *connection* are parameters of the model. The *connection* and *resting* parameters are especially important for our modelling work. *Connection*( $n, m$ ) refers to the strength of the connection from node  $n$  to node  $m$ . These parameters may be all set to the same value, or may vary by levels in the system (depending on the implementation of damage, see below). They may also vary by individual items (specifically for the target), an aspect we utilize for our frequency modelling (see below). The *resting* parameter reflects resting activation, which is an implicit part of this modelling approach (Dell, 1990), but has been set to zero for the above studies. This parameter is included in the equations because we utilize it for one of our frequency manipulations (see below).

In addition to these general assumptions, Dell et al. (1997) add several additional assumptions and specified additional parameters. A major assumption that they include is interactivity in the system. In other words, all connections between nodes are bidirectional, so activation flows forwards and backwards in the system. The other major assumption of Dell et al. (1997) is that of distributed semantic representations—that is, semantic features are represented without unitary concept nodes (see Rapp & Goldrick, 2000, for a modified architecture that includes such nodes). Another added assumption is that all connections are excitatory (not inhibitory) and that there are no connections between nodes at the same level. A related assumption that is also included is that while negative activations can exist at nodes (due to noise), they are not propagated.

The following parameters are also part of Dell et al.'s (1997) implementation: original activation of target's semantic nodes, time steps until lexical selection, jolt activation to winning node of lexical selection, and time steps until selection of segments. Lexical access is simulated in the model by setting the activation level of the semantic representations associated with the target word to the predetermined value (fixed at 10 units). The activation levels of all representations in the network are then updated for the specified number of time steps (fixed at 8 steps). After this time, the activation level of the most active lexical representation is raised to a predefined high level (a level of 100 units was used), corresponding to the notion of lexical selection. Activations are further propagated for another specified number of time steps (also 8 steps), after which the most active onset, vowel, and coda phonemes are chosen as the output of the model. The fixed number of time steps does not allow the model to predict reaction times in a straightforward manner. Through the addition of the lexical jolt, the Dell et al. implementation incorporates the idea of two separate selection steps: one lexical, and one phonological.

Without noise, the general model cannot make an error. In order to produce errors, noise is included in the first equation. The Dell et al. (1997) implementation specifies two types of noise, calculated by the following equation:

$$\text{noise} = R(\text{intrinsic}) + R(\text{activation}) \times a_t(m)$$

where  $R(x)$  represents a random sample drawn from a normal distribution with a mean of zero and a standard deviation of  $x^5$ . *Intrinsic* and *activation* are parameters in the Dell et al. implementation (fixed to 0.01 and 0.16, respectively). There is both an intrinsic noise component in each node and a noise component that depends on how highly activated the node is. Because of the contribution of noise to the activation levels, the selected lexical representation does not necessarily correspond to the target word, and even if it does, the selected segmental representations are not necessarily those associated with the selected lexical node. Along with correct responses, the model can produce the following errors: semantically related, phonologically related (formal), mixed (both semantically and phonologically related), unrelated, and nonword. By simulating many trials, one can accumulate an estimate of the probabilities of the various categories of responses and determine whether the model represents a mechanism sufficient to generate a distribution measured from a human experimental participant.

Our own work uses the Dell et al. (1997) model, but tries to generalize across two key assumptions that have been questioned in the literature: neighbourhood and level of interactivity. We discuss these assumptions and the steps that we took to generalize across them in our own modelling.

### Neighbourhood

The lexicon used by Dell et al. (1997) includes only a few word nodes. While many magnitudes

smaller than a typical speaker's lexicon, the relationships between these nodes were nevertheless designed to reflect the neighbourhood of the average word that is used in picture-naming experiments. The original neighbourhood was constructed to reflect random error opportunities in the English language. It consists of a target word (e.g., "cat"), a semantic neighbour (e.g., "dog"), two formal neighbours (e.g., "hat" and "mat"), and two unrelated neighbours (actually phonological neighbours of the semantic neighbour of the target; e.g., "log" & "fog"). There is also a slightly modified version of the network that replaces one formal neighbour (e.g., hat) with a mixed (i.e., semantically and phonologically related) neighbour (e.g., "rat"). Runs from these two neighbourhoods are sampled 90% and 10% of the time, respectively, to produce a final distribution of errors. All words in this neighbourhood (and those described below) were of the form consonant-vowel-consonant.

Rapp and Goldrick (2000) have criticized the above neighbourhood. They determined that words that are semantically related have a slightly higher chance of also being phonologically related. They pointed out that semantically related items in the original neighbourhood were slightly less phonologically related to the target. To remedy this they created a larger, 29-word lexicon for modelling, which takes this phonological overlap into account. This neighbourhood has the added advantage of producing all possible error types using a single neighbourhood (as compared to sampling from two neighbourhoods, which artificially limits the maximum number of mixed errors).

We used both these neighbourhoods in our modelling work to be confident that our results would not be tied to idiosyncratic elements of a single neighbourhood. To further confirm the reliability of our models, we also ran our simulations with a third neighbourhood<sup>6</sup> that was

<sup>5</sup> Our implementation of Dell et al.'s (1997) model uses George Marsaglia and Wai Wan Tsang's (2000) ziggurat algorithm for generating normally distributed random numbers. We thank them for making their code publicly available.

<sup>6</sup> We would like to thank Wheeler Ruml for providing us with this neighbourhood and for creating the algorithm by which it was constructed.

created using an automated search algorithm to better fit both the criteria of random error opportunity (Dell et al., 1997) and greater phonological opportunity (Rapp & Goldrick, 2000). The details of the automated search procedure can be found in Rumel et al. (2005, Appendix B). Throughout the rest of this paper, we refer to these neighbourhoods by the number of lexical nodes (N) that they contain: 6 N for Dell et al.'s (1997) original neighbourhood, 20 N for our new neighbourhood, and 29 N for Rapp and Goldrick's (2000) neighbourhood.

### *Interactivity*

As we describe above, Dell et al.'s (1997) original model is highly interactive. That is, there is feedback from the segmental level to the lexical level and from the lexical level to the semantic level. There has been disagreement for quite some time over whether the production system is interactive and, if so, to what extent (e.g., Goldrick & Rapp, 2002; Rapp & Goldrick, 2000; Rumel et al., 2005).

Even though originally conceived as being an interactive model, the implementation of the Dell model easily allows for modification of the interactivity assumption. In the current work, we included models with full interactivity, like the original model, with restricted interactivity, where there was only feedback from the segmental level to the lexical level (Goldrick & Rapp, 2002; Rapp & Goldrick, 2000), and with no interactivity, a simple cascaded model. We were interested in seeing how our frequency manipulations would be affected by changes in the interactivity of the system. Combining the three types of neighbourhood with the three types of interactivity produced nine models. When discussing these models we use the following abbreviations to refer to their level of interactivity: "H" for high interactivity, "R" for restricted interactivity, or "C" for cascaded.

### *Implementations of brain damage*

What we have described so far is the normal production model underlying our work. In order to model E.C.'s condition we must make further assumptions about how the normal system is damaged. In our current work we considered two implementations of brain damage: weight-decay and semantic-phonological connection strength (henceforth semantic-phonological). Weight-decay instantiates the global damage (or globality) hypothesis of brain damage (Dell et al., 1997), while semantic-phonological instantiates the levels of damage hypothesis of brain damage<sup>7</sup> (Foygel & Dell, 2000; see also Rumel & Caramazza, 2000).

The global damage hypothesis assumes that aphasic performance can be captured by diffuse damage to the entire model. The weight-decay implementation simulates global damage to the system by changing the global connection strength of the model and/or the rate of activation decay (in the equations presented above, these would be the *connection* and *decay* parameters, respectively). The levels of damage hypothesis, contrary to the global damage hypothesis, claim that there are independent levels in the system where damage can occur. In word production, these levels have been taken to be the semantic and phonological levels. In the semantic-phonological implementation this damage is simulated by changing the connection weights between the semantic and lexical nodes and between the lexical and segmental nodes independently. Thus semantic and phonological deficits can be more intuitively modelled in the system. There is mounting evidence suggesting that this is the preferred method for modelling patients (Foygel & Dell, 2000; Rumel & Caramazza, 2000; Rumel et al., 2000; Schwartz et al., 2006), but we include an implementation of the global damage hypothesis for the sake of completeness (cf. Dell, Lawler, Harris, & Gordon, 2004; Schwartz et al., 2006). Combining the 9 models from the previous

<sup>7</sup> We wish to stress the difference between theory and implementation, since it is possible to implement the same theory in different ways. For example, see Rapp and Goldrick (2000) for other implementations of the global damage and levels of damage hypotheses.

section with these two implementations of brain damage produced a total of 18 models. In discussing our models, we use the following abbreviations to refer to the type of brain damage implementation: Global for the global damage implementation and Levels for the levels of damage implementation.

### *Fitting patients*

Before modelling E.C.'s errors as a function of frequency, it was necessary to determine the correct damage settings for each type of model (i.e., each combination of neighbourhood type, interactivity type, and damage implementation). Fitting such models to patients has been considerably researched (e.g., Dell et al., 2004; Foygel & Dell, 2000; Rumel & Caramazza, 2000) and consists of mapping out the parameter space of the damage theory to a certain grain. At each combination of parameters, the model is run extensively (10,000 trials) to get an estimate of the distribution at those settings. With such a map, fitting a patient consists of finding the point on the map with a distribution most similar to that of the patient, using  $\chi^2$  as the measure of fit (e.g., Dell et al., 2004; Foygel & Dell, 2000). We fit our models to E.C. using a method quite similar to that of Foygel & Dell (2000).<sup>8</sup> The parameter settings of this fit for each model were used as the "average" frequency point in the data obtained from our manipulations of frequency in the models.

We fitted E.C. based on his overall performance on the first and second administrations of the Snodgrass and Vanderwart (1980) picture set and on a single administration of the Philadelphia Naming Test. The same items as

those that were used in the patient analyses (see "Patient E.C." section above) were used for the fitting procedure. E.C.'s performance across the three administrations did not differ reliably:  $\chi^2(5) = 4.60$ ,  $p = .467$ , for the first and second administrations of the Snodgrass and Vanderwart pictures;  $\chi^2(5) = 4.79$ ,  $p = .442$ , for the first administration of the Snodgrass and Vanderwart pictures and the Philadelphia Naming Test;  $\chi^2(5) = 2.41$ ,  $p = .790$  for the second administration of the Snodgrass and Vanderwart pictures and the Philadelphia Naming Test.

### *Frequency implementation*

The architecture of the general model of lexical access lends itself to three loci for frequency: the interface between semantic and lexical levels, the lexical level, and the interface between lexical and segmental levels. Lexical frequency can be implemented as either of the connection strengths between representational levels or the resting activation level of the lexical nodes, respectively (e.g., Dell, 1990; Stemmer, 1985).<sup>9</sup> Specifically, we allowed the frequency of the target to vary, keeping the rest of the network resting activations or connections as they were.

The first step was to fit each model to E.C.'s performance. Taking this fitted model, we then applied each of the above frequency manipulations in turn. We allowed target resting activation to vary from  $-0.10$  to  $+0.10$  (inclusive) in 0.01-unit steps. For connection strength (either semantic-lexical or lexical-segmental), we allowed the target connection strength to vary from 0.00 to  $+0.20$  (inclusive) in 0.01-unit steps. At each point we ran the model 10,000

<sup>8</sup> Both our mapping procedure and that of Foygel and Dell (2000) used a smart procedure to only add points to the parameter map if they were significantly different from their surrounding neighbours (based on a  $\chi^2$  criterion). The slight difference between our procedure and theirs is that in ours the decision to add each new point is done on a point-by-point basis and not for groups of five points. Our maps allowed connection strength to vary from 0.001 to 0.1 and decay to vary from 0.5 to 1.0. The maximum level of expansion (i.e., grain) that we allowed was 1/128. We created a map for each of the nine models (i.e., three levels of interactivity crossed with three neighbourhoods) with each of the two implementations of damage, for a total of 18 maps. In testing our procedure, we obtained fits for their patients that were comparable to published results.

<sup>9</sup> Another possible locus of frequency within the model is in the jolt strength. Higher frequency lexical nodes can receive a higher jolt of activation after they are selected. The results of such an implementation should be quite similar to the lexical-segmental connection locus that we do implement, since both have frequency effects occurring mostly after lexical selection. Please see our General Discussion for our reasons for not manipulating jolt strength in the model.

times to obtain a distribution of errors. Distributions for settings greater than the “average” point would reflect modelled performance on increasingly higher frequency words, and likewise points smaller than the “average” point would reflect performance on lower frequency words. We deliberately tried to capture a wide range of the model’s performance over our frequency manipulations—most likely greater than the frequency difference that could be obtained in an experimental setting.

### *Analysis*

While it is important to look at the qualitative patterns produced by the different frequency implementations to see how well they capture E.C.’s data, it is equally important to be able to quantify their goodness of fit, especially since we wish to directly compare the three frequency implementations. In our approach, we split E.C.’s data into low-, medium-, and high-frequency groups and compare within each of our 18 models how well each implementation can match E.C.’s low- and high-frequency error distributions. We discuss the details of this process and the analysis of the results below.

We looked at how well the three frequency implementations could capture E.C.’s performance. This was done in the following way. For each model we found the point that minimized the  $\chi^2$  difference between the model’s error distribution and E.C.’s low-frequency error distribution. Then, we also found the point with the best  $\chi^2$  fit to E.C.’s high-frequency error distribution. We restricted the range of the search for the best low-frequency point to be below the “average” frequency point. Likewise, the range of the search for the best high-frequency point was restricted to be above the average frequency point. We did not search for the best  $\chi^2$  fit of E.C.’s medium frequency distribution, since theoretically the “average” point was most appropriate to match this distribution.

We ran a single repeated measures analysis of variance (ANOVA) on the high and low  $\chi^2$  fits of the models. The within-models factors were frequency implementation (i.e.,

semantic-lexical connections, lexical resting activations, or lexical-segmental connections) and frequency distribution (i.e., high or low). Any significant differences observed in this analysis must have their source in the frequency implementations since the same underlying models were used each time. Based on our account of E.C.’s performance, we predicted that implementing frequency in the lexical-segmental connection strengths would give the best fits for E.C.’s performance. We did not include the fits to E.C. medium-frequency distribution, since these were exactly the same across the three frequency implementations for each model.

### **Results**

A total of 54 model runs were produced by crossing the factors interactivity (H, R, or C), neighbourhood (6N, 20N, or 29N), damage implementation (global, levels), and frequency implementation (semantic-lexical connections, lexical resting activations, and lexical-segmental connections). Some models were a better  $\chi^2$  fit to E.C.’s overall error distribution than others. The actual fits are summarized for each of the models in Table 1. Using our fitting procedure, the original model from Dell et al. (1997; i.e., global 6N-H) accounted best for patient E.C.’s overall performance, before frequency was implemented. Below we present the qualitative results of the individual runs and the quantitative analysis comparing the three frequency implementations.

### *Qualitative results*

In discussing our qualitative results, we focus on the frequency range in our modelling output that corresponds to E.C.’s performance range as we ascertained through our fitting procedures. While space here permits us to provide only several representative graphs of the semantic and nonword error trends, all modelling graphs (with all error types in colour) can be found online at: [http://www.wjh.harvard.edu/~caram/frequency\\_modeling/index.html](http://www.wjh.harvard.edu/~caram/frequency_modeling/index.html)

*Semantic-lexical connections.* We found that implementing frequency in the semantic-lexical connections produced error distributions that were different from E.C.'s pattern. Not surprisingly, implementing lexical frequency at this level in the model produced an effect of frequency on the semantic error rate whereas E.C.'s semantic error rate was not modulated by target lexical frequency. All 18 models showed this trend, along with a decrease of phonological, mixed, and unrelated errors as frequency increased, though different models predicted different rates of decrease for these errors.

With regard to nonword errors, there were two distinct patterns produced by our models. The six high-interactivity models showed nonword errors decreasing as a function of frequency, though at a slower rate than semantic errors, while all restricted interaction and cascading models showed no effect of frequency

on nonword errors. See Figure 3 for representative graphs of these two dominant patterns from the Levels 6N-H and Levels 6N-C models.

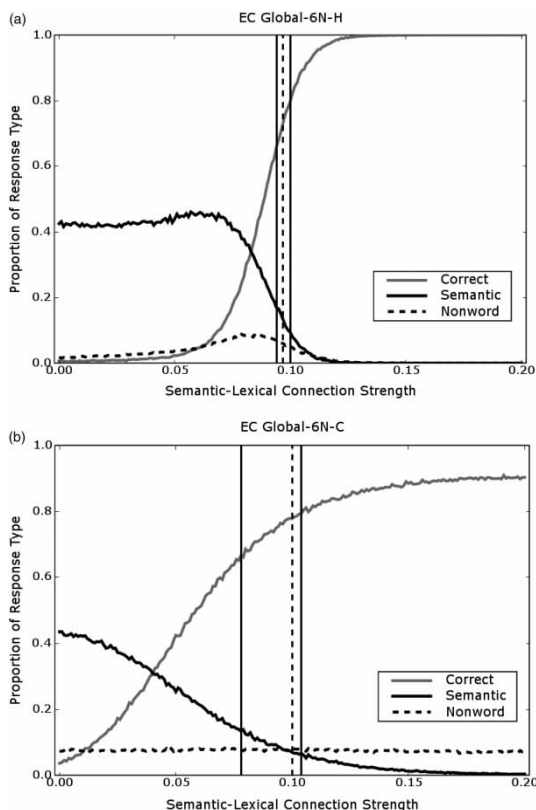
*Lexical resting activations.* Implementing lexical frequency in the resting activation levels of the lexical nodes led to a similar patterning of semantic, phonological, mixed, and unrelated errors across all 18 models. See Figure 4 for a representative graph from the Levels 6 N-H model. Just as when frequency was implemented in the semantic-lexical connections, all models produced fewer semantic errors as frequency increased. In contrast, these models showed little or no change in the rate of nonword error rates in the range of E.C.'s performance.

*Lexical-segmental connections.* When implementing lexical frequency in the lexical-segmental connections, all 18 models produced the same pattern

**Table 1.** All models fitted to E.C.'s performance to the combined data from two administrations of the Snodgrass and Vanderwart picture set and one administration of the Philadelphia Naming Test

Model	Correct	Semantic	Phonological	Mixed	Unrelated	Nonword	Sem Weight	Phon Decay	RMSD	$\chi$ -squared
EC	0.7204	0.1519	0.0259	0.0278	0.0056	0.0685				
Global 6N-H	0.7306	0.1336	0.0345	0.0350	0.0059	0.0604	0.0969	0.8516	0.010	3.79
Global 20N-H	0.6630	0.0756	0.0266	0.1122	0.0101	0.1125	0.0474	0.6836	0.055	83.92
Global 29N-H	0.7360	0.0890	0.0264	0.0788	0.0171	0.0527	0.0296	0.6172	0.035	46.16
Global 6N-R	0.7128	0.0789	0.0608	0.0081	0.0312	0.1082	0.0196	0.5820	0.039	84.25
Global 20N-R	0.6837	0.0474	0.0637	0.0495	0.0459	0.1098	0.0288	0.6016	0.054	150.93
Global 29N-R	0.7515	0.0684	0.0384	0.0456	0.0243	0.0718	0.0103	0.5195	0.038	63.30
Global 6N-C	0.7743	0.0742	0.0397	0.0090	0.0242	0.0786	0.1000	0.6602	0.041	70.40
Global 20N-C	0.7451	0.0467	0.0393	0.0404	0.0369	0.0916	0.0899	0.6484	0.047	130.88
Global 29N-C	0.7455	0.0751	0.0355	0.0445	0.0274	0.0720	0.0869	0.6445	0.035	52.64
Levels 6N-H	0.7520	0.0853	0.0557	0.0131	0.0299	0.0640	0.0211	0.0250	0.035	53.38
Levels 20N-H	0.7849	0.0521	0.0292	0.0555	0.0221	0.0562	0.0250	0.0281	0.051	107.27
Levels 29N-H	0.7194	0.0845	0.0341	0.0699	0.0198	0.0723	0.0250	0.0242	0.033	46.41
Levels 6N-R	0.7176	0.0910	0.0781	0.0109	0.0486	0.0538	0.0211	0.0296	0.038	73.55
Levels 20N-R	0.6867	0.0501	0.0702	0.0553	0.0574	0.0803	0.0265	0.0304	0.053	144.77
Levels 29N-R	0.7144	0.0661	0.0592	0.0583	0.0304	0.0716	0.0312	0.0281	0.041	82.67
Levels 6N-C	0.7403	0.0889	0.0544	0.0103	0.0401	0.0660	0.0250	0.0335	0.033	60.14
Levels 20N-C	0.7334	0.0545	0.0461	0.0507	0.0617	0.0536	0.0296	0.0420	0.048	121.36
Levels 29N-C	0.7874	0.0781	0.0317	0.0382	0.0215	0.0431	0.0366	0.0428	0.043	52.75

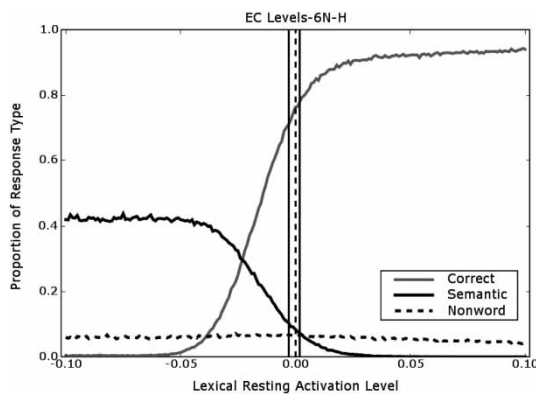
*Note:* Picture set: Snodgrass & Vanderwart, 1980. Philadelphia Naming Test: Roach et al., 1996. These were the fits of each model to the overall data before frequency implementations were tested. H = high interactivity; R = restricted interactivity; C = cascaded. N = number of lexical nodes. RMSD = root mean square deviation.



**Figure 3.** Example graphs of implementing frequency as the connection strength between semantic and lexical nodes. Panel a shows model Levels 6N-H, while Panel b shows model Levels 6N-C. The vertical dotted line shows the settings of the model that was fitted to E.C. before frequency manipulations and represents the average frequency word. The vertical solid lines to the left and right of the average line show the best model fits to E.C.'s low and high distributions, respectively.

of sharply decreasing nonword errors as target frequency increased. This decrease was larger than that observed in the two previous frequency implementations. Phonological and unrelated errors were also generally seen to decrease as frequency increased, but to a lesser degree than nonword errors.

The high and restricted interaction models produced fewer semantic and mixed errors as



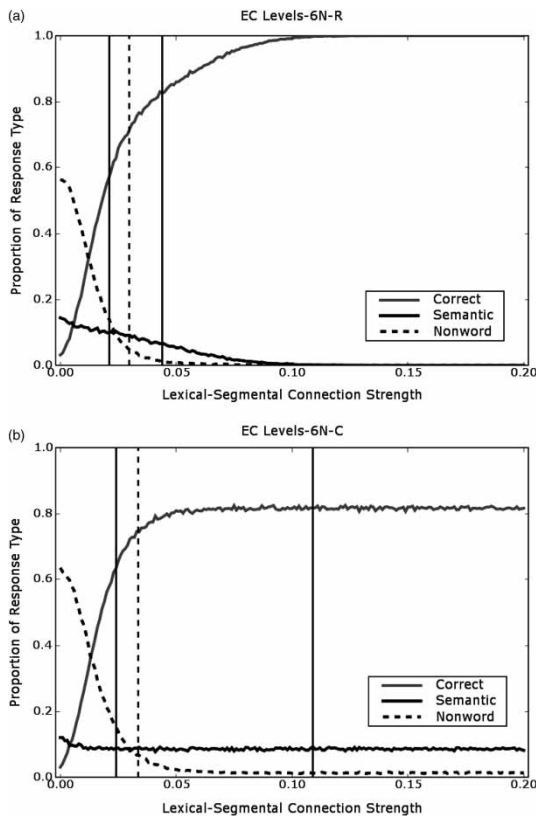
**Figure 4.** Example graph of implementing frequency as the resting activation of the lexical nodes. The model shown is Levels 6N-H. The vertical dotted line shows the settings of the model that was fitted to E.C. before frequency manipulations and represents the average frequency word. The vertical solid lines to the left and right of the average line show the best model fits to E.C.'s low and high distributions, respectively.

frequency increased, though this frequency effect was attenuated relative to the effect of frequency on the nonword errors. In contrast, the cascading models did not exhibit an effect of frequency on the semantic and mixed error rates. See Figure 5 for representative graphs of these two main patterns from Levels 6N-R and Levels 6N-C.

### Quantitative results

In order to analyse the effects of frequency implementation, we included model ( $N = 18$ ) as a random factor in a  $3 \times 2$  repeated measures ANOVA. Only the main effect of frequency implementation,  $MSE = 2,145.1$ ,  $F(1.18, 20.02) = 99.05$ ,  $p < .001$ ,<sup>10</sup> was significant. The lexical-segmental connection weights implementation of frequency produced better  $\chi^2$  fits (mean = 21.72,  $SD = 12.99$ ) than did either the semantic-lexical connection weights (mean = 31.57,  $SD = 13.21$ ) or lexical resting activation levels implementations (mean = 32.34,  $SD = 13.36$ ). The main effect of

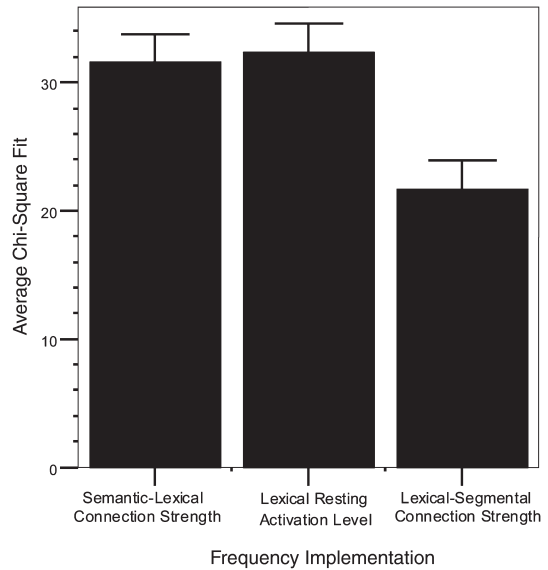
<sup>10</sup>The Huynh-Feldt correction was applied because the sphericity assumption was not met.



**Figure 5.** Example graphs of implementing frequency as the connection strength between lexical and segmental nodes. Panel a shows model Levels 6N-R, while Panel b shows model Levels 6N-C. The vertical dotted line shows the settings of the model that was fitted to E.C. before frequency manipulations and represents the average frequency word. The vertical solid lines to the left and right of the average line show the best model fits to E.C.'s low and high distributions, respectively.

frequency distribution,  $MSE = 712.7$ ,  $F(1, 17) = 1.53$ ,  $p = .233$ , was not significant. Fits to high- and low-frequency distributions were comparable. More importantly, the interaction between the two main factors was also not significant,  $MSE = 75.14$ ,  $F(1.07, 18.18) = 1.11$ ,  $p = .311$ .<sup>11</sup> The overall fits of the three frequency manipulations to E.C.'s error distributions, collapsed across low and high distributions, are shown in Figure 6.

<sup>11</sup> The Huynh-Feldt correction was applied because the sphericity assumption was not met.



**Figure 6.** Average  $\chi^2$  fits between each of the three frequency implementations and E.C.'s low- and high-frequency error distributions. Low- and high-frequency fits are combined for each frequency manipulation.

## Discussion

The purpose of this modelling work was to determine which implementation of lexical frequency would best fit E.C.'s pattern of performance. Remember, the key feature of E.C.'s performance was that his nonword error rate was modulated by lexical frequency, while his rate of semantic errors was not. The implementation of lexical frequency that best captured this feature of E.C.'s performance was the lexical-segmental connection strength implementation. Placing lexical frequency between the lexical and segmental layers gave the best fit for E.C.'s performance for both high- and low-target-frequency items; this was true for models incorporating different assumptions of interactivity and lexical neighbourhood sizes. In contrast, implementing lexical frequency in the semantic-lexical connections or in the activation levels of the lexical nodes themselves

failed to capture this critical feature of E.C.'s performance.

The results obtained for these three frequency implementations make intuitive sense. For example, if one were to strengthen the connection between the semantic and lexical nodes, this should lead to fewer semantic errors, but have little (if any) effect on the rate of nonword errors. The results of the modelling work support these intuitions. The opposite should be true when one increases the connection strength between the lexical and segmental nodes. In this case, the number of nonword errors should decrease, and there should be little to no concomitant effect on semantic errors. Again, the results of the modelling work support these intuitions. Less intuitively, implementing frequency as resting activations has a similar effect to manipulating semantic-to-lexical connections. Here the target's higher resting activation reduces competing word response, especially semantic errors, while having less of an effect on the nonword errors, since the jolt of activation that comes from lexical selection washes away most of the resting activation differences.

It may be surprising that the best fitting model for patient E.C. was the Global 6N-H, given the evidence that we have cited favouring the levels of damage hypothesis. This model was particularly successful in producing a high rate of semantic

errors. We believe this is due to the combination of a high decay rate with full interactivity, which has been shown to lead to increases in such errors (see Dell et al., 1997; Foygel & Dell, 2000; Ruml & Caramazza, 2000). Our implementation of the levels of damage hypothesis does not allow for manipulation of the decay parameter of the models. However, the success of this particular model does not mean that it is correctly simulating E.C.'s semantic errors. We concluded earlier that a good part of E.C.'s semantic errors were of a semantic origin, while the semantic errors that the model produces are only postsemantic or lexical in origin. This fact highlights an inherent limitation of all our models—namely, that they cannot produce semantic errors of a strictly semantic origin. While this limitation does not affect our investigation of the frequency effect, it is necessary to fully model E.C.'s word production deficits. We return to this point in the General Discussion.

The upshot of this modelling work is that it supports the hypothesis that at least one locus of the frequency effect should be in the connections between the lexical and segmental levels. This is the only apparent way to model E.C.'s performance successfully.<sup>12</sup> In the following section we review the data of a large set of Italian aphasic patients to determine how target lexical frequency modulates specific error types in a large population

<sup>12</sup> A potential challenge to this conclusion comes from a proposal by Dell (1990). He proposed, but never implemented, the hypothesis that lexical frequency effects in production could be explained as semantic and/or phonological neighbourhood effects. This hypothesis challenges an underlying assumption of our work, which is that the semantic and phonological neighbourhoods of a higher frequency word are the same as those neighbourhoods of a lower frequency word. This assumption is probably incorrect, at least for phonological neighbourhoods, given the significant correlation ( $r = -.49, p < .01$ ) between frequency and phonological neighbourhood size in our own data (see also Gordon, 2002). Nevertheless, while phonological neighbourhood size affects E.C.'s rates of phonological and mixed errors, significant effects of lexical frequency persist in regression analyses of E.C.'s performance, despite the inclusion of phonological neighbourhood size and other semantic and phonological measures as cofactors. Putting these analyses aside, we explored the possibility further by trying to simulate E.C.'s lexical frequency effects through modifying semantic and phonological neighbourhoods in the largest neighbourhood (that of Rapp & Goldrick, 2000). Adding more semantic neighbours led to the models (H, R, and C) producing more semantic and nonword errors, with a much larger increase in semantic errors. Overall accuracy also decreased. This pattern was quite unlike E.C.'s. In the case of adding more phonological neighbours, models produced slightly fewer nonword and semantic errors, but instead of leading to fewer errors as a function of frequency (as in E.C.'s case), this decrease led to a corresponding increase in phonological errors (see also Dell & Gordon, 2003). It is clear from these investigations that semantic and phonological neighbourhoods cannot explain the lexical frequency effects that we observe in E.C.'s performance.

of patients. This review reveals that E.C.'s pattern of performance is not an isolated pattern. This review also reveals a patient whose performance is best explained by implementing lexical frequency in the semantic-to-lexical connections.

## ITALIAN PATIENTS

In the first section of this article, we reported patient E.C., who showed a clear effect of frequency on his overall naming performance. This effect was carried by a decrease in his nonword errors as target frequency increased; his rate of semantic errors, in contrast, was not modulated by lexical frequency. In the present section, we investigate a large group of aphasic patients to see whether E.C.'s pattern of errors replicates in other patients. Furthermore, we investigate the possibility that some patients might exhibit other patterns that lend themselves to clear interpretation with regard to understanding frequency. To this end, we reanalysed the picture naming data from 48 out of the 50<sup>13</sup> aphasic participants reported in Ruml et al. (2005) with regards to frequency.

### *Participants*

The 48 aphasic patients in this study were 33 males and 17 females, of whom 46 are right-handed, and 2 are left-handed postmorbidly. A total of 40 suffered from CVAs, 3 from primary progressive aphasia (PPA), and 5 from herpes virus encephalitis (HVE). Detailed information about each participant is provided in Appendix A of Ruml et al. (2005).

### *Picture naming*

Pictures ( $N = 128$ ) consisting of 11 semantic categories (animals, body parts, fruits and vegetables, food, professions, kitchenware, clothing, tools, furniture, means of transportation, and musical instruments) were used to assess naming performance. Participant F.D.I. completed three administrations of this list, while participants

P.G.E., E.M.A., G.M.A., and S.F. also named additional stimuli.

Scoring was again in line with Dell et al. (1997) and similar to the system used for E.C., with the following exceptions for correct and phonological (word) error classification. In classifying correct responses, changes in inflectional morphology were still considered correct. In the case of phonological errors, using the loose criteria that Dell et al. use to classify these errors in English would lead to an inordinate amount of errors being classified as phonological in Italian. This is due to two main differences between the languages: (a) Italian words tend to be longer and do not contain any unstressed vowels like the English schwa (~), and (b) most words in Italian end in one of four possible vowels (a, e, i, o). Ruml et al. (2005) made the following modified criteria to obtain a more reasonable measure of phonological overlap in Italian: The target and response must share: (a) one third of the phonemes irrespective of sequence, (b) two initial phonemes, (c) two phonemes from the first syllable in any sequence, or (d) three phonemes in the same sequence in any position. In all cases the inflectional (final) vowel was not counted.

### *Analyses*

We analysed the Italian participants in a manner similar to that of E.C. The following analyses were done for each participant individually. We performed an overall binary logistic regression on correct responses versus errors with frequency (Bortolini, Tagliavini, & Zampolli, 1972) and letter length.<sup>14</sup> We also ran separate binary logistic regressions for each error type. Thus there was a separate analysis for semantic, phonological, mixed, unrelated, and nonword errors. As in our analyses of E.C.'s performance, we again did not include "other" errors in these analyses, since this was a heterogeneous category and not able to be simulated in current instantiations of speech production models. The independent variables were again log frequency

<sup>13</sup> We were unable to obtain enough items with frequency ratings for patients A.S. and C.L.B.

<sup>14</sup> Given Italian's shallow orthography, letter length is very close to the phoneme length.

and letter length.<sup>15</sup> Since we conducted five analyses on the same dataset (the overall analysis and analyses of semantic, phonological, mixed, and nonword errors), we adjusted our critical  $\alpha$  level to .01 using Bonferroni correction. We report significant,  $\alpha \leq .01$ , and marginal values,  $.01 < \alpha \leq .05$ . Confirmatory analyses of the other factors using the larger dataset without imageability ratings were held to the same critical alpha level.

## Results

A total of 11 of the patients (out of 48) showed a significant effect of log frequency on overall naming; 6 more showed a marginal effect. For each of these patients, error rates decreased as target lexical frequency increased. A total of 3 also showed a significant or marginal length effect; 3 other patients showed significant or marginal effects of length, but not of frequency. All patients showing frequency effects made fewer errors as frequency was increased, and likewise, all patients showing length effects made fewer errors as length decreased.

Looking now at specific error types, starting with semantic errors, we found 5 patients (A.C.O., G.I.M., E.M.A., G.M.A., and S.F.) who showed significant log frequency effects on semantic errors and another (W.M.A.) who showed a marginal effect. In all cases, the rate of semantic errors increased as target frequency decreased. There was also a patient (M.I.O.) who showed a significant length effect, where longer words led to more semantic errors.

When looking at nonword errors, patients G.N.I. and I.F.A. exhibited a marginal frequency effect. For these patients lower frequency targets led to more nonword errors. A total of 3 other patients (D.R.U., F.S., and S.F.) showed marginal length effects on nonword errors. Longer words led to more nonword errors for all patients.

Patients made very few phonological, mixed, and unrelated errors. For each of these error

types, patients did not exhibit any effects of either frequency or length.

## Discussion

The results of these analyses revealed 2 patients, who, similar to E.C., showed frequency effects on nonword errors. Neither showed a frequency effect on semantic errors, also like E.C. However, this finding is not surprising, since G.N.I. and I.F.A. did not make many semantic errors.

In contrast to E.C. and the above 2 patients, we also found 6 patients who showed frequency effects on semantic errors, and none of these patients showed a frequency effect on their nonword errors. Again, this was to be expected for 5 of them, since they did not make many nonword errors. However, W.M.A., who made 68 nonword errors, also did not exhibit a frequency effect on these errors. Thus, patient W.M.A. shows the opposite pattern to that of patient E.C. (see Figure 7).

As we discussed in the Introduction, nonword errors can stem from two loci: either in segmental processing in the lexical access stage, or at any postaccess stage, such as articulation. W.M.A. had severe difficulties with reading and repeating both words and nonwords (see Miceli, Benvegna, Capasso, & Caramazza, 1997, for a full case description of W.M.A.). Furthermore, W.M.A. also had similar difficulties in picture naming, though his performance was better than in reading and repetition. In this regard W.M.A.'s profile is similar to E.C.'s. The important difference between these patients is that W.M.A. has "severe dysarthria" (p. 52), while E.C. does not. Such a peripheral disorder easily explains both why W.M.A. produced so many nonword errors and why there was no frequency effect on their production.

There is a similar explanation for why W.M.A. shows a frequency effect on his semantic errors,

<sup>15</sup> Unfortunately we were unable to get norms for enough of our items on semantic or other phonological factors to merit an analysis including them. Since we did not include imageability and phonological neighbourhood size into our analyses of the Italian patients, any frequency effects that we describe here could potentially be effects of imageability and/or phonological neighbourhood size, even though those factors were not responsible for the frequency effects observed in patient E.C.

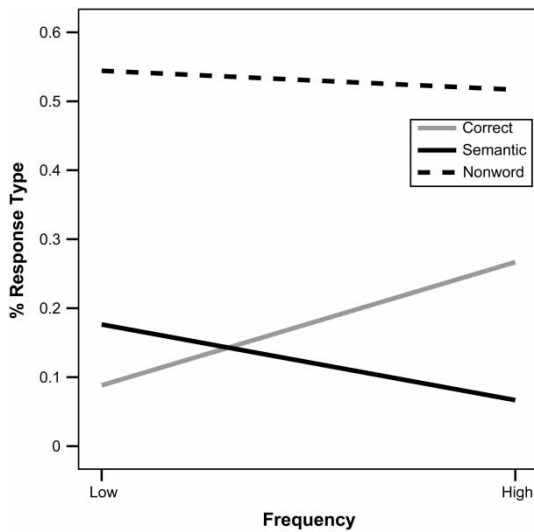


Figure 7. *W.M.A.'s response types by frequency. Low-frequency ( $N = 68$ ) items have a frequency  $\leq 3$ , and high-frequency items ( $N = 60$ ) have a frequency  $> 3$ .*

but E.C. does not. Semantic errors, like nonword errors, may also have several sources. We have already argued that E.C.'s semantic errors mainly come from his damaged semantic system and thus do not exhibit a lexical frequency effect. On the other hand, while W.M.A. clearly has semantic damage, he may also have a further deficit in the lexicon and/or in the connections between semantics and the lexicon. Either of these deficits is consistent with the frequency effects observed on W.M.A.'s semantic errors.

In explaining W.M.A.'s pattern of performance, we are led to consider a different locus of frequency in the production system in order to explain the frequency effect observed on his semantic errors. His pattern resembles the patterns exhibited by our models in which frequency was implemented in either the semantic-lexical connections (see Figures 3a and 3b) or at the lexical resting activations (see Figure 4). In fact, W.M.A.'s pattern, together with those of several other Italian patients and of other patients reported in the literature (e.g., Caramazza & Hillis, 1990), provides support for placing an additional locus of the frequency effect in either

the connections between semantics and the lexical level or at the lexical level itself.

## GENERAL DISCUSSION

We have reported results from three separate lines of research. First, we reported findings from an in depth investigation of a single patient's (E.C.) error types in a picture-naming task as a function of lexical frequency. The principal finding of E.C.'s performance was that his nonword errors (generally phonological approximations of the targets) exhibited a very strong effect of lexical frequency whereas his semantic errors did not. Second, we reported findings from a computational modelling approach in which we contrasted several different implementations of lexical frequency to determine which implementation best captured E.C.'s performance. The principal finding of this work was that implementing lexical frequency in the lexical-segmental connections (but not the semantic-lexical connections or the resting activation levels of lexical representations) best captured E.C.'s performance. To determine the generality of E.C.'s pattern of findings, we analysed the specific error types of 48 different aphasic patients in a picture-naming task. This analysis revealed that 2 of the 48 patients exhibited a similar pattern to that of E.C., suggesting that it is not unusual to find an influence of lexical frequency on nonword error rates.

The finding that an individual's likelihood of making a nonword error is modulated by the frequency of the to-be-named item is striking and has important implications for how we conceptualize the dynamical properties of the speech production system. To the extent that nonword error rates are modulated by the target's lexical frequency, then it must be that these errors arise at a level of processing in which lexical frequency is represented. Because it is very unlikely that lexical frequency is represented in phonological segments, the most reasonable way to account for this finding is to place lexical frequency in the lexical-segmental connections. As we mentioned

above, the modelling work confirmed this conclusion.

Of course, lexical frequency is highly correlated with other variables, such as length (in this case, number of phonemes) and number of neighbours (phonological), and it may be that these correlated variables, not frequency, modulate nonword error rates. To test for this possibility, we included each of these variables as predictors of E.C.'s nonword errors in a simultaneous regression analysis. We found that, though each variable contributed to E.C.'s nonword error rate, lexical frequency explained 2.5% of the variance independently of length (1.6% variance independently) and neighbourhood size (0.2% variance independently). As such, we are confident that, while variables correlated with lexical frequency undoubtedly contribute to the likelihood of observing a nonword error, the effect of lexical frequency on this error type is reliable and independent of these correlated variables.

#### *Alternative accounts to the lexical-segmental hypothesis*

We have argued that the most straightforward way to account for E.C.'s performance (and the performance of the three patients who showed a similar profile) is to implement lexical frequency in the lexical-segmental connections. But is this the only way to capture E.C.'s performance? As we have noted in our modelling discussion, the lexical jolt assumed by our modelling work (Dell et al., 1997) washes out frequency differences when frequency is implemented as lexical resting activations or semantic-lexical connections strengths. Removing this jolt of activation would allow such implementations to also strongly affect nonword errors through cascading (as they already do to a much smaller extent). When they reduced the strength of the jolt in their simulations, Rapp and Goldrick (2000) found a corresponding increase in nonword errors.

The gain from removing the jolting mechanism might be a single locus account of frequency effects on both semantic and phonological nonword error rates. However, removing the lexical jolt of activation would require a sizeable

reconceptualization of this modelling framework, since the jolt of activation represents the lexical selection process. The jolt would either have to be replaced with a different selection mechanism that does not dilute activation from frequency differences, or one would have to take the more radical step of eliminating lexical selection from their theory of lexical access. Clearly further research needs to be done in this area before we can be certain of whether such a step is the right explanation for E.C.'s pattern.

A related alternative explanation, also supported by Rapp and Goldrick's (2000) modelling work, is that perhaps E.C. has a damaged jolting mechanism. Instead of total elimination, it may be sufficient to weaken the amount of jolt activation that the winning lexical item receives. This type of damage should also lead to more nonword errors for higher frequency words, even if the locus of frequency is in the lexical nodes or the semantic-lexical connections. Such an account needs only to assume a single locus of frequency to explain our results.

While such an account may seem appealing, using the jolting mechanism of the model as a theoretical explanation at this time is premature for several reasons. Currently, the jolting mechanism is used to add seriality to the model by enhancing the winner of lexical selection, but it is not the only possible implementation of seriality. As Rapp and Goldrick point out (2000, Footnote 9, p. 480) a competitive output mechanism can serve the same purpose.

Furthermore, it is unclear what theory of damage is instantiated by damaging the jolting mechanism. The jolt for lexical selection has been viewed as a signal from the syntactic frame to produce a particular word in an utterance (Dell et al., 1997), so most likely, damage to the jolting mechanism would thus reflect a theory of syntactic damage. It is possible that E.C.'s syntactic system is damaged in some way and is leading to his word production deficits. However, we cannot produce any finer grained predictions for his syntactic performance from the model, since the syntactic system of the theory has never been implemented. We are

only left with the prediction about rates of nonword errors, but this is just a small part of the predictions that a syntactic theory of damage would make about performance. Without the rest of these predictions for comparison, it is impossible to truly evaluate whether damaging the jolting mechanism is the proper way to model E.C.'s syntactic damage. Such difficulties make it premature to consider damaging the jolting mechanism as an alternative account to the one that we have proposed.

#### *One locus or several?*

We have suggested that our findings support a lexical-segmental locus of the frequency effect, but we should be clear that we do not think that this is the only place that lexical frequency exerts its effects. For example, in the sample of 48 patients, one (W.M.A.) exhibited a pattern of performance that is best explained by implementing frequency in the semantic-lexical connections. We do not take this finding to be contradictory to our suggestion of a lexical-segmental locus for frequency. Rather, we suggest that lexical frequency is probably represented throughout all of the levels and connections that participate in the lexical access process. This notion of a "distributed" lexical frequency follows naturally from the activation framework upon which models of speech production are constructed. Word production involves the translation of one's intentions to speak into motor movements. Thus, to the extent that each unique utterance involves processes to be carried out over unique representations and connections, it follows that the frequency with which one translates a particular intention into articulated speech (e.g., "dog" or "eagle") should have a bearing on all processes essential to the translation of those intentions.

#### *Implications for extant models of word production*

Our results have serious implications for how frequency effects should be modelled in lexical production. Models such as those by Levelt et al. (1999) and Dell (1990) split the lexical level into two stages to account for morphological processes.

Dell claimed that the locus of the frequency effect is at the whole-word or lemma stage, while Levelt et al. have argued for placing the frequency effect at the subsequent morpheme or lexeme stage. We claim that neither of these accounts can accommodate our findings.

We have explicitly shown that the model by Dell and colleagues (1997) does not explain frequency effects on nonword errors with frequency only at the lexical level, but this version of the model has only a single lexical level, while the version used by Dell (1990) had two lexical levels. Nevertheless, the addition of another lexical layer should not change the finding that implementing frequency at either lexical level cannot explain frequency effects on nonword errors as long as there is some seriality mechanism between the lexical selection(s) and the selection of segments.

It is very difficult to extrapolate our modelling results to the model by Levelt and colleagues (1999), given not only the differences in number of lexical levels, but also in the general processing dynamics. Furthermore, their model also has two lexical levels and does not lend itself to modelling patient errors. Nevertheless, their current assumptions with regard to frequency seem incapable of theoretically explaining frequency effects on sublexical processes. We direct our criticisms to specific features of their own modelling architecture. The first alternative for implementing frequency in their modelling framework, proposed by Jescheniak and Levelt (1994), represents frequency as the selection threshold for lexeme (i.e., morpheme) representations. Higher frequency lexemes have lower selection thresholds than lower frequency lexemes. Frequency, implemented in this way, would not affect the amount of activation reaching a word's segments and should not affect their chance or rate of selection. The second alternative, proposed by Roelofs (1997), implements frequency in the process of verifying that the correct lexeme has been selected, with higher frequency lexemes being verified faster than lower frequency lexemes. This implementation also seems to fail in explaining a frequency effect on nonword errors, since the verification

process does not affect the amount of activation flowing to a word's segments based on its frequency.

The model by Dell and colleagues (1997) can be extended to cover our results by adding a frequency locus in the connections between lexical nodes and their segments. However to still explain the frequency effects on semantic errors by patients and the work on frequency effects in normals, this new locus must be in addition to the current locus of frequency that these models currently posit. The model by Levelt and colleagues (1999) may be modified in a similar manner to obtain frequency effects, but since it allows cascading from the lexeme node to its segments without a jolt of activation interceding, a simpler modification is possible. Frequency could be implemented as lexeme resting activation levels, instead of the current proposals of activation threshold or the verification time. Higher frequency words would have higher resting level activations and thus more activation would flow to their segments, making their selection more likely. Such a modification would straightforwardly explain frequency effects on nonword errors, but it would seem necessary to include an additional locus of frequency in the model to explain frequency effects on semantic levels, perhaps at the level of the abstract word form (lemma), since there is no cascading of information from the lemma level to the lexeme level. While our proposed modifications to these models do not go against their major architectural assumptions, these modified models must be thoroughly tested to see whether they perform as well as the original models, specifically with regard to frequency effects, but also with regard to other effects when frequency is concurrently modelled.

Our results may also be indicative of how frequency is represented in general by cognitive systems. Having several loci of frequency in a system is fundamentally problematic for models that can only accommodate a single locus of frequency. Specifically, our findings should be seen as problematic for the class of search models that have been proposed in the domain of word

recognition (e.g., Forster, 1976). These models implement lexical frequency effects as an ordered search through the lexicon. Higher frequency words are checked against the input before lower frequency words. While such a search can explain frequency effects at a single level quite well (Murray & Forster, 2004), extending it to explain effects at two or more levels becomes difficult. It is easiest to explain this difficulty with respect to a hypothetical search model of production. Such a model would easily explain frequency effects on reaction times and, with minor extensions, on errors as well. However, the errors that it would be able to explain are word errors, not nonword errors. Another mechanism would be necessary to explain nonword errors. However, the problem is that such a model would predict no difference in the rate of nonword errors as a function of the lexical frequency of the target. Since frequency is represented in the order of the items, once an item is selected, there is no residual effect (e.g., activation) that remains to influence access to its segments. Though the search model currently attempts to only explain frequency effects in word recognition, our results question its ability to generalize to other domains in explaining frequency effects.

#### *Future directions*

A question that our research does not resolve is whether the other locus of frequency is best modelled as connection strength between the semantic and lexical system or as the resting activation level of lexical nodes. Our own simulations show these two loci produce very similar qualitative results. Both predict a stronger effect of frequency on semantic errors than nonword errors such that relative to the total error distribution semantic errors decrease while nonword errors increase. Future work will have to take a more fine-grained approach to differentiating between these two accounts and the third possibility that both are involved in representing frequency.

While we have modelled the frequency effects observed in our patients' performance to a large extent, another limitation of our work is that our

model cannot capture our patients' word production disorders in their entirety. Our model is restricted to the lexical system, but we have argued that E.C. and several of our Italian patients have both prelexical semantic damage and postlexical phonological damage in addition to lexical damage. In order to model these latter types of damage it is necessary to extend our model with a semantic system (see Rapp & Goldrick, 2000, for an implementation of such a system) and a system for sublexical phonology. Such an extended model would have the additional benefit of allowing us to simultaneously model several loci of frequency at the same time, since we would have sources of errors that would not be affected by any lexical loci of frequency.

## CONCLUSION

We have reported patient E.C., who made predominantly semantic and phonological nonword errors in picture naming. In an analysis of his error types as a function of the lexical frequency of the to-be-named picture, we found that E.C.'s nonword errors exhibited an effect of frequency (high-frequency targets elicited fewer nonword errors) and that his semantic errors did not. We suggested that the best way to account for this finding is to place lexical frequency in the lexical-segmental connections. Support for this conclusion was provided through a computational modelling approach in which we contrasted several different implementations of lexical frequency and in which we found that implementing lexical frequency in the lexical-segmental connections (but not the semantic-lexical connections or the resting activation levels of lexical representations) best captured E.C.'s performance. Finally, we reported an investigation of 48 additional aphasic patients, which revealed that two of the 48 patients exhibited a similar pattern to that of E.C., suggesting that it is not unusual to find an influence of lexical frequency on nonword error rates.

It is important to reiterate that while we take these findings to suggest that lexical frequency

should be located in the lexical-segmental connections, we have not argued that this is the only possible locus of lexical frequency. In contrast, we have suggested that lexical frequency may be distributed throughout the lexical access system and that lexical-segmental connections constitute just one possible locus. If this conclusion is correct, this would have important implications for extant models of word production in capturing the dynamical properties of lexical access.

Manuscript received 17 January 2007

Revised manuscript received 4 June 2007

Revised manuscript accepted 11 June 2007

First published online 9 September 2007

## REFERENCES

- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX Lexical Database [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *Quarterly Journal of Experimental Psychology*, *50A*, 560–585.
- Bartram, D. J. (1976). Levels of coding in picture-picture comparison tasks. *Memory & Cognition*, *4*, 593–602.
- Bates, E., D'Amico, S., Jacobsen, T., Szekeley, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*, 344–380.
- Bird, H., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2000). The rise and fall of frequency and imageability: Noun and verb production in semantic dementia. *Brain & Language*, *73*, 17–49.
- Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*, 289–314.
- Bortolini, U., Tagliavini, C., & Zampolli, A. (1972). *Lessico di frequenza della lingua italiana contemporanea* [Frequency dictionary of the contemporary Italian Language]. Milan: Garzanti.

- Brown, G. D., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*, 208–216.
- Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology*, *12*, 65–85.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177–208.
- Caramazza, A., & Hillis, A. E. (1990). Where do semantic errors come from? *Cortex*, *26*, 95–122.
- Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the “tip-of-the-tongue” phenomenon. *Cognition*, *64*, 309–343.
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1–34.
- Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: Predictors and errors. *Brain & Language*, *82*, 344–365.
- Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 345–361.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language & Cognitive Processes*, *5*, 313–349.
- Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*. New York: Mouton.
- Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*, 125–145.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838.
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 515–523.
- Feyereisen, P., Van der Borgh, F., & Seron, X. (1988). The operativity effect in naming: A re-analysis. *Neuropsychologia*, *26*, 401–415.
- Finocchiaro, C., & Caramazza, A. (2006). The production of pronominal clitics: Implications for theories of lexical access. *Language & Cognitive Processes*, *21*, 141–180.
- Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 27–85). Amsterdam: North Holland.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning & Verbal Behavior*, *12*, 627–635.
- Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory & Language*, *43*, 182–216.
- Gagnon, D. A., Schwartz, M. F., Martin, N., Dell, G. S., & Saffran, E. M. (1997). The origins of formal paraphasias in aphasics’ picture naming. *Brain and Language*, *59*, 450–472.
- Gerhand, S., & Barry, C. (1999a). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, *73*, B27–B36.
- Gerhand, S., & Barry, C. (1999b). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & Cognition*, *27*, 592–602.
- Gilhooly, K. J., & Gilhooly, M. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition*, *7*, 214–223.
- Goldrick, M., & Rapp, B. (2002). A restricted interaction account (RIA) of spoken word production: The best of both worlds. *Aphasiology*, *16*, 20–55.
- Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, *102*, 219–260.
- Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain & Language*, *82*, 113–145.
- Hillis, A. E., Rapp, B. C., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, *7*, 191–243.
- Hirsh, K. W., & Ellis, A. W. (1994). Age of acquisition and lexical processing in aphasia: A case study. *Cognitive Neuropsychology*, *11*, 435–458.
- Howard, D., Patterson, K., Franklin, S., Morton, J., & Orchard-Lisle, V. M. (1984). Variability and consistency in picture naming by aphasic patients. *Progress*

- in aphasiology. In F. C. Rose (Ed.), *Advances in neurology* (Vol. 42). New York: Raven Press.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word-frequency effects in speech production—retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 824–843.
- Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition*, *13*, 789–845.
- Kay, J., & Ellis, A. (1987). A cognitive neuropsychological case study of anomia. Implications for psychological models of word retrieval. *Brain*, *110*, 613–629.
- Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning & Verbal Behavior*, *23*, 39–66.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–37.
- Lewis, M. B. (1999). Age of acquisition in face categorisation: Is there an instance-based account? *Cognition*, *71*, B23.
- MacKay, D. G. (1987). *The organization of perception and action. A theory for language and other cognitive skills*. New York: Springer Verlag.
- Marsaglia, G., & Tsang, W. W. (2000). The ziggurat method for generating random variables. *Journal of Statistical Software*, *5*, 1–7.
- Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture–word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 1146.
- Miceli, G., Benvegna, B., Capasso, R., & Caramazza, A. (1997). The independence of phonological and orthographic lexical forms: Evidence from aphasia. *Cognitive Neuropsychology*, *14*, 35–69.
- Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 116–133.
- Murray, W., & Forster, K. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, *111*, 721–756.
- Nickels, L. (1995). Getting it right? Using aphasic naming errors to evaluate theoretical models of spoken word production. *Language & Cognitive Processes*, *10*, 13–45.
- Nickels, L., & Howard, D. (1994). A frequent occurrence? Factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology*, *11*, 289–320.
- Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, *33*, 1281–1303.
- Oldfield, R., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17A*, 273–281.
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460–499.
- Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia naming test: Scoring and rationale. *Clinical Aphasiology*, *24*, 121–133.
- Roelofs, A. (1997). The weaver model of word-form encoding in speech production. *Cognition*, *64*, 249–284.
- Ruml, W., & Caramazza, A. (2000). An evaluation of a computational model of lexical access: Comment on Dell et al. (1997). *Psychological Review*, *107*, 609–634.
- Ruml, W., Caramazza, A., Capasso, R., & Miceli, G. (2005). Interactivity and continuity in normal and aphasic language production. *Cognitive Neuropsychology*, *22*, 131–168.
- Ruml, W., Caramazza, A., Shelton, J. R., & Chialant, D. (2000). Testing assumptions in computational theories of aphasia. *Journal of Memory and Language*, *43*, 217–248.
- Sailor, K., Antoine, M., Diaz, M., Kuslansky, G., & Kluger, A. (2004). The effects of Alzheimer's disease on item output in verbal fluency tasks. *Neuropsychology*, *18*, 306–314.
- Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*, 228–264.
- Schwartz, M. F., Wilshire, C. E., Gagnon, D. A., & Polansky, M. (2004). Origins of nonword phonological errors in aphasic picture naming. *Cognitive Neuropsychology*, *21*, 159–186.
- Silveri, M. C., Cappa, A., Mariotti, P., & Puopolo, M. (2002). Naming in patients with Alzheimer's disease: Influence of age of acquisition and categorical effects. *Journal of Clinical & Experimental Neuropsychology*, *24*, 755–764.

- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 174–215.
- Stadler, M. A. (1992). Statistical structure and implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 318.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language production* (Vol. 1). London: Lawrence Erlbaum Associates.
- Thompson-Schill, S. L., Gabrieli, J. D. E., & Fleischman, D. A. (1999). Effects of structural similarity and name frequency on picture naming in Alzheimer's disease. *Journal of the International Neuropsychological Society*, *5*, 659–667.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 735–747.
- Weide, R. L. (1994). *CMU pronouncing dictionary*. From speech at Carnegie Mellon University. Retrieved 7 February 2005, from <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>
- Wilshire, C. E. (2002). Where do aphasic phonological errors come from? Evidence from phoneme movement errors in picture naming. *Aphasiology*, *16*, 169–197.
- Wilshire, C. E., & Fisher, C. A. (2004). "Phonological" dysphasia: A cross-modal phonological impairment affecting repetition, production, and comprehension. *Cognitive Neuropsychology*, *21*, 187–210.
- Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, Version 2.00. *Behavior Research Methods, Instruments & Computers*, *20*, 6–10.
- Wingfield, A. (1967). Perceptual and response hierarchies in object identification. *Acta Psychologica*, *26*, 216–226.
- Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology*, *81*, 226–234.
- Zingeser, L. B., & Berndt, R. S. (1988). Grammatical class and context effects in a case of pure anomia: Implications for models of language production. *Cognitive Neuropsychology*, *5*, 473–516.