Journal of Memory and Language 63 (2010) 541–559

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Naming and repetition in aphasia: Steps, routes, and frequency effects

Nazbanou Nozari [a,*], Audrey K. Kittredge [a], Gary S. Dell [a], Myrna F. Schwartz [b]

[a] Beckman Institute, University of Illinois at Urbana-Champaign, 405 N. Matthews Ave., Urbana, IL 61801, USA
[b] Moss Rehabilitation Research Institute, 1200 West Tabor Road, Philadelphia, PA 19141, USA

ARTICLE INFO

Article history:
Received 20 November 2009
revision received 26 July 2010
Available online 1 September 2010

Keywords:
Lexical access
Aphasia
Repetition
Picture naming
Computational models
Case-series
Word frequency

ABSTRACT

This paper investigates the cognitive processes underlying picture naming and auditory word repetition. In the two-step model of lexical access, both the semantic and phonological steps are involved in naming, but the former has no role in repetition. Assuming recognition of the to-be-repeated word, repetition could consist of retrieving the word's output phonemes from the lexicon (the lexical-route model), retrieving the output phonology directly from input phonology (the nonlexical-route model) or employing both routes together (the summation dual-route model). We tested these accounts by comparing the size of the word frequency effect (an index of lexical retrieval) in naming and repetition data from 59 aphasic patients with simulations of naming and repetition models. The magnitude of the frequency effect (and the influence of other lexical variables) was found to be comparable in naming and repetition, and equally large for both the lexical and summation dual-route models. However, only the dual-route model was fully consistent with data from patients, suggesting that nonlexical input is added on top of a fully-utilized lexical route.

Introduction

While language production models differ in many respects, most of them are in agreement about the two-step nature of lexical access (Dell & Reich, 1981; Fromkin, 1971; Garrett, 1975; Levelt, 1989). The first step, the semantic step, involves mapping meaning onto an abstract representation of a word, and the second step, the phonological step, involves mapping this abstract lexical representation onto its phonology.

In this paper, we investigate how two of the most widely-used neuropsychological tests, picture naming and auditory word repetition, sit within the two-step architecture of the lexical access system. Although these two tasks are often utilized in clinical aphasia assessment, the relation between them is much more than a clinical issue. Imitation (repetition of heard words) and production from meaning are two of the key activities that shape the development of lexical skill in infants and even adult second-language learners. How these tasks support one another relates to theoretical issues such as the processing levels in the lexical system, the nature of verbal short-term memory, the form of phonological representations and the extent to which they are shared between comprehension and production, and the role of corrective feedback in acquisition (e.g. Gupta & Tisdale, 2009; Plaut & Kello, 1999).

Picture naming is often viewed as involving both semantic and phonological steps of the production process (Schriefers, Meyer, & Levelt, 1990). Word repetition, on the other hand, might be accomplished by direct mapping of input phonology (sounds of the heard word) to output phonology, without the involvement of either of the two steps. Or, it could take advantage of the phonological step, so that the heard word would be first mapped onto its abstract representation and then its output phonology would be retrieved via the phonological step. If the latter case is true, picture naming and word repetition share the second step of lexical retrieval.

* Corresponding author. Fax: +1 217 333 2922.
E-mail address: nnozari2@illinois.edu (N. Nozari).

0749-596X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2010.08.001

The idea that word repetition *could be* accomplished through the second step of naming is not new. This proposal has been illustrated in box-and-arrow models of lexical processing (e.g. Hanley, Kay, & Edwards, 2002; Hillis & Caramazza, 1991) and has even been simulated in computational models of lexical access in production (e.g. Dell, Martin, & Schwartz, 2007; Martin, Dell, Saffran, & Schwartz, 1994). At this point, most researchers agree that this second step plays some role, but there is no consensus on what that role is (e.g. Baron, Hanley, Dell, & Kay, 2008; Ruml, Caramazza, Capasso, & Miceli, 2005). The research presented in this article combines quantitative measurement of the degree of lexical involvement in aphasic naming and word repetition with a model-driven analysis of what those measurements could mean. As a result, we are able to garner support for a general characterization of the role of the phonological step in naming and repetition, and defend a specific model of both tasks.

The first goal of the paper is empirical. We aim to determine whether picture naming and word repetition are differentially sensitive to indices of lexical retrieval affecting the phonological step of lexical access. To this end, we present actual data from 59 aphasic patients on naming and word repetition tasks, and compare the effect of word frequency on their performance on the two tasks. Testing aphasic as opposed to unimpaired speakers is both theoretically and practically motivated. On the theoretical side, Freud's claim that aphasic errors differ from normal speakers' errors only quantitatively, but not qualitatively, has been empirically tested and supported (e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). This "continuity thesis" of errors justifies the use of aphasic errors in the study of normal processes of language production. On the practical side, normal speakers rarely err in single word repetition. Analysis of speech errors requires a significant number of them, which aphasic speakers readily provide.

Frequency, as an index of lexical retrieval, is well-suited for our goal, because a large number of studies have shown the major locus of the effect to be the phonological step (e.g. Jescheniak & Levelt, 1994), the step that naming and repetition hypothetically share. If in fact the two tasks share this step, one would expect similar effects of frequency in both of them, provided that the analysis can be focused on performance that arises from that step. Although we treat frequency as the principal index of lexical involvement, we also examine the influence of other lexical variables such as phonological neighborhood density and age-of-acquisition.

Once we show that word repetition does, in fact, routinely use the phonological step of the lexical access system, we turn to a second goal: to determine which cognitive architecture is most compatible with this finding. To accomplish this goal, we simulate picture naming and auditory word repetition (the latter, using three different models), analyze the behavior of the models with regard to the frequency effect, and compare and contrast different architectures. Finally, we simulate a sample of "virtual patients", whose data we analyze using the same statistical methods as the ones used for the actual patients.

For repetition, three distinct models are examined: the *lexical-route* model (which describes repetition purely as the phonological step of lexical access), the *nonlexical-route* model (which assumes that repetition is merely mapping input onto output phonology, without the necessary involvement of either the semantic or the phonological step) and the *summation dual-route* model (which combines the activity of both lexical and nonlexical routes). We borrow the term "summation" from Hanley et al. (2002) and Hillis and Caramazza (1991), who proposed that the lexical and nonlexical sources of activation combine to produce the response at the output phonemes. It is instructive to note that the distinction between nonlexical and lexical routes, and the possibility that both are used, are issues that also arise in models of printed-word naming, which is just like word repetition, except that the stimulus is visual. The nonlexical route corresponds to some kind of orthographic to phonological mapping, while the lexical route corresponds to the recognition of the orthographic string as a word and the ensuing contribution of lexical and possibly semantic knowledge to the output. At present, the dominant models of word reading recognize the importance of both mappings, though there is considerable disagreement about how they function and interact with one another (e.g. Coltheart, Curtis, Atkins, & Haller, 1993; Perry, Ziegler, & Zorzi, 2007; Plaut, McClelland, Seidenberg, & Patterson, 1996). Both lexical and sublexical processes also appear to be important in spelling (Rapp, Epstein, & Tainturier, 2002).

If repetition routinely takes advantage of the phonological step of the lexical access system, the analysis of the nonlexical-route model's output should fail to match the pattern observed in patients' data, while the lexical-route model should fit it well. It is uncertain whether the summation dual-route model's behavior will pattern more like the lexical-route or the nonlexical-route models, and hence the simulation work with this model and its degree of match to the data will be particularly important.

Since the two-step nature of lexical access is the foundation of this study, we first review the evidence in favor of the two-step process and introduce the interactive two-step model of lexical access (Dell & O'Seaghdha, 1992), used as the naming model for our simulations, from which the different variations of the repetition models are derived. Next, we explain our focus on a specific type of speech error (nonwords), and discuss in greater detail why word frequency is a useful index for studying the architecture of the lexical access system. Finally, we present our analysis of aphasic patients' data followed by the simulations, and conclude by comparing these results.

The two-step model of lexical access

Why two steps? To produce a word, its meaning needs to be translated into sound. However, because of the unsystematic nature of such a mapping (similar meanings are not necessarily described by similar sounds), to establish the connection one needs an intermediate layer or some other mechanism that nonlinearly transforms the outcome of the initial mapping from semantics.

There is also empirical evidence endorsing the two-step process of lexical retrieval. For example, the tip of the tongue phenomenon suggests that one might get stuck in the middle of the process (Meyer & Bock, 1992). However, the most solid evidence in support of a two-step process comes from the study of speech errors (Fromkin, 1971; Garrett, 1975). A two-step process can explain why people sometimes say DOG when they mean to say CAT, replacing a complete word with another (a *lexical error*). The concepts DOG and CAT share semantic features, which are represented in the first step of the process. It also explains why they sometimes say ZAT instead of CAT, replacing part of the word (a *sublexical error*). In the process of phonological encoding during the second step, the onset /k/ is replaced by the onset /z/, and the nonword ZAT is created.

Apart from the dual nature of errors (lexical vs. sublexical), which ostensibly points to two steps of processing, there is yet another argument in favor of a two-step structure. Lexical errors tend to be of the same grammatical class as the target (e.g. target utterance: *does it SOUND different?* error: *does it HEAR different?* where the verb "sound" is replaced by a lexical error of the same grammatical category; Ferreira & Humphreys, 2001; Garrett, 1975; MacKay, 1982), while sublexical errors do not show such a tendency (e.g. target utterance: BLACK BOX; error: BLACK BLOX, where a consonant from an adjective moves to a noun). This dissociation in sensitivity to grammatical class shows that some stage of the process must bear grammatical information, but that such information is not important at another stage of processing (Garrett, 1975).

The idea of two steps in lexical retrieval has formed the backbone of a number of computationally implemented models (Cutting & Ferreira, 1999; Dell, 1986; Dell & O'Seaghdha, 1992; Levelt, Roelofs, & Meyer, 1999; Rapp & Goldrick, 2000). Among these, the interactive two-step model is unique in that it predicts response categories in picture naming and word repetition (Dell et al., 1997). For a given set of parameters, the model predicts a certain percentage of responses in the correct category, as well as in different error categories. As argued above, errors can be powerful indicators of the involvement of specific steps of lexical retrieval, so such a model is suitable for our study.

The structure of the interactive two-step model is shown in Fig. 1. This model is a localist connectionist model constructed to *name* objects whose name is a consonant–vowel–consonant (CVC) syllable. There are three layers in this model: the *semantic* layer, the *word* layer and the *phoneme* layer. The semantic layer comprises nodes representing semantic features of objects or concepts (e.g. for the concept CAT, the semantic features would be animal, feline, furry, meows, etc.). The word layer consists of nodes that represent *lemmas* (Kempen & Huijbers, 1983), abstract forms of words that link to grammatical information. And finally, the phoneme layer has nodes representing phonemes, organized into three clusters: the onset cluster, the vowel cluster, and the coda cluster. Each node in the model has an activation level, a number that describes how strongly that node can stimulate other nodes. Nodes in each layer in the model are connected directly to the nodes in the adjacent layers via connections that have certain weights. The weights determine how strongly two nodes are related and how strongly the activation of one node affects the activation of the other.

The model names a picture in two steps: during the first (semantic) step, the semantic nodes corresponding to the to-be-named object's meaning are activated. The activation spreads down to the word nodes, each receiving node's activation being determined by a linear activation function with decay. From here activation spreads further down, but also flows backwards to the upper level. The two-way spread of activation in the model makes it an interactive model, meaning that although there are two steps of processing, the steps are not completely modular and each step can, and does, influence the other. Evidence for such interactivity exists (e.g. Nozari & Dell, 2009), but interactive accounts of lexical access remain controversial (e.g. Levelt, Roelofs, & Meyer, 1999; Rapp & Goldrick, 2000). The activation flows in the model, with random noise perturbing the activation of nodes at each time step, until a certain number of time steps have passed. At this point, the most active node at the word level is selected and the first step is concluded. The grammatical class of the target word puts a restriction on this selection, so that if the target word has to be a noun, only nouns can be selected at the word level. This property of the model is endorsed by the empirical finding that word errors (word exchanges, substitutions, anticipations, or perseverations) often match the grammatical class of the target word they replace.

The second (phonological) step starts when the node chosen at the word level receives an extra *jolt* of activation, which sends activation through the network again. At the end of the second step, after the preset number of time steps are completed, the most active phoneme in each of the three phoneme clusters (onset, vowel and coda) is chosen; the resulting phoneme combination is considered the model's response.

The model's response can fall into one of six categories: correct response, semantic error, formal error, mixed error,

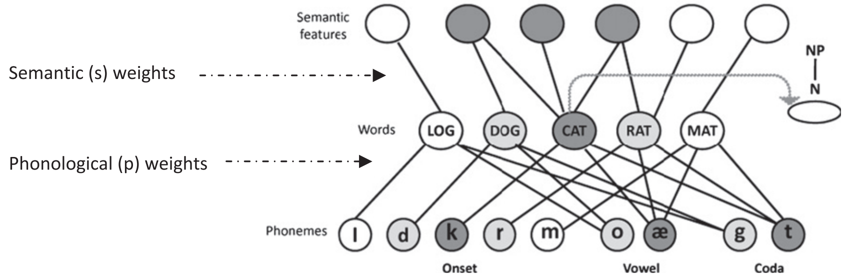

Fig. 1. The interactive two-step model of word production.

unrelated error and nonword error. Semantic errors have their origin in the first step, when confusion arises in the process of choosing between words that have common elements of meaning. Formal errors are related in sound to the target word (e.g. MAT for CAT), and can arise at the second step, when a correctly chosen word goes through erroneous phonological encoding, or at the first step, when another word activated by feedback from the target's phonemes is erroneously chosen (Schwartz, Dell, Martin, Gahl, & Sobel, 2006). Mixed errors are words that are similar to the target in both meaning and phonology (e.g. RAT for CAT) and are, in the model, produced primarily during the first step, when feedback from the phonemes results in the choice of a semantic competitor at the word layer by increasing its activation (Dell & Reich, 1981; Harley, 1984; Martin, Weisberg, & Saffran, 1989; see Rapp & Goldrick (2000) for a mechanism creating mixed errors during the second step). Unrelated errors bear no resemblance to the target, either meaning-wise or sound-wise (e.g. LOG for CAT) and usually occur during the first step. Finally, nonword errors are any combination of phonemes, similar or dissimilar in sound to the target, which do not make a legitimate word in the language implemented in the model (e.g. FAP for CAT). In contrast to all the other errors produced by the model, nonwords necessarily arise during the second rather than the first step of production (or in later post-lexical phonetic processes, e.g. Goldrick & Rapp, 2007). Thus, given our interest in the role of the second step in naming and repetition, we will focus on nonword errors in our simulations and subsequent analyses.

The model's history in simulating aphasia

The interactive two-step model of lexical access was first applied to a single patient with deep dyslexia (Martin et al., 1994). The model's success in simulating this patient's response pattern motivated its further use with groups of aphasic patients studied using the case-series method (Dell et al., 1997, 2007; Schwartz et al., 2006), where it has been applied in two ways: data-fitting and prediction-making. In data-fitting, the aim is to adjust the model's parameters so that it will closely simulate the pattern of data generated by a given patient on a given task. In the newest version of the interactive two-step model (Foygel & Dell, 2000), there are semantic (s) and phonological (p) parameters, corresponding to the weight of the bidirectional connections between semantic and lexical units, and between lexical and phonological units, respectively. The computational case-series approach entails fitting each patient's picture naming data and finding the s and p parameters that best define the performance of that patient. Prediction-making takes advantage of data-fitting. It uses the estimated s and p parameters (recovered from naming) to make predictions about other aspects of the individual patient's performance, in particular, word repetition.

In order to predict repetition from naming, recent studies have assumed some degree of structural overlap between naming and repetition. The predictions of two models have been tested in this way. In the lexical-route model of repetition, the word is assumed to have been accurately recognized (the "perfect recognition assumption"[1]). The word node is then mapped onto output phonology using the phonological step of the two-step model. Thus, repetition errors are considered to be errors of output that occur during the second step of lexical access. The other implemented repetition model was a dual-route model. It was created to explain the behavior of two patients with similar naming abilities, but vastly different repetition abilities (Hanley, Dell, Kay, & Baron, 2004) and was later tested in a computational case-series study (Dell et al., 2007). The implemented dual-route model was of the summation type; activation from both routes converged onto the phoneme nodes (e.g. Hanley et al., 2002; Hillis & Caramazza, 1991).

Whether the lexical or the dual-route model is the correct model of word repetition is controversial. In some studies, it seems that adding a nonlexical route does not help improve the model's accuracy of predicting repetition from naming, but will instead result in an overshoot in the estimation of correct responses (Baron et al., 2008; Dell et al., 2007). However, the dual-route model does provide a good fit for patients whose repetition ability is unexpectedly good, given their naming performance (Abel, Huber, & Dell, 2009; Hanley et al., 2004).

In summary, the combined use of computational modeling and the case-series method has been useful for investigating the relation between picture naming and auditory word repetition. By creating different versions of a repetition model and setting the parameters to the ones recovered by simulating individuals' naming performance, one can look at how each model predicts repetition performance from naming. Those predictions can then be compared to the actual repetition performance of the aphasic patients, and the model that provides the best fit can be said to offer a more realistic picture of the relationship between naming and repetition. However, this approach is limited largely to predicting overall repetition correctness and completely ignores the influence of lexical properties on the process. As a result, the relation between these two tasks remains very much an open question.

In the present paper, we compare the naming model and three versions of repetition models, by exploring how the models vary in their effects of frequency on the probability of commission of nonword errors. This

[1] Earlier models assumed shared phonological input and output units (Martin et al., 1994). This means that the phonological nodes that represented a heard word to be repeated and the phonological nodes that governed spoken output for repetition (as well as naming) are the same. Although this assumption of shared input and output nodes holds up well in some cases (e.g. the patient modeled in Martin et al., 1994), there is evidence that phonological input and output can be dissociated, and, specifically, that intact phonological input processing can coexist with deficient output processing (Martin, 2003; Martin, Lesch, & Bartha, 1999). Therefore, later models (e.g. Dell et al., 1997) instead assume perfect recognition. The *perfect recognition assumption* stipulates that, although there are separate input and output phonological nodes, all patients, regardless of their problems with outputting words, recognize auditorily presented words perfectly, at least to the extent that the correct word node is accessed. Although the perfect recognition assumption does not always hold (Schwartz et al., 2006), patients with deficient input processing can be easily identified and eliminated from tests of the model, as was done in the present study.

comparison speaks directly to the architectural overlap between models of naming and repetition. A similarly-sized frequency effect is expected with full overlap in the step of phonological retrieval (i.e. the lexical-route model) while almost no effect is expected in the model that does not use this step, the nonlexical model. Of particular interest, is the behavior of the summation dual-route model, in which lexical and nonlexical routes jointly contribute to repetition.

Implementing the three models of auditory word repetition

Fig. 2 illustrates the relationship between the naming model and the three repetition models. The lexical-route model simply involves traversing the second step of the naming model. Where a naming trial begins with a jolt of activation to the semantic units, a repetition trial begins with the jolt supplied to the target word. This simulates recognition of the target, in keeping with the perfect recognition assumption.

The nonlexical-route model takes very little advantage of the steps involved in naming. It works simply by mapping input phonology onto output phonology. Since there is no input phonology implemented in the model, this is achieved by sending activation from a "nonlexical node" outside the model to its relevant output phonology in the model, via connections with weights that represent the strength of the nonlexical route. This strength is the model parameter "nl"; when modeling a particular patient, nl is estimated by the accuracy of nonword-repetition (for details, see Hanley et al., 2004). Finally, the summation dual-route model *sums* the contribution of both routes.

Why word frequency?

The effects of lexical variables such as frequency (rate of occurrence in a speech corpus) and imageability (rating of how easily a word generates a mental image) have played a large role in debates about the architecture of word production, because they are considered surefire indices of lexical access. Consistent with this interpretation, frequency and imageability effects have been found in the picture naming of both normal (e.g. Bonin, Barry, Meot, & Chalard, 2004) and aphasic subjects (e.g. Nickels & Howard, 1994). Authors have used the presence of these same effects in repetition to argue that it, too, bears traces of lexical access. Normal subjects show a frequency effect in word list repetition (e.g. Hulme et al., 1997), and aphasic subjects are often more accurate when repeating single words that are high in frequency and imageability (e.g. Jacquemot, Dupoux, & Bachoud-Levi, 2007; Jeffries, Crisp, & Lambon Ralph, 2006; Lallini, Miller, & Howard, 2007; Martin, Saffran, & Dell, 1996). However, in some

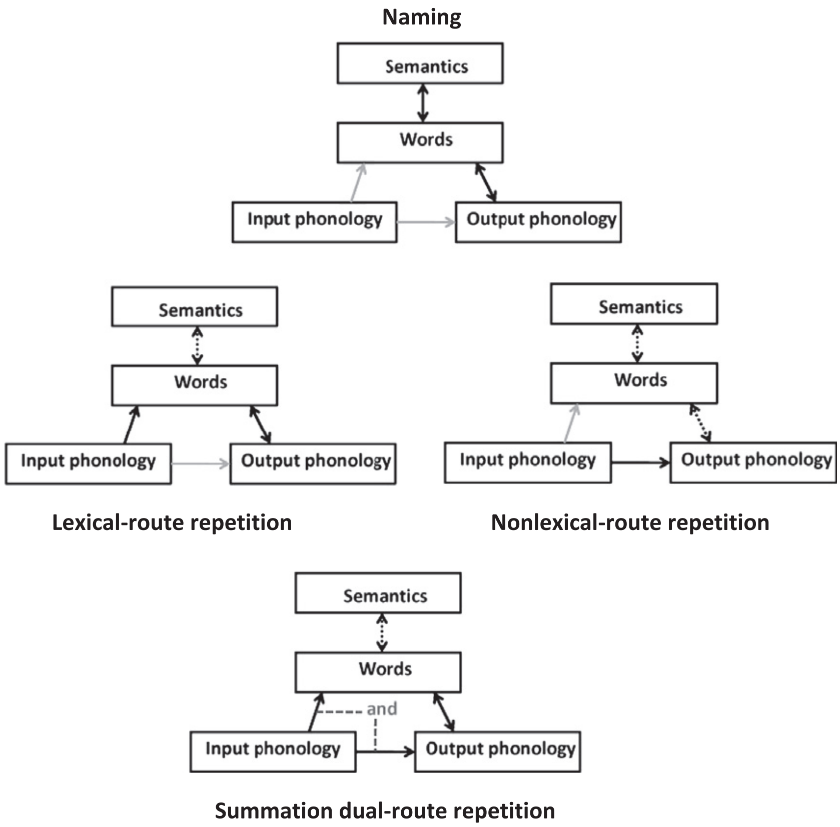

Fig. 2. Relationship between picture naming and three possible accounts of auditory word repetition. Dashed arrows represent involvement through feedback only. Gray arrows represent an inactive route.

of the aphasia studies, the patients' input processing abilities were either poor or not assessed, and thus some of the lexical influence found may be associated with perception, rather than production. Moreover, there are some case studies that report a complete lack of frequency effects in aphasic repetition (e.g. Ackerman & Ellis, 2007; Biederman & Nickels, 2008; Goldrick & Rapp, 2007; Hanley et al., 2002; Wilshire & Fisher, 2004). Positive evidence for a frequency effect in repetition unequivocally demonstrates *some* involvement of lexical information in repetition. Importantly, though, such evidence does not speak to the degree of the lexical route's involvement, and consequently does not clearly delineate the architecture of word repetition in aphasia. In contrast, our investigation directly compares the effect sizes of a lexical variable (frequency) on aphasics' nonword errors in repetition and naming.

Why focus on word frequency, rather than another lexical variable such as imageability? The differences between the hypothesized models of repetition hinge on their reliance on the phonological step of naming. While imageability probably has its effect at higher, semantic levels of processing (Plaut & Shallice, 1993), there is much evidence that frequency is firmly associated with the phonological step (Dell, 1990; Jescheniak & Levelt, 1994). This is not to say that frequency is *only* associated with the phonological step of lexical access. There are frequency effects in normal production tasks that involve the semantic step (e.g. noun phrase production and gender decision; Alario, Costa, & Caramazza, 2002; Navarrete, Basagni, Alario, & Costa, 2006), as well as on the production of semantic speech errors by normal and aphasic subjects (see Kittredge, Dell, Verkuilen, & Schwartz, 2008; Knobel, Finkbeiner, & Caramazza, 2008, for reviews). However, the effect of frequency on the phonological step is stronger and more widely documented, and our own work is consistent with this view of the literature. In a regression analysis of 50 aphasics' picture naming errors, Kittredge et al. (2008) found that higher frequency targets had fewer semantic and phonological errors. However, the effect of frequency on semantic errors proved to be considerably smaller than that on phonological errors, as well as less robust (Kittredge et al., 2008, for a replication).

The fact that the facilitative effect of frequency is so robust on the phonological step of naming indicates that in any production task that originates predominantly at that step, we should see a frequency effect on errors, specifically, nonword errors.[2] Our models of repetition make different predictions about the relationship between nonword errors in naming and repetition. If the repetition of a correctly recognized word is carried out by exactly the same processes as occur during the phonological step of naming (the lexical-route model), the size of the frequency effect on nonword errors in repetition should be comparable to that in naming. In contrast, there should be very little effect of frequency on repetition if it is not lexically influenced (the nonlexical model). As for the summation dual-route model, adding the activation of the two routes together would imply that the frequency effect generated through the lexical route will be present in the dual-route model. However, the presence of the nonlexical route might also be expected to dilute the contribution from the lexical route and consequently decrease the overall frequency effect in the dual-route model. Ultimately, the extent to which any frequency effect is maintained in the dual-route model depends on the details of the model, including its activation function and selection mechanisms. The simulations will thus shed light on the predicted magnitude of the frequency effect in the dual-route model in comparison to the other two models.

[2] Not all nonword errors occur during phonological access; some arise during phonetic processes that occur after the production processes that we are modeling (e.g. Goldrick & Rapp, 2007). The occurrence of phonetic nonword errors would be expected to diminish frequency effects calculated on all nonword errors, but critically, this diminution should be equal in naming and repetition and thus not impact our comparisons.

Patient data

Participants

Data were re-analyzed from 59 of 94 patients who took part in Dell, Schwartz, and colleagues' computational studies of naming and repetition (Dell et al., 2007; Schwartz et al., 2006). They represent a subset of the 65 repetition-study participants, whom Dell et al. (2007) classified as not having phonological input processing deficits based on tests of phoneme discrimination, auditory lexical decision, and picture-name verification with phonological foils (for details, see Dell et al., 2007; Martin, Schwartz, & Kohen, 2005). Limiting the study to these 59 patients increases the chance that the models' assumption that the input word is accurately recognized is a reasonable one. Table 1 presents background information on these 59 patients.

Stimuli

The stimuli were the items of the Philadelphia Naming Test or "PNT" (a set of 175 pictures) and the Philadelphia Repetition Test or "PRT" (spoken versions of the 175 PNT picture names). We were able to obtain complete lexical characterizations for 124 of these items: base 10 log frequency values from the lemma corpus of the online and published CELEX lemma databases (see Baayen, Piepenbrock, & van Rijn, 1993), base 10 logarithms of density values (the number of phonologically similar words or "neighbors") from the online version of the Hoosier Mental Lexicon (see Luce & Pisoni, 1998), ordinal age-of-acquisition (AoA) values (1 = acquired between 8 and 16 months, 2 = acquired between 17 and 30 months, 3 = acquired above 30 months) based on the American version of the MacArthur Communicative Development Inventory (Fenson et al., 1994) and obtained from the online International Picture-Naming Project database (Szekely et al., 2004), rated imageability values from an online version of the MRC Psycholinguistic Database (Wilson, 1988), length in phonemes, and name agreement from a sample of 60 age-matched normal control participants (Dell et al., 1997).

Table 1
Background information on the 59 patients.

Patient code	Sex	Months post CVA	Age	Aphasia category	Naming			Repetition			Naming		Repetition	
					% NW	% Other	% Correct	% NW	% Other	% Correct	% NW to LF words	% NW to HF words	% NW to LF words	% NW to HF words
BAL	F	15	77	Anomic	1	54	45	1	4	95	27	19	5	0
BAT	M	1.5	34	Anomic	8	77	15	0	2	98	2	0	0	2
BBC	F	2	78	Wernicke	42	30	28	42	18	40	63	50	50	29
BI	M	7	63	Wernicke	25	67	8	23	19	59	24	8	11	8
BQ	M	106	24	Broca	2	43	55	4	0	96	8	8	0	0
BT	M	9	61	Broca	20	69	11	0	2	98	44	40	55	29
BW	M	4	54	Broca	0	5	95	4	4	92	23	27	27	18
CAC	F	2	43	Anomic	16	16	68	4	6	90	11	6	2	0
CE	F	12	59	Broca	12	60	27	2	0	98	47	31	39	10
CK	F	4	79	Anomic	9	14	77	7	6	86	34	27	44	24
CT	F	71	40	Broca	1	18	81	6	2	92	3	2	5	3
DAN	F	11	73	Conduction	32	44	23	4	9	87	19	21	0	0
DD	M	18	54	Broca	6	53	40	12	7	81	13	3	11	0
DN	M	1	47	Anomic	0	45	55	0	0	100	0	0	6	2
DX	M	4	75	Conduction	17	19	65	12	15	73	26	6	3	5
EAC	F	19	47	Anomic	5	41	54	4	6	90	2	0	3	2
EBC	M	120	62	Anomic	0	16	84	0	5	95	19	10	6	3
EC	F	76	47	Broca	5	41	54	1	2	97	16	11	19	8
EE	F	93	55	Broca	2	32	66	1	1	98	2	0	0	0
EL	F	51	38	Broca	6	8	85	7	3	90	2	0	2	0
FAG	M	31	72	Broca	21	35	44	3	1	96	3	0	3	0
FAH	M	8	60	Conduction	15	82	3	10	8	82	15	3	23	10
FAM	F	9	40	Broca	1	15	85	1	3	96	13	11	3	2
FBH	M	5	69	Wernicke	25	26	49	19	26	56	39	15	32	10
FG	F	8	68	Conduction	10	15	75	6	5	89	8	5	5	2
FM	F	14	57	Conduction	23	35	42	2	6	92	0	2	3	5
FT	F	4	51	Conduction	56	30	14	40	15	45	13	5	11	3
HN	M	74	53	Conduction	16	51	33	10	16	74	0	2	8	5
HO	M	41	55	Conduction	9	10	81	1	2	98	34	31	5	3
KAC	M	2	60	Conduction	39	20	41	24	19	57	5	0	3	0
KAM	M	2	61	Wernicke	31	65	4	34	30	36	11	2	18	6
KAX	M	111	39	Conduction	8	16	76	6	2	93	0	0	0	0
KCC	M	82	65	Anomic	1	4	95	2	17	81	6	2	3	0
KD	F	6	49	Conduction	15	7	78	5	4	91	0	0	0	0
KK	M	1	58	Wernicke	14	77	10	14	13	73	3	3	5	3
MBC	F	12	66	Anomic	1	6	93	0	3	97	10	2	24	10
MD	M	28	48	Broca	1	8	91	1	2	98	2	0	0	0
MG	M	51	81	Anomic	2	28	70	2	4	94	0	2	2	2
MO	M	70	54	Broca	9	13	78	16	2	82	18	16	18	6
MQ	F	21	41	Conduction	27	15	58	21	8	71	5	5	6	2
MX	F	43	35	Broca	6	16	77	3	2	94	0	0	0	0
NAC	F	4	74	Trans-cortical sensory	1	26	73	4	10	85	2	0	3	3
ND	F	39	76	Broca	2	36	61	2	1	98	3	2	0	0
NH	M	19	59	Broca	4	8	88	2	3	95	0	0	2	3
NN	M	8	48	Anomic	0	4	96	0	0	100	5	5	2	0
NU	M	4	40	Anomic	3	4	93	4	3	93	3	0	2	0
NX	F	20	54	Conduction	6	41	53	17	4	79	11	2	13	2
OE	M	9	59	Anomic	1	21	78	0	6	94	6	2	2	0
SAM	M	55	76	Anomic	1	17	82	2	6	93	13	2	0	0
SI	M	39	61	Anomic	1	53	46	3	6	90	0	0	0	2
SL	M	16	48	Anomic	2	11	86	0	2	98	15	13	6	2
SM	M	29	72	Anomic	0	16	84	2	15	82	21	21	5	2
SS	M	33	57	Trans-cortical sensory	4	83	13	1	1	98	15	15	10	10
ST	M	4	78	Conduction	7	5	88	0	6	94	2	0	0	2
TAT	M	10	67	Anomic	0	10	90	1	0	99	40	10	29	8
TE	F	38	54	Broca	14	38	48	4	2	94	8	13	2	2
TG	F	5	66	Anomic	10	19	70	2	3	95	16	11	6	2
TT	M	36	54	Broca	14	39	48	4	3	93	16	5	10	3
XD	M	86	54	Broca	21	71	8	6	5	90	19	23	8	3

Procedure

The PNT pictures were presented one at a time on a computer for the patient to name. Trials ended when the patient responded or after 30 s if there was no response. At trial termination, the experimenter said the target name. Only the first complete response produced on each trial was scored, and only exact matches to the target were counted as correct except when patients had clinically obvious articulatory-motor impairments (for further details see Schwartz et al., 2006).

Stimuli for the PRT were recorded onto audiotape at a rate of 1 per 5 s. The participant's task was to repeat each word immediately after hearing it. In rare cases when subjects were still responding towards the end of the interval, the tape was stopped to allow the response to be completed. Requests for repetition of the stimulus were denied. All responses were tape-recorded. During the test, the examiner, an experienced speech-language pathologist, transcribed the responses. The audiotapes were then transcribed by a research assistant. Any discrepancies between the two transcriptions were resolved jointly by the two transcribers.

For the purposes of this paper, each patient's response to a given picture (or word in the case of repetition) was classified as either *correct*, a *nonword* error (including both responses phonologically similar to the target and those not similar to the target), or an *other* error that did not fall into either of these two categories (e.g. a semantic error such as saying "pie", or an omission error such as "I don't know", in response to a picture of a cake).

Statistical analysis

In order to compare the frequency effects in naming and repetition, we must first establish that frequency affects performance on each of the tasks. On average, when naming the pictures, patients made 8.39% nonword responses when the target was high-frequency (in the upper half of the 124 test items) and 12.93% nonword responses when the target was low-frequency (in the lower half of the items). The same pattern was observed in auditory word repetition: patients made on average 4.26% nonword errors when the to-be-repeated word was high-frequency and 9.49% nonword errors when the target was low-frequency (see Table 1 for each patient's performance). It is evident in both tasks that as the target frequency becomes lower the probability of making a nonword error increases. However, frequency is correlated with many other lexical variables. Are there frequency effects when these other variables are taken into account, and do these effects differ for naming and repetition?

To address this question, we performed binomial, hierarchical, multiple logistic regression[3] using the statistical *R* package. The goal was to compare the degree to which high-frequency picture names reduced the likelihood of making nonword errors in naming and repetition. The log odds of making a nonword error relative to any other type of response was predicted from the target's mean-centered log frequency, log density, AoA, imageability, length, name agreement, and item order. To compare performance in naming and repetition, we included a predictor coding for the type of task (which took on a value of 0 in naming and 1 in repetition), as well as interactions of task with each mean-centered lexical variable. Table 2 presents the descriptive statistics for the predictors in the model. Log density and length have a correlation of −.84, but diagnostic measures (calculating variance inflation factors and obtaining similar log frequency effects with models that exclude either log density or length) suggest that multicollinearity is not present in our dataset.

We included crossed random effects (see e.g. Quené & van den Bergh, 2008) in the model, to model the correlation among responses made by the same subjects to the same items. We also included random slopes for log density, age-of-acquisition, task, and the interaction of log density and task, because the effect of these predictors varied significantly by subject. We assessed the need for randomly varying fixed effects by subject in the following manner: random slopes for each predictor were added to the model one by one, in a fixed order. We dropped the random slope term if the model appeared to be overparameterized (i.e. if there was a correlation above .95 between any two random effects). Otherwise, we compared the model with the newly added term to the preceding model by performing a chi-square test of the deviance in model log likelihoods. If this was not significant, we dropped the term from the model and continued adding other random slopes. We also tested the need for a randomly varying fixed effect of mean-centered order by items in the same way. Only random slopes whose addition significantly decreased the model's log likelihood in this way were included in the final model.

Regardless of which repetition model is correct, we should expect to see an effect of frequency (and other lexical variables; see Kittredge et al., 2008) on picture naming, since this is a lexically influenced process. Thus, we would expect a negative-going coefficient for the effect of frequency on naming (that is, a one-unit increase in log frequency should predict a significant decrease in the log odds of producing a nonword). Importantly, the key result for testing different repetition models against each other is the coefficient estimating the interaction between frequency and task, i.e. the change to the

Table 2
Descriptive statistics for predictors in the regression models.

Variable	Mean	Range
Item order	88.56	1–175
Log frequency	1.38	0–3.21
Log phonological density	.85	0–1.60
Age-of-acquisition	2.02	1–3
Imageability	591.89	369–644
Length in phonemes	4.27	2–10
Name agreement	.97	.82–1.00

[3] We code the response as binomial (nonword vs. any other response) because we are only theoretically interested in nonword errors. However, since complete coding of the responses recognizes more than two types, we conducted an additional multinomial logistic regression analysis and obtained the same pattern of results.

frequency effect in repetition. If repetition is a lexically influenced process, the effect of frequency should be comparable to that in naming. In that case, the coefficient for the interaction should not be significantly different from zero. If repetition is not lexically influenced, the effect of frequency should be smaller in repetition, and hence the change to the frequency effect in repetition should be opposite in direction to the effect in naming. In this case, the coefficient should be positive-going and significantly different from zero.

Results

We found that pictures with high-frequency names were significantly less likely to elicit nonword errors than low-frequency picture names (coefficient = −.272, standard error = .109, p = .013). Expressing this coefficient as an odds ratio provides a more intuitive estimate of the effect size: with every 10-fold increase in the frequency of the target, the odds of making a nonword error relative to any other response decreased by about a quarter (odds ratio = .762). This effect emerged even after effects of other correlated lexical variables and the order of item presentation were accounted for, and replicates other effects of frequency on form-related picture naming errors in the literature[4] (e.g. Gordon, 2002). In addition, the subsidiary findings that nonword errors were less likely to occur for target words from phonologically dense neighborhoods and early acquired words are consistent with similar effects on phonological errors found in previous work (e.g. Gordon, 2002; Kittredge et al., 2008).

Most importantly, the regression analysis revealed that the frequency effect in repetition is comparable to the frequency effect in naming. That is, there was no statistically significant change in the log odds of making a nonword error in repetition, as compared to naming, due to a one-unit increase in log frequency (coefficient = −.167, standard error = .148, p = .258). Adding this coefficient to that of the frequency effect in naming will give us an estimate of the size of the frequency effect in repetition (−.439) which can be directly compared to the size of the effect in naming (−.272). Recall that the difference is not statistically significant. To verify that there is an effect in repetition, independently of the naming data, we ran a similar hierarchical logistic regression on the PRT data alone (predicting only repetition from the seven lexical variables and including random intercepts for subject and item, as well as random slopes for log density and AoA). This supplementary analysis found a significant frequency effect in repetition (−.417, standard error = .141, p = .003), confirming the results of the larger, comparative regression. Also in confirmation of the main analysis, AoA, length, and phonological density affected nonword-error probability in the expected directions, although the density effect was only significant when no random slopes for that variable were included in the model.

The comparable frequency effects in naming and repetition emerged despite a baseline difference in the rate of nonword errors between the two tasks (coefficient = −.572, standard error = .174, p = .001). This shows that the log odds of making a nonword error were less in repetition than in naming, an important finding that we return to in the general discussion. It is also noteworthy that the other lexical factors that significantly affected the production of nonword errors (density, AoA, and length) did not interact with task either; their influence was the same in naming and repetition (as can be seen in Table 3, the corresponding coefficients are not significantly different from zero). These findings all point to a repetition process that is as lexically influenced as naming.

Discussion

Our regression analysis shows that the frequency effect found in picture naming is present in auditory word repetition to a similar degree. The comparable magnitudes of the effects in naming and repetition point to a high degree of overlap between the two tasks within the architecture of lexical access. However, one possibility needs to be ruled out before we draw conclusions about repetition models. It is conceivable that the patient sample, for whatever reason, is comprised mostly of individuals who lack functioning nonlexical routes. In that case, showing that such patients primarily use their lexical route for word repetition would not be especially informative.

To address the issue of the efficacy of the nonlexical route in the sample, we repeated the main regression analysis including a measure of the strength of each patient's

Table 3
Results of regressions on patients' nonword error data. * Indicates significance with $p < .05$, ** indicates significance with $p < .001$.

Description	Estimate[a] (standard error)
Effects of other lexical variables	
Log density effect in naming	−.594 (.226)*
AoA effect in naming	.266 (.068)**
Imageability effect in naming	.003 (.002)
Length effect in naming	.174 (.064)*
Name agreement effect in naming	−.360 (1.264)
Change to effects of other lexical variables in repetition	
Change to log density effect in repetition	.216 (.275)
Change to AoA effect in repetition	.067 (.083)
Change to imageability effect in repetition	−.0001 (.002)
Change to length effect in repetition	.023 (.077)
Change to name agreement effect in repetition	3.317 (1.735)
Item order effect in naming	−.0003 (.001)
Change to item order effect in repetition	.001 (.002)
Intercept	−3.174 (.230)**
Random effect of items	.117
Random effect of subjects	2.734
Random slope for log density	.629
Random slope for AoA	.036
Random slope for task	1.125
Random slope for log density*task	.711

[a] For the random error term, the number represents the variance of the random error term's distribution.

[4] The size of the corresponding multinomial regression coefficient (−.478) is comparable to that found by Kittredge et al. (2008) in a similar multinomial regression analysis of nonword errors in aphasic picture naming (−.484). This analysis performed by Kittredge et al. included 34 (58%) of the aphasics who participated in the present study.

nonlexical route. The ability to repeat nonwords is the classic index of the nonlexical route (Hanley et al., 2002). Consequently, we included a nonword-repetition score for each patient as an additional predictor. This score was available for 30 of the 59 patients and was collected using a task comprising 60 phonologically legal strings, which were derived from concrete one- and two-syllable (average of 1.37 syllables) words by changing both a vowel and a consonant. For this sample, the mean repetition accuracy was 65% with a range of 27–92%, thus demonstrating that the group as a whole has some degree of functioning for the nonlexical route, but also that this functioning is quite variable.

The key question in this analysis is whether nonword-repetition scores modulate frequency effects, particularly for word repetition. If good nonword-repetition causes an individual to leave out the lexical route (i.e. the patient repeats by the pure nonlexical-route model), then frequency effects should be diminished in word repetition relative to naming as nonword-repetition scores increase, that is, there should be a triple interaction of task, frequency and nonword-repetition. If, on the other hand, patients' word repetition is consistently lexically influenced, frequency effects should not be modulated by nonword-repetition score and task.

The three-way interaction between frequency, task and nonword-repetition score was essentially nonexistent in the data (coefficient = .115, $p = .902$). The analysis did, however, replicate the overall frequency effect (coefficient = $-.328$, $p = .017$) and the fact that this did not interact with task (coefficient = $-.045$, $p = .828$). The absence of the three-way interaction thus suggests that nonword-repetition ability (the strength of the nonlexical route) does not change the effect of frequency on word repetition. This null result is unlikely to be due to lack of power or lack of validity of the nonword-repetition measure, because the measure was potent and valid in another way: nonword-repetition strongly interacted with task (coefficient = -2.487, $p = .002$). Specifically, the tendency for good nonword-repetition ability to prevent nonword errors to lexical stimuli was significantly stronger in word repetition than in naming. In short, good nonword-repetition implies good word repetition rather than good naming, but good nonword-repetition is irrelevant to frequency effects in both tasks. We will return to these findings in the general discussion as they are highly relevant to the contrast between the lexical-route and the summation dual-route models.

So far, the similarity of the frequency effects for naming and repetition in our patient data supports the lexical-route model of word repetition, and goes against the nonlexical model, or any model in which the nonlexical route, operating alone, is used for a large fraction of repetition trials. In the next section, we confirm these assertions with model simulations. But recall that the summation dual-route model also predicts an effect of frequency to the extent that the lexical route is used. Without knowing that extent, however, it is unclear whether the present results could be accommodated within the framework of the summation dual-route model. We answer this question as well with simulations.

Model simulations

Exploratory simulations

Two sets of simulations were run, one exploring the frequency effect and the other the proportion of nonword errors generated by each model. In each simulation set, the dependent variable is explored in a square space of *s* (semantic) and *p* (phonological) weights, where the particular values tested are .01, .02, and .03 for each parameter. The choice of this parameter range was motivated by the fact that most of the aphasic sample's *p*-weights, which are primarily responsible for nonword errors, are within or close to this interval. Moreover, nonword-error probability begins to approach the model's floor or ceiling when larger or smaller *p*-weights are used with some of the model types. The *s*-weights were varied across the same range of aphasic values, largely to demonstrate that they have little effect on the model's nonwords in comparison to the *p*-weights. The model's nonword error probabilities as a function of manipulations of weights are smooth and monotonic and so there is no need for more fine-grained testing (Foygel & Dell, 2000).

Each naming and repetition model was run through 1000 simulated trials. Half of the trials used high-frequency targets and the other half low-frequency targets. Because empirical studies localize the bulk of the frequency effect in naming to the phonological step of lexical access, we implemented frequency in the models' *p*-weights. A high-frequency target kept the baseline weights (the chosen *s* and *p* parameters), and a low-frequency target's *p*-weights were reduced by multiplying them by .8, while keeping the rest of the weights in the model at baseline values. Other parameters, such as the decay rate (.6), were held constant across simulations.

The naming process started with the semantic features of the target word receiving a jolt of activation (10 units of activation to each of the three semantic features connected to the target word). The activation was then allowed to spread in the model for eight time steps, after which the most active node in the word layer was selected, thus completing word access. Then, phonological access commenced. The selected word received an extra jolt of 100 units, and activation continued to spread for another eight time steps. At this point the most active phoneme in each of the onset, vowel and coda clusters was identified and the combination of those phonemes defined the model's final response. The activation of nodes during each time step was subject to random noise as described in Dell et al. (1997).

Repetition was simulated using three repetition models: the lexical-route repetition model was simply the second step of the naming model with no additional parameters. The process started by giving the jolt of activation (100 units) to the target word instead of the semantic features, eliminating the chance of first-step errors. The activation then spread in the model over the course of eight steps and the final response was determined in the same fashion as for the naming model.

In the simulated lexical route, the perfect recognition assumption is in force: the modeling assumes that the word to be repeated is correctly recognized and gets its full jolt of activation. This is clearly not true for all patients, but the assumption is, nonetheless, a reasonable one for many. As stated above, our sample was selected to have minimal input processing problems (e.g. good lexical decision, phoneme discrimination). Even granting this, there are several ways in which the perfect recognition assumption might cause problems for the analysis. First, perhaps some patients, despite meeting our inclusion criteria, do have subtler input processing deficits that cause them to mishear the word stimuli, particularly low-frequency words, thus introducing a frequency effect at the word recognition step of repetition. An analysis in Dell et al. (2007) suggests that even if this occurred, it would not influence the current analysis: They compared the repetition errors of patients with and without input processing deficits. For patients without input processing deficits (i.e. the present sample), nonword error rate was nearly exactly predicted from such errors in naming (i.e. the *p*-weight in modeling terms), suggesting that these errors arose during output. The same was true for the patients who had input processing deficits. An input processing deficit, however, mattered with regard to *formal* errors (phonologically related word errors). These were more likely than expected for the patients with an input processing deficit. It appears that poor input processing causes aphasics to mishear a stimulus word as another word and to repeat that wrong word. Thus, input processing deficits create an excess of formal errors, not nonwords. Because our analysis focuses on nonwords, there is unlikely to be a problem. Perhaps, however, somehow poor input processing does lead to poorer representations of correctly recognized words that then somehow increase the chance of nonword-repetition errors. All that we can say about this is that it was not observed in the study of Dell et al.'s patients with poor input processing. The patients did not have an unexpected excess of nonword errors.

For the nonlexical-route and the summation dual-route models, an additional parameter (nonlexical or "nl"), the strength of the nonlexical pathway mapping input to output phonology, was defined. This parameter determines the weight of a bidirectional connection between a node that represents heard input and the target phonemes, a path that represents the nonlexical route of repetition (see Hanley et al., 2004, for details). Two values of nl were tested for each nonlexical and dual-route model. .01 represent a relatively weak nonlexical route, while .02 implements a stronger route.

In the nonlexical-route model, repetition started by giving a jolt of activation (100 units) to the external node, with the activation spreading to the target phonemes and flowing freely in the model for eight steps, at the end of which the most activated phonemes in each cluster defined the model's response. Although we expected that the nonlexical model would not show much of a frequency effect, and hence fail to match the patient data, running the model is necessary because phonological–lexical interaction in this model could, in principle, allow for a nonlexical route to provide some lexical effect.

Finally, the summation dual-route model fully combined the activity of both the lexical and the nonlexical routes. In keeping with all other simulations of this model (Dell et al., 2007; Hanley et al., 2004) and the theoretical notion of summation (Hillis & Caramazza, 1991), the target word and the external node each received a 100-unit jolt of activation and the model's response was generated in the same number of time steps as the other repetition models.

Frequency effect maps

For each model, the frequency effect was calculated by taking the difference between the logit transform of the probability of making a nonword error $\left(\ln\left(\frac{p}{1-p}\right)\right)$ when the target was low-frequency and that when the target was high-frequency. This measure would be proportional to a logistic regression coefficient. Note, however, that our subtraction makes for positive values when frequency has its expected effect. The contour plots in Fig. 3 show the frequency effect as a function of *s* and *p*-weights, where darker represents higher values. The scale minimum was just below 0 because, due to chance factors, negative effects are possible, and the maximum was set to the largest empirically obtained frequency effect on phonological errors, that we know of: 8.5 times more likely on low, than high-frequency targets (Dell, 1990). Note that the frequency effect is always present in naming, as well as in the lexical-route and the two dual-route repetition models (nl = .01 and nl = .02). More importantly, the magnitude and distribution of the effect across parameters is similar for naming and lexical and dual-route repetition. The mean effects, given above each figure panel, are quite close, and the distributions are such that there is a tendency for the effect to increase with *p*-weights for all the models. The correlation between frequency effect and *p*-weight was .78, .74, and .79 for the naming, lexical-route, and (pooled) dual-route models, respectively, while there was no reliable effect of *s*-weights for any model. This is to be expected, given that frequency is a proportional manipulation on *p*-weights. The nonlexical-route model's frequency distribution stands in sharp contrast to these other models; its effect is near zero across the tested parameter space.

Our finding that the lexical-route model of repetition shows a comparable frequency effect to the naming model was hardly surprising, since the effect was calculated on nonword errors, which inhabit the step shared between naming and repetition through the lexical route. However, the fact that the summation dual-route shows frequency effects of the similar size to these two models is striking. Note, particularly, that the size of the effect does not change when the nonlexical route is stronger. This last finding is, at first blush, at odds with intuitions about dual-route models, where the expectation is that the behavior of the overall model should reflect some sort of an average of its component routes. In general, how a dual-route model behaves depends on how the lexical and the nonlexical routes interact. There are two possibilities. The first is that the two routes divide the labor. An example of division of labor is what we call the independent dual-route model. On a particular trial, a word is repeated by either the lexical route or the nonlexical-route, but not both. Thus, the predictions of this dual-route

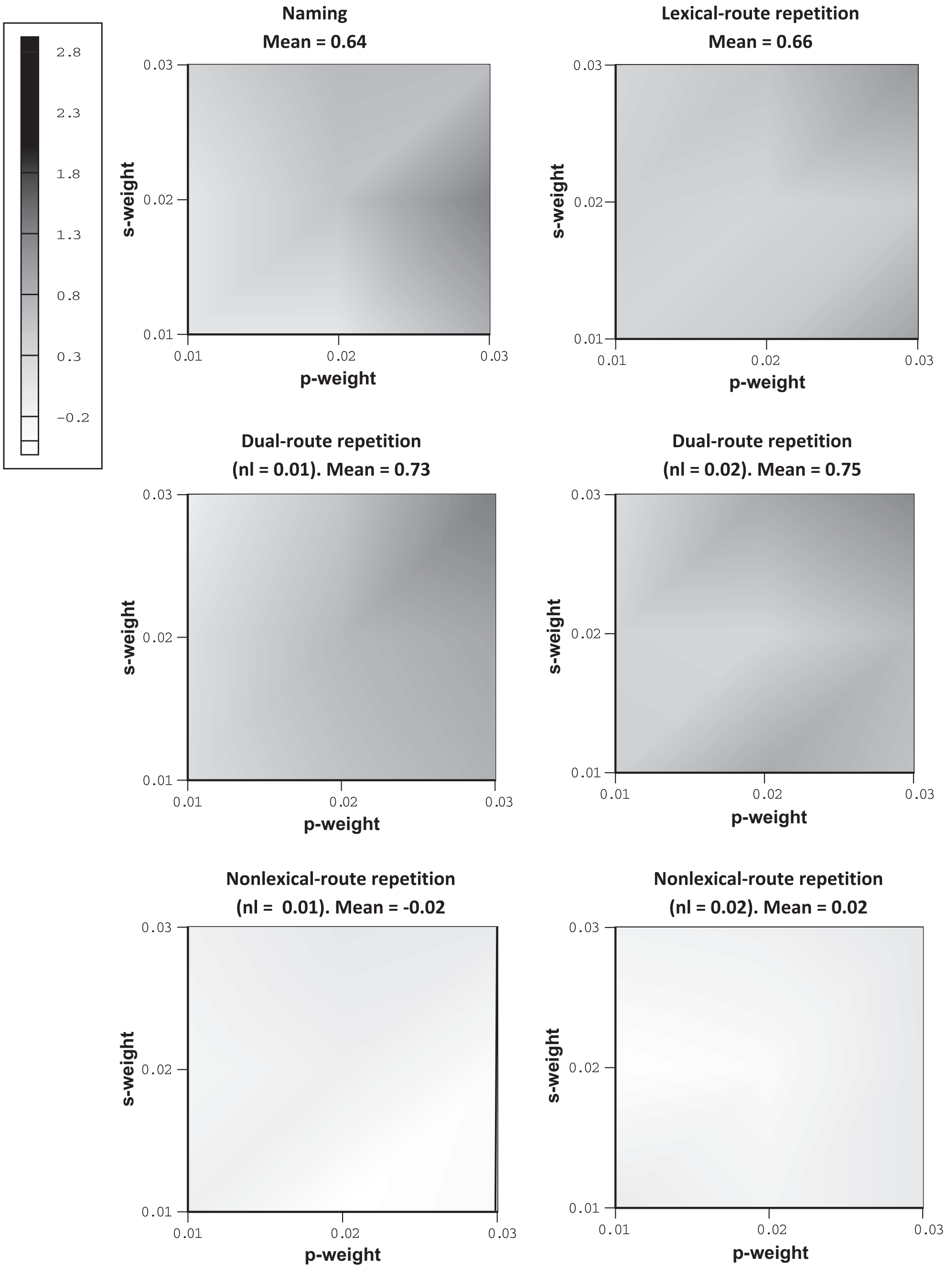

Fig. 3. Frequency effects in naming and three repetition models in the *s*–*p* parameter map. The dual-route and nonlexical-route models have each been simulated using two different strengths of the nonlexical route (nl = .01 and nl = .02). White represents low and black represents high values. Note that the frequency effect is sizeable and similar in naming, lexical-route and dual-route models, while the nonlexical-route model shows insignificant effect.

model would be a weighted average of the lexical-route and nonlexical-route models. If, say, each is used equally often, the frequency effect for repetition would then be half the size of that for naming. A more sophisticated division of labor is possible as well, in which both routes may contribute, but in such a manner that neither makes a full contribution and, the stronger one route is, the weaker the other is. Again, this predicts a decreased frequency effect in repetition relative to naming, as the lexical route becomes weaker. The other possibility is that the two routes combine without any loss. When the nonlexical route is added, the lexical route contributes to the same degree as it would in the absence of the nonlexical route. Thus, any lexically-mediated difference in activation due to frequency is not offset by a contribution from the nonlexical route and, hence, there is no difference in the frequency effect

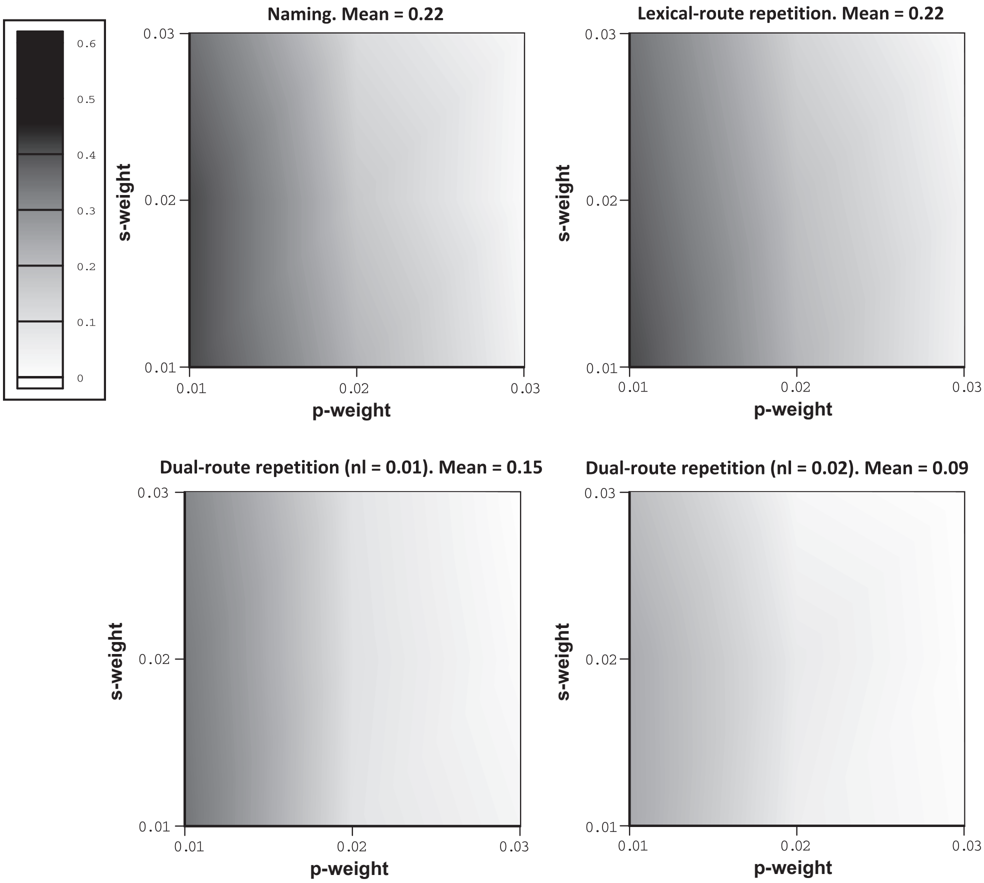

Fig. 4. Probability of nonword generation by naming and three repetition models in the *s–p* parameter map. White represents low and black represents high values. Note that as the contribution of the nonlexical route becomes stronger, the probability of nonword error generation decreases.

between the dual-route and lexical-route models. This is what we observed with the model simulations. Of course, one must recognize that if the contribution of the nonlexical route is sufficiently powerful to remove *all* nonword errors, then, the dual-route model loses its frequency effect on nonword-error probability. However, we did not see any diminution of the effect because we used the logit transformation, which eliminates the natural nonlinearity of the dependent variable as nonwords become rare, and because we used parameter values that kept nonwords off the floor. A similar point was made for a parallel distributed processing model of impaired word reading (Dilkina, McClelland, & Plaut, 2008). Although word frequency interacts with regularity of spelling if the dependent variable is percentage correct, the interaction goes away if the analysis is logistic. This demonstrates that, in the reading model, frequency and regularity contribute in an additive way. Our finding is analogous. The frequency effect in logistic regression does not interact with the presence or absence of an additional source of activation (in our case, the nonlexical route).

The similarly-sized frequency effect in the lexical and summation dual-route models implies that the lexical route is actively and fully used in both models. So, where do the models differ? To answer this question, we map out the likelihood of nonword errors, instead of the frequency effect for these models.

Nonword maps

Fig. 4 shows the probability of making a nonword error as a function of *s* and *p* parameters in naming, as well as in repetition, modeled using the lexical-route and the dual-route models (nl = .01 and nl = .02). The probability of nonwords is graphed only for the high-frequency targets. Overall, the maps show the expected sensitivity to *p*-weight, with nonword probability dramatically decreasing as *p*-weight increases. The important contrast, though, is between the models. Naming and lexical-route repetition, as expected, have essentially identical distributions of nonwords. The dual-route differs dramatically. As the contribution of the nonlexical route becomes stronger, the probability of making a nonword error decreases. The mean nonword probability is .22 when there is no nonlexical route. It drops to .15 in the dual-route model with nl of .01, and even further to .09 when the strength of the nonlexical route is .02. In short, although the dual-route model has the same frequency effect as the lexical-route model, it

Table 4
Proportions of nonword responses of the virtual patients simulated by the naming, lexical-route, dual-route and nonlexical-route models. Responses to high-frequency words are contrasted with those to low-frequency words with frequency implemented in the *p*-weights.* Indicates parameters from actual patients.

Parameter values			Frequency manipulation	Naming	Repetition		
s	*p*	nl			Lexical-route	Dual-route	Nonlexical-route
.019*	.009*	.026*	High frequency	.552	.566	.174	.474
			Low *p*-weight frequency	.642	.618	.234	.498
.009	.019	.026	High frequency	.214	.196	.054	.472
			Low *p*-weight frequency	.31	.336	.108	.468
.032*	.008*	.028*	High frequency	.52	.53	.14	.494
			Low *p*-weight frequency	.59	.584	.202	.464
.022	.018	.028	High frequency	.196	.198	.036	.448
			Low *p*-weight frequency	.302	.296	.062	.442
.02*	.016*	.026*	High frequency	.274	.276	.058	.488
			Low *p*-weight frequency	.374	.358	.1	.522
.010	.026	.026	High frequency	.078	.078	.012	.41
			low *p*-weight frequency	.12	.178	.034	.47
.034*	.011*	.022*	High frequency	.358	.35	.104	.57
			Low *p*-weight frequency	.452	.464	.166	.568
.024	.021	.022	High frequency	.122	.118	.036	.57
			Low *p*-weight frequency	.206	.2	.078	.512

produces many fewer nonword errors as a result of the nonlexical route's contribution.

Virtual patient sample

Although the exploratory simulations provided a thorough description of the models' behavior, we elected to check whether our conclusions about the differences among the models would stand up in the same kind of regression analysis that we did with the real patients. Consequently, we created a small sample of eight virtual patients, each defined by *s*, *p*, and nl values (the latter only relevant for the dual-route and nonlexical-route repetition models), to represent some degree of individual difference. For four of these virtual patients, the parameters were taken from those determined for actual patients reported in Dell et al. (2007). The parameters of these particular patients were chosen because their predicted number of nonword errors in the three repetition models was large enough to be statistically useful (nonword errors >4% of responses), but not near the model's ceiling for nonword responses (75% of responses). To increase the sample size we generated four additional virtual patients with *s* and *p* parameters in the same range (see Table 4 for the parameters). Thus, the virtual patients were like real patients who make a fair number of nonword errors.

Simulations were run in a similar fashion to those used to generate frequency and nonword contour graphs described earlier, with each virtual patient run through 1000 trials of naming and 1000 trials of repetition (with the three repetition models). Half of the trials had high-frequency targets and the other half, low-frequency targets. Frequency was manipulated in the *p*-weights.[5]

Just as for the real patients, we performed binomial, hierarchical, multiple, logistic regression analyses (predicting the log odds of making a nonword error relative to any other response type). The log odds of making a nonword error relative to any other type of response was predicted from the mean-centered frequency of the target (.1 when high-frequency and −.1 when low-frequency), the type of task (the predictor took on a value of 0 in naming and 1 in repetition), and the interaction of mean-centered log frequency and task. Notice that there is no need for predictors related to AoA, phonological density, and so on because frequency was manipulated in a completely controlled fashion in the model (e.g. CAT as a high-frequency target is compared to CAT as a low-frequency target). As with the real patients, we also included an intercept term, and a random error term for subjects to account for the hierarchical nature of the dataset (there was no random error term for items because the model simulations were run on the same target item). This regression was performed with the data from each model type (lexical-route, nonlexical-route, and dual-route) for a total of three different regressions.

We found a significant frequency effect in naming for all models, such that lower frequency targets generated more nonword errors (coefficient = −2.196, standard error = .255, $p < .001$ for the lexical-route model; coefficient = −1.993, standard error = .243, $p < .001$ for the nonlexical-route model; coefficient = −2.205, standard error = .256, $p < .001$ for the dual-route model).[6] As expected, there was no effect of frequency in the non-

[5] We also created a low-frequency version of the models in which frequency was implemented in both *s*-and *p*-weights. The results were the same as the models in which frequency was implemented only in the *p*-weights.

[6] Note that the values of the coefficients are not comparable to those discovered for the real patients because frequency is based on the number of instances in a corpus for the real patients, while in the virtual patients frequency is arbitrarily represented as a fractional change in weights. For the same reason, we do not re-express the coefficient in terms of an odds ratio because the effect size is a function of the particular values of the high and low frequency weights used in the simulation and therefore not meaningful.

Table 5
Coefficients (not discussed in the text) of regressions on data from virtual patient simulations with frequency manipulated on the *p*-weights. * Indicates significance with $p < .05$, ** indicates significance with $p < .001$.

Model	Description	Estimate[a] (standard error)
Non-lexical route	Task (change to nonword error rate in repetition)	.697 (.033)**
	Intercept	−.731 (.138)**
	Random effect of subjects	.149
Lexical route	Intercept	−.809 (.285)*
	Random effect of subjects	.644
Dual route	Intercept	−.818 (.295)*
	Random effect of subjects	.692
	Random slope for task	.026

[a] For the random error term, the number represents the variance of the random error term's distribution.

lexical-route model. In this model, the change to the frequency effect in repetition (coefficient = 1.946, standard error = .333, $p < .001$) is opposite-signed and large enough to effectively cancel out the naming coefficient. For the lexical-route model, there was no significant change to the effect of frequency in repetition compared to naming (coefficient = −.123, standard error = .360, $p = .732$). Critically, the summation dual-route model behaved exactly like the lexical-route model in showing a comparable effect of frequency in repetition and in naming (coefficient = −.516, standard error = .466, $p = .268$). The dual-route model also exhibited significantly fewer nonwords in repetition than in naming (coefficient = −1.600, standard error = .075, $p < .001$), an effect that we also observed in the real patients. The lexical-route model, in contrast, showed no difference between naming and repetition in the rate of nonword production (.011, standard error = .036, $p = .762$). Other coefficients estimated by these regressions are reported in Table 5.

Comparing model predictions with patient data

Summarizing the patient results, we found an effect of frequency in naming, and no evidence that the frequency effect was diminished in repetition. These findings suggest a fully-lexically-influenced repetition process. In the patient study, our assessment of the effects of other lexical variables that were not explicitly modeled, such as phonological density and age-of-acquisition, further supports this interpretation. Whenever an effect of a lexical variable was discovered in naming, its magnitude was the same as in repetition. It appears that nonword errors in naming and repetition are very similar in their properties. Moreover, we ruled out the possibility that our results were due to the patients' incompetence in using their nonlexical routes. A stronger nonlexical route was shown to increase repetition accuracy, but not to interact with the effect of frequency.

Fig. 5 shows the frequency effect in naming and repetition observed in real patients as well as the effects simulated by the naming and repetition model regressions. Simulations confirmed what was concluded from the analysis of patient data. Despite the interactive nature of the model, no frequency effect is observed in the nonlexical-route model. In contrast, the lexical-route model correctly demonstrates comparable frequency effects in naming and repetition. But, so does the summation dual-route model. So, which one is right? We turn to this in the general discussion.

General discussion

Word repetition must, at least partly, take advantage of the lexical retrieval system used in production. It is fair to say that, based on recent research, there is little doubt about this point. The issue that we addressed was related, but fundamentally different. We sought to measure the degree of lexical influence in repetition by an explicit comparison to naming, and to relate this measurement to architectural issues inspired by two-step accounts of lexical access in production via a formal model-driven analysis. All existing models of word repetition are consistent with the fact that there is some lexical involvement in word repetition. Even a pure nonlexical-route model could in principle involve lexical knowledge through interaction. Thus there is a need for greater precision. We need, first, to determine the actual extent of lexical involvement in repetition and, second, to determine the extent of involvement that is predicted by the various proposed

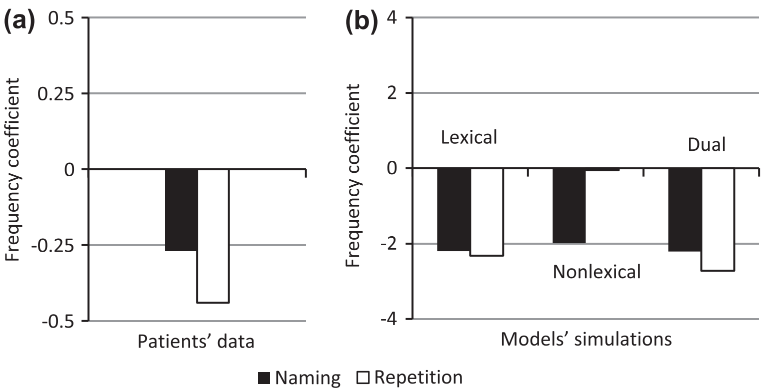

Fig. 5. Frequency effect in naming and repetition in patients' data (a) and in models' simulations with the lexical, nonlexical and summation dual-route repetition models (b). Comparison between the two panels shows that both the lexical and the summation dual-route models reflect the pattern observed in real patients.

architectures of the lexical access system. We addressed the first of these goals with an empirical study, and the second with simulations.

By analyzing frequency and other lexical effects in the patients we showed that the nonword errors made in naming and repetition were remarkably similar. In both tasks, nonword-error probability was sensitive to frequency, phonological neighborhood density, length, and age-of-acquisition. These influences did not interact with task; that is, the regressions were compatible with the conclusion that the strengths of these variables' influences were the same in naming and repetition. This is good evidence that repetition routinely uses the second step of the naming architecture.

We then used implementations of the three repetition models to understand why these effects (focusing on frequency) would be similar in naming and repetition, and found, as expected, that this result is consistent with the pure lexical-route model and incompatible with a pure nonlexical-route model. Surprisingly, though, the modeling revealed that the magnitude of the frequency effect in the summation dual-route model is comparable to that in the lexical-route model. We explained this by pointing out that there is an additive relation between the two routes within the dual-route model, so that recruitment of the nonlexical route increases repetition's accuracy, but does not decrease processing through the lexical route.

Thus, we have two models that appear to match the pattern of aphasic patients' data with regard to the influence of lexical variables. Which one is right? Recall that the patient data found a main effect of task: nonword-repetition errors were significantly less likely in repetition than in naming. Thus, although the errors were similarly sensitive to frequency in the tasks, they were less numerous in repetition. This is exactly what we found for the summation dual-route model: It makes repetition less error-prone, but does not diminish the frequency effect. Furthermore, when we entered the nonword-repetition scores as a predictor in our regression, we found additional evidence in favor of the summation dual-route model. If a nonlexical route contributes to *word* repetition, patients with better nonword-repetition scores should have fewer nonword errors in word repetition, and critically, this effect should interact with task; it should be more in evidence in repetition than in naming. These predictions were confirmed. To review, the main effect of task was significant (coefficient = −.630, fewer nonword errors in repetition than in naming), as was the interaction between task and nonword-repetition scores (coefficient = −2.487, better nonword-repetition ability reduces nonword errors in word repetition more than in naming). However, we found the three-way interaction of task, frequency and nonword-repetition score to be minimal. The independence of the frequency effect from nonword-repetition ability is inconsistent with any model in which the nonlexical route "steals from" the lexical route when it is operating, because such a model would predict a diminishing frequency effect as a function of a stronger nonlexical route. In the same vein, a pure lexical-route model does not explain why patients with better nonword-repetition ability should also do better in word repetition. Taken together, the results of the analysis are most consistent with a summation dual-route model: patients who scored higher on the nonword-repetition task also produced fewer nonword errors in the word repetition task. However, nonword-repetition ability did not modulate the frequency effect in word repetition.

Although our results show that the summation dual-route model provides the best account for our findings, it does not mean that it is the "right" model for every case. Some modeling studies, in fact, have concluded that the lexical-route model is more appropriate than the dual-route model (e.g. Baron et al., 2008; Dell et al., 2007). To reconcile our findings with these studies, we propose that there are individual differences, with some patients repeating via the dual-route model and others via the lexical-route model.

To document these hypothesized individual differences, we surveyed all of the studies of naming and repetition that have used the model-fitting/prediction method that we described in the introduction. Recall that in this method an individual patient's naming is fitted to the model, yielding *s* and *p* parameters. Then the parameterized model is used to predict the patient's word repetition. Unlike the study done here, though, there is no consideration of the influence of lexical variables, just a straight prediction of accuracy. The four studies using this method tested a total of 88 aphasic speakers (Abel et al., 2009; Baron et al., 2008; Dell et al., 2007; Hanley et al., 2004). In these studies, most of the patients' repetition accuracies were equally well predicted by *both* models, thus demonstrating the limitations of this method for model discrimination. We then looked for cases in which one model's repetition prediction was correct while the other's was much worse. Thus, we looked for patients where one model was a clear winner. There were 15 such cases, and these included nine "dual-routers", and six "lexical-routers". Fig. 6 shows the mean predicted correct repetition from both models and the actual repetition of both groups of patients. There are clear differences that need to be explained. For the dual-routers, their word-repetition performance requires that the model add in the nonlexical contribution, as determined by their ability to repeat nonwords (parameter nl). For the lexical-routers, when you add in the appropriate nonlexical contribution, the model badly overpredicts.

Thus, it appears that, although many if not most aphasic individuals are repeating words as proposed in the summation dual-route model, there are some who are not adding in the nonlexical route when repeating words.

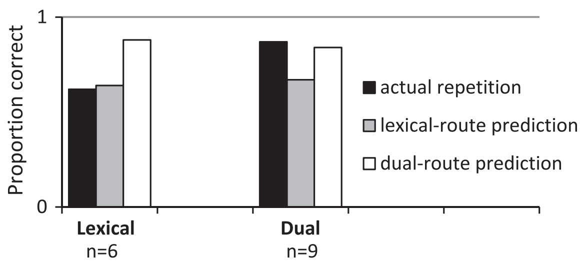

Fig. 6. Demonstration of model fit to repetition performance of six lexical and nine dual-routers (see text for details).

Where do these individual differences come from? We cannot answer this question but we can initiate the discussion. Let us first consider how unimpaired speakers repeat verbal stimuli.

Most theories of word repetition in list contexts assume that the lexical route is used (e.g. Gupta & MacWhinney, 1997). Upon hearing a meaningful word (e.g. elephant), a normal speaker would access the lexical item and accomplish repetition by simply encoding the phonology of the retrieved word. However, if a stimulus is not recognized as a word (e.g. ablemandon) the speaker must do something different, namely map input onto output phonology as accurately as s/he can. There are many models of nonword production by unimpaired speakers (e.g. Gupta & Tisdale, 2009), but for our purposes, they all implement something like the nonlexical route. The sounds are recognized and the sequence is produced by sending that sequence to output phonology in some manner. It is important to note that the nonlexical route is more than just a mechanism for echoing stimuli; it has an essential function in the learning of new vocabulary by both mature and developing speakers (see particularly, Plaut & Kello, 1999).

These considerations suggest that the *comprehension* of a verbal stimulus could affect which route(s) is used. When access to meaning fails (as when the stimulus is a nonword), the nonlexical route is recruited. It is therefore conceivable that the hypothesized normal mechanisms may apply to aphasic word repetition. If comprehension fails when a word to be repeated is presented, that word will most likely be repeated with the contribution of the nonlexical route. Thus, patients who have difficulty accessing the meaning of words in general may be more likely to recruit the nonlexical route, and thus differences in comprehension ability could end up explaining the variation in the suitability of the summation dual-route and the lexical-route models.

However, comprehension cannot be not the only factor determining the word-repetition model. There are cases of aphasic patients with intact comprehension and blocked access to phonology during production (e.g. Badecker, Miozzo, & Zanuttini, 1995; Vigliocco, Vinson, Martin, & Garrett, 1999), who nonetheless repeat words very well. If comprehension was the only determining factor, these patients should have relied on their lexical routes (which are arguably broken down in the second step). But it appears that they successfully recruit their nonlexical route to achieve excellent word-repetition performance. Here, the process of route recruitment might be based directly on the estimated quality of the production system (in the case of these patients, how good the *p*-weight is), and if it is judged to be suboptimal, then the nonlexical route is recruited. In such cases, if the lexical route is highly dysfunctional, then the summation dual-route model acts like a pure nonlexical-route model.[7]

[7] In anomic patients such as these, phonological blockages during naming can often be overcome by cueing the first phoneme of the target word. Such a cue functions much like the nonlexical route does in word repetition as characterized by the dual-route summation model. The nonlexical input (whether a phonemic cue in naming or an auditory word) adds its activation to that generated from the second step of lexical access and the result is often successful production of the target.

We conclude by acknowledging that while our discussion of individual differences is preliminary, our central conclusion is not. Aphasic word repetition is fully lexically influenced. This influence is as strong in repetition as it is in naming, as evidenced by the many similarities between error properties in naming and repetition, in frequency effects, but also in the influence of other lexical variables. This suggests that repetition routinely takes advantage of the phonological step of naming. Moreover, we showed that the pure lexical and the summation dual-route models are both consistent with the fact that the patients show comparable frequency effects in naming and repetition. The summation dual-route model, however, is uniquely consistent with the finding that aphasic repetition has, on average, fewer nonword errors than naming, and that, if a patient has better nonword-repetition ability (a more functional nonlexical route), then word repetition is benefited much more than naming. These findings demonstrate that the nonlexical route adds to the lexical route in a dual-route repetition structure for a sizeable portion of our sample. Finally, our results make clear that future research efforts should not be directed at the question of *whether* the lexical route or summation dual-route model is correct. Instead, there should be a focus on crafting a powerful, predictive account of when the nonlexical route adds to the performance of the lexical route in word repetition. Such an account would be a key component of any general theory of word production, with ramifications for lexical acquisition, comprehension, reading, verbal short-term memory, as well as theories of language impairment.

Acknowledgments

This research was supported by the National Institutes of Health, DC000191 and HD44458. We would like to thank Randi Martin, Stefanie Abel, Robert Hartsuiker, Victor Ferreira and two anonymous reviewers for their valuable comments.

References

Abel, S., Huber, W., & Dell, G. S. (2009). Connectionist diagnosis of lexical disorders in aphasia. *Aphasiology, 23*, 1–26.

Ackerman, T., & Ellis, A. W. (2007). Case study: Where do aphasic perseverations come from? *Aphasiology, 21*, 1018–1038.

Alario, F. X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes, 17*, 299–319.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Pennsylvania, PA: Linguistic Data Consortium, University of Pennsylvania.

Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition, 57*, 193–216.

Baron, R., Hanley, J. R., Dell, G. S., & Kay, J. M. (2008). Testing single and dual route computational models of auditory repetition with new data from six aphasic patients. *Aphasiology, 22*, 62–76.

Biederman, B., & Nickels, L. (2008). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex, 44*, 683–697.

Bonin, P., Barry, C., Meot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50*, 456–476.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589–608.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 318–344.

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 and Cognitive Processes, 5*, 313–349.

Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language, 56*, 490–520.

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

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20*, 611–629.

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

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients. *Cognitive Neuropsychology, 25*, 136–164.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 71*, 310–322.

Ferreira, V. S., & Humphreys, K. R. (2001). Syntactic influences on lexical and morphological processing in language production. *Journal of Memory and Language, 44*, 52–80.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43*, 182–216.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47*, 27–52.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.). *The psychology of learning and motivation* (Vol. 9). New York: Academic Press.

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 and Language, 82*, 113–145.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267–333.

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language, 61*, 481–502.

Hanley, J. R., Dell, G. S., Kay, J., & Baron, R. (2004). Evidence for the involvement of a nonlexical route in the repetition of familiar words: A comparison of single and dual route models of auditory repetition. *Cognitive Neuropsychology, 21*, 147–158.

Hanley, J. R., Kay, J., & Edwards, M. (2002). Imageability effects, phonological errors, and the relationship between auditory repetition and picture naming: Implications for models of auditory repetition. *Cognitive Neuropsychology, 19*, 193–206.

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science, 8*, 191–219.

Hillis, A., & Caramazza, A. (1991). Mechanisms for accessing lexical representations for output: Evidence from a category specific semantic deficit. *Brain and Language, 40*, 106–144.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G., Martin, R., & Stuart, G. (1997). Word-frequency effects on shortterm memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 1217–1232.

Jacquemot, C., Dupoux, E., & Bachoud-Levi, A. C. (2007). Breaking the mirror: Asymmetrical disconnection between the phonological input and output codes. *Cognitive Neuropsychology, 24*, 3–22.

Jeffries, E., Crisp, J., & Lambon Ralph, M. A. (2006). The impact of phonological or semantic impairment on delayed auditory repetition: Evidence from stroke aphasia and semantic dementia. *Aphasiology, 20*, 963–992.

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.

Kempen, G., & Huijbers, P. (1983). The lexicalisation process in sentence production and in naming: Indirect election of words. *Cognition, 14*, 185–209.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*, 463–492.

Knobel, M., Finkbeiner, M., & Caramazza, A. (2008). The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production. *Cognitive Neuropsychology, 25*, 256–286.

Lallini, N., Miller, N., & Howard, D. (2007). Lexical influences on single word repetition in acquired spoken output impairment: A cross language comparison. *Aphasiology, 21*, 617–631.

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

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

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 1–36.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review, 89*, 483–506.

Martin, R. C. (2003). Language processing: Functional organization and neuroanatomical basis. *Annual Review of Psychology, 54*, 55–89.

Martin, N., Dell, G. S., Saffran, E. M., & Schwartz, M. F. (1994). Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain and Language, 47*, 609–660.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language, 41*, 3–29.

Martin, N., Saffran, E. M., & Dell, G. S. (1996). Recovery in deep dysphasia: Evidence for a relation between auditoryverbal STM and lexical errors in repetition. *Brain and Language, 52*, 83–113.

Martin, N., Schwartz, M. F., & Kohen, F. P. (2005). Assessment of the ability to process semantic and phonological aspects of words in aphasia: A multi-measurement approach. *Aphasiology, 20*, 1–13.

Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurence of naming errors: Implications for models of lexical retrieval. *Journal of Memory and Language, 28*, 462–485.

Meyer, A., & Bock, K. (1992). The tip of the tongue phenomenon: Blocking or partial activation. *Memory & Cognition, 20*, 715–726.

Navarrete, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology, 59*, 1681–1690.

Nickels, L., & Howard, D. (1994). A frequent occurrence? Factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology, 11*, 289–320.

Nozari, N., & Dell, G. S. (2009). More on lexical bias: How efficient can a "lexical editor" be? *Journal of Memory and language, 60*, 291–307.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*, 273–315.

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and speech production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381–415). Mahwah, NJ: Erlbaum.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study in connectionist neuropsychology. *Cognitive Neuropsychology, 10*, 377–500.

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*, 413–425.

Rapp, B. C., Epstein, C., & Tainturier, M. J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology, 19*, 1–29.

Rapp, B. C., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*, 460–499.

Ruml, W., Caramazza, A., Capasso, R., & Miceli, G. (2005). Interactivity and continuity in normal and aphasic language production. *Cognitive Neuropsychology, 22*, 131–168.

Schriefers, H., Meyer, A., & Levelt, W. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*, 86–102.

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.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*, 247–250.

Vigliocco, G., Vinson, D. P., Martin, R. C., & Garrett, M. F. (1999). Is "count" and "mass" information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language, 40*, 534–558.

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. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers, 20*, 6–11.