This article was downloaded by:[Swets Content Distribution]
On: 26 July 2007
Access Details: [subscription number 768307933]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Aphasiology

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713393920>

Anomia is simply a reflection of semantic and phonological impairments: Evidence from a case-series study

Matthew A. Lambon Ralph; Lynne Moriarty; Karen Sage

Online Publication Date: 01 January 2002

To cite this Article: Ralph, Matthew A. Lambon, Moriarty, Lynne and Sage, Karen (2002) 'Anomia is simply a reflection of semantic and phonological impairments: Evidence from a case-series study', *Aphasiology*, 16:1, 56 - 82

To link to this article: DOI: 10.1080/02687040143000448
URL: <http://dx.doi.org/10.1080/02687040143000448>

PLEASE SCROLL DOWN FOR ARTICLE

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

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

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

© Taylor and Francis 2007

Anomia is simply a reflection of semantic and phonological impairments: Evidence from a case-series study

Matthew A. Lambon Ralph

MRC Cognition and Brain Sciences Unit, Cambridge, UK

Lynne Moriarty

Sheffield Speech and Language Therapy Agency, UK

Karen Sage

University of Manchester, Manchester, UK

The York Speech Therapy Interest Group

A number of recent studies have attempted to explain patterns of normal and impaired performance in a variety of different language tasks with respect to the same set of “primary” systems rather than resorting to explanations in terms of dedicated processes, specific to each and every language activity. In this study we consider whether the same approach can be taken to patterns of impaired single-word speech production. Specifically, using cross-sectional data from 21 aphasic patients we tested the hypothesis that the degree and nature of anomia can be explained using independently derived measures of the integrity of the patients’ phonological and semantic/conceptual representations, without postulating a role for an abstract lexical level of representation. At a global level, we found that these two measures explained 55–80% of the variance in the patients’ naming accuracy, a figure which approaches that found for test reliability. There was also a close fit between observed and expected naming accuracy for all individual patients. The same two measures also predicted the rate of different types of anomic error across individuals. Measures intended to assess lexical integrity did not explain any additional, unique variance in naming accuracy. We discuss these results and the theoretical approach with respect to existing theories of speech production, and evaluate the case-series methodology itself, both as a tool to reveal the underpinnings of speech production and as a neuropsychological technique in general.

Address correspondence to: Prof. M.A. Lambon Ralph, Department of Psychology, University of Manchester, Oxford Road, Manchester M13 9PL, UK. Email: matt.lambon-ralph@man.ac.uk

We would like to thank all the patients who contributed to this analysis. The study was conducted while MALR was supported by a grant from the National Institutes of Health, USA (to J.L. McClelland and K. Patterson), and LM by a grant from the Sheffield Speech and Language Therapy Agency.

The York Speech Therapy Interest Group meets on a regular basis to facilitate clinical–academic interaction. The present study would not have been possible without the contributions of the following members of the group: Tessa Ackerman, Kate Arul, Helen Bird, Jenni Crisp, Sian Davies, Andy Ellis, Ruth Harold, Pat Heaney, Meryl Jones, Sally Knapp, Hannah Luff, Ruth McKinnon, Julie Morris, Alison Newton, Ruth Raynor, Sam Sharpe, Lucy Southwell, Janine Swinson, Lynn Taylor, Emma Walker, and Elaine Whitton.

In a recent review of the neuroimaging literature, Neville and Bavelier (1998, p. 258) concluded that "the functional role of the language-related areas is more accurately characterized in terms of . . . systems, such as phonology, syntax and semantics, than in terms of activities, such as speaking, repeating, reading and listening". This would also be an accurate summary of the approach taken by a number of researchers to explaining normal and impaired performance across a variety of language tasks. It is possible, for example, that normal reading and acquired dyslexias reflect interactive processing between three "primary" systems: conceptual knowledge or semantic memory, phonology, and the visual processing required to decode patterns of orthography (Patterson & Lambon Ralph, 1999; Plaut & Shallice, 1993). Under this view, phonological dyslexia¹ is a product of a generalised phonological deficit not specific to reading (Farah, Stowe, & Levinson, 1996; Patterson, Suzuki, & Wydell, 1996) and in severe cases, this impairment may lead to deep dyslexia² (on the assumption that the two disorders are points along a continuum rather than two discrete entities: Friedman, 1996). Pure alexia or letter-by-letter reading³, under this hypothesis, is attributed to a deficit in orthographic processing, which is just one manifestation of a general visual processing impairment (Behrmann, Nelson, & Sekuler, 1998a, Behrmann, Plaut, & Nelson, 1998b). The set is completed by assuming that surface dyslexia⁴ follows from a significant disruption to semantic memory (Graham, Hodges, & Patterson, 1994; Patterson & Hodges, 1992).

Opponents to this view might argue that models postulating components dedicated to reading do just as well, if not better, at describing performance in normal or alexic populations. The potential of explanations in terms of primary systems is revealed, however, when the same assumptions can be applied successfully to describe the nature of processing underlying other language activities. This approach, for instance, explains not only why patients with selective semantic impairment are surface dysgraphic⁵ as well as surface dyslexic but also why the deficit observed in spelling is more pronounced than that seen in reading (Graham, Patterson, & Hodges, 2000). Very recent studies have also applied this method to inflectional verb morphology. Rather than assuming a dual-mechanism of lexical look-up and rule-based generation for producing the past tense of English verbs (Pinker, 1999; Ullman et al., 1997), it is possible to construct a computational model of past tense transformation based solely on phonological and semantic processes (Joanisse & Seidenberg, 1999). In this domain, patients with semantic deficits are predicted to, and do, reveal a marked and frequency-modulated difficulty in producing the past tense of irregular verb forms to which they generate mainly regularisation errors such as *go* → "goed" (Patterson, Lambon Ralph, Hodges, & McClelland, 2001).

The principal aim of the present study was to extend previous attempts to apply this approach to speech production. Influential contemporary models of speech production

¹ A disorder characterised by a significant difference in oral reading accuracy for words over nonwords.

² This reading disorder typically presents with a number of different features of which the cardinal one is the production of semantically related responses in single word reading aloud.

³ Patients with pure alexia or letter-by-letter reading show a pronounced effect of letter length on word reading times and may read out the individual letters within a target word before giving the pronunciation.

⁴ This form of alexia is characterised by a relative impairment of reading words with atypical spelling-to-sound correspondences. For these words, the elicited pronunciation tends to match that expected for the typical spelling-to-sound correspondence (e.g., reading *pint* as though it rhymed with *mint*).

⁵ This form of dysgraphia mirrors that for surface dyslexia: poor spelling for words with atypical sound-spelling correspondences and errors that match the more typical spelling (e.g., "cough" →

COFF).

assume that there are one (Dell & O'Seaghdha, 1992), two (Levelt, 1992), or even three (Levelt, Roelofs, & Meyer, 1999) dedicated sets of lexical representations that mediate between conceptual knowledge and segmental phonological representations. There are a number of parallel distributed processing (PDP) networks, however, that include direct mappings between semantic and phonological representations (Plaut, 1999; Plaut & Kello, 1999; Plaut & Shallice, 1993). This raises the possibility that speech production, like reading aloud and inflectional morphology, could be characterised as a product of a direct interplay between semantic and phonological representations without recourse to sets of specific lexical representations (Lambon Ralph, 1998; Lambon Ralph, Sage, & Roberts, 2000). If so, word-finding difficulties in aphasia should be explicable in terms of impairments to one or both of these primary systems. In a review of the literature on word-finding difficulties in aphasia, Lambon Ralph et al. (2000) concluded that the anomia in nearly all patients could be ascribed to deficits to semantic and/or phonological representations.

A direct test of this hypothesis would assess how well independently derived measures of semantic and phonological impairment predict the degree and nature of anomia observed in a case-series of aphasic patients. Case-series data also bring a number of methodological advantages over single-case studies. In addition to the sort of quantitative analyses proposed here, case-series data allow direct comparison between individual patients because all are tested on the same materials. Among other things, this allows researchers to investigate the function that relates severity of a damaged system (e.g., level of semantic impairment) to the resulting impaired behaviour (e.g., degree of anomia). Given such a function, it may be possible to adjudicate between those fluctuations in performance that are due to random variation in test scores (such data would fall within the confidence limits of the function) versus those that reflect additional cognitive deficits (reflected in data points significantly beyond the estimated function).

Lambon Ralph et al. (2001) combined longitudinal data from 16 patients with semantic dementia to study the effect of semantic impairment on picture naming. Until very late stages in the disease, the patients have impaired comprehension and pronounced anomia without any of the significant deficits that often accompany semantic impairment in other aetiologies; these preserved cognitive domains include day-to-day memory, orientation, nonverbal problem solving, perceptual and spatial skills, and (within the language domain) syntactic and phonological processing (Hodges, Patterson, Oxbury, & Funnell, 1992). Overall Lambon Ralph et al. (2001) found that naming performance was strongly modulated by the degree of semantic impairment. Even small semantic deficits led to significant anomia, characterised by omission and co-ordinate semantic errors but no phonologically related paraphasias.

Another aspect of the Lambon Ralph et al. study provides a useful example of the central thesis of this paper. While there was a clear and direct relationship between selective semantic impairment and word-finding difficulties, the degree of anomia varied considerably between individual patients. Particularly at the mid-range of disease severity, there were several patients with roughly equivalent levels of impaired comprehension yet different degrees of anomia. Given that these patients do not seem to have any disruption of phonological representations, this finding could be taken as evidence of an additional post-semantic and pre-phonological impairment contributing to the anomia found in some cases (Graham, Patterson, & Hodges, 1995), perhaps even supporting an argument for abstract, word-specific representations or lemmas. This "two-impairment" hypothesis, however, has several limitations. By using a case-series approach, it became clear that all differences between the patients were ones of degree

rather than type. Longitudinal profiles all began and especially ended in very similar ways (with combined anomia and semantic impairment); and rather than the patterns falling into two neat, non-overlapping groups, there was a continuum of profiles from a purely semantically driven anomia to the disproportionate progressive anomia true of some individuals. Another unsatisfactory part of the two-impairment hypothesis is that it offers no basis for predicting which patients should fall into which group: there is no independent factor—psychological, neuroanatomical, or otherwise—that would enable one to predict where within the continuum a patient's longitudinal trajectory should fit. We shall return to this point—namely the lack of an independent measure of post-semantic “lexical” integrity—later.

Lambon Ralph et al. (2001) were able to show that the relative asymmetry of the bilateral temporal lobe atrophy provided an independent factor capable of predicting the degree of anomia given a certain level of semantic impairment. As the atrophy became more pronounced in the left rather than the right temporal lobe, the degree of word-finding difficulties increased. The possible causal relationship between laterality of atrophy and degree of anomia was demonstrated using a computational model. This network encapsulated three key assumptions: (1) progressive semantic impairment is the one and only key deficit in all these cases; (2) conceptual representations are supported by bilateral temporal lobe structures but phonology is strongly left-lateralised; (3) two communicating functional systems supported by regions on the same side of the brain will be more tightly coupled than if the supporting neural substrates are in different hemispheres (Plaut, 1999). Although we will not be concerned with the distribution of atrophy in the present paper, the critical finding from study of patients with semantic dementia is that their anomia was accurately predicted by the *degree* and *distribution* of impairment within the semantic system alone.

The purpose of the present study was to test how well measures of semantic and phonological impairment can predict the degree and nature of anomia as measured across a case-series of (non-progressive) patients with aphasia—predominantly following cerebral vascular accidents (CVA). Previous studies have typically highlighted the phonological impairments present in this population (Gagnon et al., 1997; Goodglass et al., 1997) and thus these patients would provide complementary evidence to that described earlier for semantic dementia. It is clear that CVA can lead to impaired conceptual knowledge too, especially when the infarct is sufficiently large to include Brodmann areas 22, 21, and 37 in the middle and posterior temporal lobe (Chertkow, Bub, Deaudon, & Whitehead, 1997). We therefore included assessment of both “primary” systems—semantics and phonology—in the test battery.

METHOD

Patients

Patients were selected from the current combined case load of the members of the York Speech Therapy Interest Group. In order to give our hypothesis the hardest test, we endeavoured to include a wide variety of aphasic patients. Consequently a minimum number of selection criteria were used. Aphasia classification and aphasia severity were not used as methods of selection. Patients were recruited just so long as they performed outside the normal range on one or both of the picture naming tasks (see later). No minimal score criterion was set. Only patients with aphasia following acute brain injury were included, and those with dysarthria and dyspraxia were excluded.

Assessment

Each patient was asked to complete nine assessments. These were conducted over a number of short testing sessions within a 2–3 week period. The following tests were included within the battery:

1. Naming

- (a) *Boston Naming Test (BNT: Kaplan, Goodglass, & Weintraub, 1976)*: The BNT was administered without the standard systematic cueing and was used as an initial screening test for anomia. Administration was discontinued if a patient produced six consecutive naming errors.
- (b) *100-item naming test (Lambon Ralph, Ellis, & Sage, 1998)*: This assessment includes 100 simple line drawings drawn primarily from two corpora (Riddoch & Humphreys, 1992; Snodgrass & Vanderwart, 1980). The patients were asked to name each picture. A 10-second time limit was applied to each trial and only the first full response (a word or nonword response containing at least one syllable) was used in the subsequent analyses. Errors were classified and analysed separately (see later for details).

2. Comprehension

- (a) *Pyramids and Palm Trees test (PPT: Howard & Patterson, 1992)*: The all-picture version of this test was given. Patients are required to select one of two alternative pictures that is most closely associated with a target item (e.g., for a *pyramid*, the subject has to select a *palm tree* and not a *fir tree*).
- (b) *100-item written word–picture matching (WWPM: Lambon Ralph et al., 1998)*: This task uses the same 100 concepts as for the naming test just described. Each target picture is presented with four close semantic foils and the written name of the target. Subjects are required to pick which picture matches the written name.
- (c) *100-item spoken word–picture matching (SWPM: Lambon Ralph et al., 1998)*: This task is identical to that just described except that the name of the target item is spoken by the examiner rather than written on the page.
- (d) *Synonym judgement (Kay, Lesser, & Coltheart, 1992)*: This task contains 60 trials. In each a pair of written words is presented and the subject has to indicate whether they have approximately the same meaning or not. For this test, the written words were also read aloud by the examiner.

3. Reading aloud and repetition

- (a) *Surface List (Patterson & Hodges, 1992)*: The 84 low-frequency words from this test were given to each subject for reading aloud. Half of the words have atypical and the other half typical spelling-to-sound correspondences.
- (b) *Nonword reading (Test 24: Kay et al., 1992)*: This reading test includes 24 monosyllabic nonwords varying in length from three to six letters.
- (c) *Word and nonword repetition (Franklin, Turner, & Ellis, 1992)*: This assessment of single item repetition includes 40 words and 40 nonwords. Words and nonwords are presented in a random, mixed order.

RESULTS

Analyses of the case-series data are split into four sections, each designed to test aspects of our working hypothesis. If this proposal is correct then we should find clear evidence that phonological and semantic impairments are present in the case-series. We then go on

to derive estimates of the degree of *generalised* semantic and phonological deficits for each case (using three different methods) and test how well these measures predict naming performance. In the third section, two additional estimates of the integrity of abstract lexical representations are computed. If such representations play a role in speech production then the measures of lexical integrity should significantly improve the prediction of naming accuracy over and above the variation explained by semantic and phonological factors alone. Finally, we investigate the relationship between different types of naming error and the degree of semantic/phonological impairment.

General characteristics of the case-series and overall results

A total of 21 patients met the selection criteria and were able to complete the nine tests. The only exception was patient RN who appeared to be guessing on the synonym judgment task, and administration was discontinued. Consequently a value of 30/60 was substituted for all the following analyses. There were 10 male and 11 female patients ranging in age from 20 to 81 years (mean age 65.2; SD 13.8). Months post-onset varied between 3 and 88 (mean 18.9; SD 21.8). All patients were right-handed with one exception (patient BT). All but one (TS: head injury) were aphasic following CVA. Unfortunately information from structural scanning is limited, but left hemisphere infarction was confirmed by CT in 15 cases.

Results for the 21 patients across the nine-test aphasia battery are shown in Table 1. The patients are ordered from left to right in terms of their accuracy on the 100-item naming test. A wide range of anomia was covered by this cohort, ranging from patients with profound word-finding difficulties (e.g., PW, BT, EM) through to those cases with only a very mild degree of anomia (e.g., JN). In addition the battery highlighted concurrent semantic and, in particular, phonological impairments in this patient group. While it is clear that patients with semantic dementia have a central semantic deficit affecting all modalities of input (Bozeat et al., 2000) the same is not always true of aphasic subjects. Some patients can have specific difficulties accessing conceptual knowledge from spoken (e.g., Franklin et al., 1996) or written modalities (Lambon Ralph et al., 1998; Lambon Ralph, Sage, & Ellis, 1996). To allow for these possible modality differences, we included a range of comprehension assessments that varied in reliance on spoken words, written words, and pictorial input. Central semantic deficits would be reflected, therefore, in abnormal performance on two or more of these assessments. If one accepts the view that abstract lexical forms mediate verbal comprehension as well as expression then it is possible, at least theoretically, to have selectively impaired verbal access to meaning for both written and spoken words (a semantic lexicon/lemma impairment: Butterworth, Howard, & McLoughlin, 1984). Central semantic deficits would be reflected, under this formulation, in abnormal performance on the all-picture PPT and at least one other assessment (Cherkow et al., 1997). Eight cases met the latter criteria (BT, EM, RN, TH, ML, EB, PC, & TJ) while another seven performed abnormally on two or more of the non-PPT tasks (PW, BMa, PS, FW, BMb, TS, DH, & DM). In summary, the comprehension tasks highlighted mild semantic deficits in a considerable number of cases. Although the degree of semantic impairment was not as pronounced as typically found in semantic dementia, it is possible that these deficits would have significant impact on speech production. Even with low levels of semantic impairment, Lambon Ralph et al. (2001) found a marked effect on picture naming both in a group of patients with semantic dementia and with a computational model of object naming.

TABLE 1
Results for the nine-test aphasia battery

Patient	PW	BT	EM	RN	AC	IB	TH	ML	EB	BMa	PS	PC	JM	TJ	FW	BMb	TS	DH	DM	KG	JN	Controls	
Age (years)	69	68	77	64	76	49	45	67	70	63	72	81	69	57	74	69	20	68	78	74	59		
Months post-onset	10	7	16	9	6	88	3	17	7	3	62	6	3	32	45	10	24	5	7	13	6		
Max.																							
Score																							
Naming																							
100-item	1	4	6	18	36	37	39	40	41	41	45	56	61	63	67	71	72	75	76	78	87	93–100	
BNT	3	0	2	10	29	8	32	13	19	18	6	34	37	26	9	22	38	24	40	23	50	42–60	
Comprehension																							
PPT	52	51	39	39	50	49	48	45	52	49	43	51	48	49	50	50	50	49	50	51	52	49–52	
WWPM	100	98	40	76	78	98	87	96	91	94	90	97	100	94	97	95	95	84	96	97	96	96–100	
SWPM	100	94	63	83	82	97	86	99	91	90	83	93	100	97	92	94	96	88	91	98	98	96–100	
Synonyms	60	52	37	36	30*	60	52	49	42	49	58	46	56	55	53	57	47	48	55	55	58	56–60	
Reading																							
Regular	42	0	0	0	10	17	10	29	0	16	15	12	39	24	31	39	24	19	36	36	40	41.4 (1.5)	
Exception	42	0	0	0	11	16	14	28	0	18	11	16	41	26	31	39	20	19	32	31	42	40.2 (1.9)	
Nonword	24	0	0	0	4	0	11	0	0	1	5	17	13	7	0	12	1	1	17	11	22	22–24	
Repetition																							
Word	40	29	37	20	39	26	33	40	39	31	36	19	39	34	39	37	40	40	37	38	39	40	35–40
Nonword	40	16	26	12	31	21	22	33	34	13	22	6	36	22	30	19	40	34	31	30	28	37	30–38

BNT = Boston Naming Test; PPT = Pyramids and Palm Trees test; WWPM = written word–picture matching; SWPM = spoken word–picture matching.
Scores shown in *italic bold* type are outside the normal control range.
* substituted value (patient was performing no better than chance).

The battery revealed evidence for phonological impairment in the vast majority of cases except for JN, the patient with the mildest degree of anomia. There was no evidence for surface dyslexia in the group but rather many showed a reading pattern consistent with phonological dyslexia. Only JN performed within the normal range on the nonword reading assessment. This reading disorder is apposite given that phonological dyslexia is reliably associated with phonological impairments (Patterson & Lambon Ralph, 1999). Further evidence for a generalised phonological impairment was highlighted by the word and nonword repetition task, even though it may not be the most sensitive measure of this type of deficit (in comparison to tasks such as phonemic blending, segmentation, and generation of spoonerisms). Given the relatively transparent mapping between input and output in the tasks of reading, and especially repetition, impairments to phonology may have to be severe before they affect overt repetition/reading performance. Given the arbitrary mapping between semantics and phonology, however, one would expect the same level of phonological impairment to have a greater impact on picture naming (Croot, Patterson, & Hodges, 1998). In this regard, it is important to note that of the 18 cases with impaired word and nonword reading, 11 also demonstrated deficits on word and/or nonword repetition.

How well can measures of semantic and phonological impairment predict the degree of anomia?

The association of anomia with semantic and/or phonological deficits does not necessarily imply a causal link. A more direct test of our central hypothesis would be to assess how closely word-finding deficits can be predicted from the degree of impairment in these two “primary” systems. We used three slightly different methods to derive these two measures. Despite our initial concerns for modality/material differences in comprehension, we found high correlations between all the semantic tests (r between .67 and .93, all $p < .001$). Likewise there was a high and significant correlation between word and nonword repetition ($r = .87$, $p < .001$). Neither word nor nonword repetition was correlated with accuracy on any of the comprehension tests, however (r between .06 and .16, all n.s.). Reading and, most importantly, naming correlated both with the comprehension tests (r between .41 and .79, p between .06 and $< .001$) and with the repetition assessments (r between .48 and .58, p between .03 and .006). These differing intercorrelations are neatly summarised by a principal component analysis after varimax rotation⁶. This descriptive analysis explained 77% of the variance in the patients’ overall performance and identified two main factors with an eigenvalue greater than 1. The loading of each task on the two principal components is shown in Table 2 (see “Full analysis”).

As can be seen in Table 2, the high loading of the semantic tests and no loading of repetition on the first factor suggest that this component primarily reflects the degree of semantic impairment. The reverse is true of the second factor, which reflects a strong

⁶ These principal component analyses consider all the patient data across all tasks simultaneously and extract trends or factors that appear in these scores. If these factors have an “eigenvalue” greater than 1 then it indicates that the trends or factors are statistically meaningful. Varimax rotation is a statistical method for reorientating these factors in order to maximise and minimise the loading of each task on each factor—the effect of this is that it is easier to interpret the psychological meaning of the extracted factors. In this analysis, after rotation, all the semantic tasks load on one factor and the phonologically related assessments on the other. The analyses suggest, therefore, that the degree of semantic and phonological impairment are critical underlying trends in the patient data.

TABLE 2
Component matrices for three principal component analyses after varimax rotation

Task	Full analysis		Analysis for Regression 1		Analysis for Regression 2	
	Principal Components		Principal Components		Principal Components	
	Semantic	Phonological	Semantic	Phonological	Semantic	Phonological
BNT	.493	.733	Not included	Not included	Not included	Not included
100-item naming	.579	.661	Not included	Not included	Not included	Not included
PPT	.868	.115	.870	.085	.886	.090
WWPM	.921	.060	.928	.040	.935	.051
SWPM	.869	.156	.880	.138	.898	.150
Synonyms	.848	.259	.865	.236	.840	.281
Word reading	.478	.756	.501	.719	Not included	Not included
Nonword reading	.376	.703	.405	.669	.322	.805
Word repetition	−.060	.852	−.018	.886	Not included	Not included
Nonword repetition	−.063	.900	−.019	.926	−.015	.889

phonological component. Word and nonword reading load much more heavily on the phonological than semantic component. Importantly, naming implicates both the semantic and phonological factors.

We used this descriptive technique to produce global, generalised measures of semantic and phonological integrity. This has a number of advantages over using individual test results: (1) Each assessment can never be a pure measure of semantic or phonological impairment—for example, in addition to phonological activation reading requires visual decoding of the orthographic string and repetition necessitates acoustic processing of the spoken stimulus. However, the statistical procedure uncovers the common variation in these different measures. This is what we hope to measure—the degree of impairment to the phonological representations that mediate reading, repetition, and naming—rather than poor task performance that reflects other concurrent deficits (e.g., poor visual/acoustic processing). (2) We also know that assessments have different sensitivities to a specific impairment—e.g., repetition of concrete, early acquired, frequent words may be very insensitive to mild phonological impairment while production of spoonerisms may be very sensitive. On the other hand, for more severe impairment such demanding tasks may become redundant because of floor effects. The PCA procedure provides a way of weighting each of the appropriate test results to derive a single generic measure. (3) Finally, it is clear that there will be uninteresting fluctuation in test scores due to idiosyncratic factors such as fatigue that may vary over time. By taking more than one measure of semantic and phonological integrity, of course, such irrelevant variation in test scores will tend to be averaged out.

Two slightly different estimates of generalised semantic and phonological impairments were used subsequently to predict naming accuracy and rates of different error types (see later). It would be an odd and circular procedure to include measures of naming within the principal component analysis (PCA) to use later for predicting naming performance. We repeated the analysis, therefore, without the naming measures. This revised analysis explained 78% of the variance within the patients’ performance and again produced two factors with eigenvalues over 1. The loading of each task is shown in Table 2 (see “Analysis for Regression 1”), which reveals a very similar pattern to that

from the full analysis. By using a linear combination of the loading of each task on the two factors, we produced two global measures reflecting overall semantic and the phonological integrity for each patient. These two sets of scores were then used to predict the patients' naming performance (see later).

Given the central hypothesis under examination in this study, it could be argued that word reading and repetition should not be included within these factors. Those theories that posit a role for abstract lexical forms would presumably expect the integrity of these representations to influence reading and repetition for real words in particular (in fact we test this hypothesis explicitly later). We repeated the PCA without these two measures. Again a high degree of variance was explained by a two-factor solution⁷ and the loading of each task is shown in Table 2 (see "Analysis for Regression 2"). From this analysis, the same linear combination technique was used to produce overall semantic and phonological factor scores.

A bonus of using this PCA-based technique is that it should produce a more stable measure of each patient's semantic and phonological deficit (on the assumption that, for example, the four comprehension tasks are basically tapping the same underlying central semantic system). We are, in effect, producing global semantic and phonological scores by combining each of the relevant test results, using the PCA to give a weighted average. A critic may argue that these derived measures are too opaque—the clinically transparent test scores being lost in the PCA, producing instead abstract factor scores. With this in mind, we also selected two tests to serve directly as semantic and phonological predictors. We selected the PPT as the semantic indicator—to guard against any comprehension deficit due specifically to verbal impairment in these aphasic subjects—and we chose nonword reading as a relatively sensitive but specific measure of phonological impairment⁸.

Before we report how well these measures predicted naming accuracy, it is worth considering how much variance we should expect to be able to explain. The maximum target variance would be estimated by using either a split-half reliability test—in effect using the accuracy on one half of the naming test to predict the patients' score on the other half—or by using one naming test to predict accuracy on the other (a measure of test validity). The correlation between the two naming tests was 0.74, indicating a shared variance of 55%. A split-half correlation using the 100-item naming test (correlating scores for the odd and even items) produced $r^2 = .90$.

The three different measures of semantic and phonological impairment were used in a linear regression analysis to predict the patients' naming accuracy both on the 100-item naming test and the BNT. The observed and predicted naming scores across the case-series are shown for the BNT in Figure 1 and 100-item naming test in Figure 2. Figures 1a and 2a show the results for the analysis that used the PCA-based semantic and phonological scores including word reading and repetition data. For the BNT, these two measures explained 68% of the variance in naming accuracy, and case-diagnostic analysis showed that there were no differences between observed and expected scores greater than two standard residuals. For the 100-item naming test again 68% of the naming accuracy variance was explained by a linear combination of these two factors and only one case showed a significant difference between observed and expected scores

⁷ Note that we did not impose a two-factor solution in any of these PCAs. The standard method of extracting components with eigenvalues greater than one was adopted throughout.

⁸ In fact in a stepwise regression entering all test scores to predict BNT scores, PPT and nonword reading were the only two tasks to load significantly.

(PW's naming was 2.3 standard residuals below that predicted). Figures 1b and 2b show the results for the PCA-based scores that excluded word reading and repetition data. For the BNT (Figure 1b), 72% of the naming variance was explained by these two estimates of semantic and phonological impairment. Case diagnostics found no significant differences between observed and expected naming scores across the individual patients. For the 100-item test (Figure 2b), 60% of the variance was explained, with the case-diagnostic analysis again showing a significant difference between observed and expected scores for only one patient (again PW's expected score was 2.3 standard residuals greater than that observed). Finally using the PPT and nonword reading scores as predictors, 65% of the variance in the BNT performance was explained (Figure 1c) with the predicted score for one case deviating significantly from that observed (TS's BNT score was 2.5 standard residuals greater than that expected). For the 100-item naming test, 55% of the naming variance was explained by these two measures with only one significant difference between observed and expected scores (PW's score was 2.6 standard residuals below that expected).

In summary, the three different measures of semantic and phonological impairment produced a relatively close fit between observed and predicted naming scores both for the BNT and the 100-item naming test. Of course one would not predict a perfect fit—all measures are imperfect, potentially varying from session to session, or according to premorbid individual differences, and so on. It is important, therefore, to view the analyses in light of the correlations between and within naming tasks. The three different measures of semantic and phonological integrity were able to explain at least as much variance as that explained by using one naming test to predict the other (55%) although none was able to reach the level produced by using performance on one half of the 100-item to predict scores on the other half (90% shared variance). Additional evidence for the close fit provided by measures of the primary systems was reflected in the case-diagnostic analyses. These revealed either no significant difference between observed and expected scores or a significant difference for only 1 out of the 21 patients. Clearly given an $\alpha = 0.05$, this is no different from that expected by chance.

Does a specific measure of “lexical” processing add any predictive power?

If speech production did involve a separate set of abstract lexical representations and these were compromised in at least some aphasic patients, then one might expect a measure of this level of processing to contribute significantly to predicting the patients' naming accuracy. This possibility raises a new problem: what is an independent measure of the integrity of the abstract lexical representations? Obviously one needs to derive such a measure from tasks other than naming and we have attempted to use characteristics of reading and repetition for three such estimates. It should be noted from the outset, however, that such measures are theory-dependent and some models of speech production say little, if anything, about which components of the postulated speech production system are used in reading aloud or repetition. The first estimate is derived from the notion that word repetition can be supported by these abstract lexical representations but nonword repetition cannot (on the assumption that nonwords are not repeated by some form of “lexical” analogy, which is unlikely in this framework as abstract lexical representations, by definition, have no phonological content). The second estimate comes from the assumption that words with atypical spelling-to-sound correspondences are more reliant on lexical integrity whereas regular words and

A

B

C

Figure 1. Predicting BNT naming accuracy from three estimates of semantic and phonological impairment. Predictions 1–3 used three different methods to derive estimates of semantic and phonological integrity (see text for details).

nonwords can activate segmental phonology directly using grapheme-to-phoneme conversion rules (Coltheart, Curtis, Atkins, & Haller, 1993).

We note here as an example of the theory-dependent nature of these estimates that the same assumptions do not make any sense in the PDP-based networks of reading (Plaut, McClelland, Seidenberg, & Patterson, 1996). In these frameworks pronunciation of all written forms is computed for the most part by a single pathway that converts orthography to phonology directly, with an additional, relatively minor, contribution from

Figure 2. Predicting 100-item picture naming accuracy from three estimates of semantic and phonological impairment. Predictions 1–3 used three different methods to derive estimates of semantic and phonological integrity (see text for details).

semantics that is especially critical for lower-frequency words with inconsistent spelling-to-sound correspondences. In these models there is neither a separate lexical (non-semantic) route for reading nor an abstract level of lexical representation.

A series of additional regression analyses were used to test whether these estimates of abstract lexical integrity improved the prediction of naming scores on the BNT (Table 3) and 100-item naming test (Table 4). In each case the additional predictor was added to the existing baseline regression model that included PPT and nonword reading as measures of semantic and phonological impairment (Regression 1). If the hypothesis were true then one or both of the measures of abstract lexical processing should be a significant additional predictor of naming over and above the baseline model.

TABLE 3
Testing the increases in predictive power associated with estimates of "lexical" integrity for BNT accuracy

Regression	Factors included in "baseline" model	Independent predictive power			Total variance explained by regression model
		Additional factor:	t	p	
1	PPT, nonword reading	—	—	—	65%
2	PPT, nonword reading	Word repetition	−1.62	0.12	70%
3	PPT, nonword reading, nonword repetition	Word repetition	0.06	0.96	71%
4	PPT, nonword reading, regular word reading	Exception word reading	−0.64	0.53	69%

Regressions 2 and 3 tested the effectiveness of adding lexical estimates derived from repetition performance. The improved fit of the model after word repetition was added to the equation (Regression 2) failed to reach significance. A stricter test of this estimate is given when both word and nonword repetition performance are added (Regression 3). In some frameworks (Franklin, 1989) there are three routes available for repeating words: (1) a semantically mediated route; (2) a direct route from input to output phonology; and, (3) a lexical (non-semantic) route. Nonwords are presumed to rely exclusively on the direct computation of output from input phonology (route 2). The integrity of the lexical route should be reflected, therefore, by the variance explained by word repetition over and above that already explained by nonword performance (in addition to the existing baseline equation). From Tables 3 and 4 (see Regression 3) it is clear that this estimate added no significant predictive power to the equation. Regression 4 tested the power of estimates derived from reading performance (by adding exception and regular word reading to the baseline equation). This is based on the assumption that in dual-route theories of reading, words with inconsistent/exceptional spelling-to-sound correspondences are largely reliant on the lexical route whereas regular words can be read by either pathway. Consequently exception over regular word performance should give an estimate of lexical integrity. Again this measure failed to add any significant predictive power to the baseline model.

TABLE 4
Testing the increases in predictive power associated with estimates of "lexical" integrity for 100-item naming test accuracy

Regression	Factors included in "baseline" model	Independent predictive power			Total variance explained by regression model
		Additional factor:	t	p	
1	PPT, nonword reading	—	—	—	55%
2	PPT, nonword reading	Word repetition	1.96	0.07	64%
3	PPT, nonword reading, nonword repetition	Word repetition	0.98	0.34	64%
4	PPT, nonword reading, regular word reading	Exception word reading	0.52	0.62	80%

Are certain types of naming error associated with semantic or phonological deficits?

Returning to the central hypothesis under consideration in this study, it is interesting to know whether semantic and phonological measures can predict not only naming accuracy but also the quality of the aphasic responses—i.e., can these two measures predict the rate of different types of naming error. A semantic impairment, for example, might lead to semantically related errors due to inaccurate or underspecified conceptual representations incorrectly activating semantically related word forms. It is also possible that underspecified conceptual representations would lead to omission errors because the reduced semantic input fails to activate any word forms sufficiently strongly to drive overt production. It is relevant to note here that omissions and semantically related responses predominate in the naming errors observed in semantic dementia (Lambon Ralph et al., 2001). Conversely phonological deficits should lead to phonological distortions of target word forms. With mild degrees of impairment patients might produce word and nonword responses with considerable overlap with the target form, whereas more severe deficits could lead to a much greater distance between response and target word.

The naming errors produced by each patient to the 100-item naming test were sorted into the following categories: *omissions*; *semantic*—including coordinate errors, which made up the vast majority of this category (e.g., *duck* → “swam”), superordinates (e.g., *beetle* → “insect”), and associate responses (e.g., *pig* → “pork”); *circumlocutions* (e.g., *pepperpot* → “something for shaking it out”, or partial information such as *ant* → “like a spider”); *phonologically related word and nonword errors*—in which the response contained at least half of the phonemes found in the target in any order (e.g., *pig* → “pick” or *cat* → “clat”); *phonologically unrelated word and nonword errors*—in which the response contained less than half the number of phonemes in the target (e.g., *star* → “easy” or *yacht* → “jus”); *other*—this included a very small number of other specific errors types such as visually related errors (e.g., *razor* → “hammer”) and gestures (e.g., *sock* → points to foot)—this type was produced rarely, so these errors were excluded from further analysis. The number of each error type for each patient is shown in Figure 3.

Table 5 shows the correlations between the various different error types and the three different measures of semantic and phonological integrity used earlier to predict naming accuracy. As it is unclear whether one should look at the errors in terms of raw number or percentage of total errors, we conducted the correlational analysis with both measures. In practice the pattern was nearly identical and thus for the sake of brevity we only show the results for the raw number of errors.

Three results follow the simple predictions noted earlier. Omissions were significantly correlated with degree of semantic impairment (the greater the semantic impairment, the higher the number/rate of omissions). Nonword responses, both phonologically related and unrelated, were correlated with the phonological impairment (those cases with a more severe phonological deficit produced a greater number of nonword responses). There were also three unexpected results. Semantic errors were found to correlate *negatively* with the degree of phonological ability; phonologically related word errors correlated positively, and unrelated errors correlated negatively, with semantic integrity. Given the very small numbers/rate of phonologically related and unrelated word errors (see Figure 3), the correlations should be treated with some caution and, therefore, we do not offer any specific explanations. There were a reasonable number of semantic naming

Figure 3. Error types for each patient.

TABLE 5
Correlations between naming errors (raw number) and measures of semantic and phonological integrity

	<i>Omissions</i>	<i>Semantic</i>	<i>Phonol-Related Word</i>	<i>Phonol-Related Nonword</i>	<i>Unrelated Word</i>	<i>Unrelated Nonword</i>	<i>Circumlocutions</i>
Semantic-1	-.603**	-.015	.435*	.294	-.762**	-.011	-.026
Semantic-2	-.606**	-.014	.443*	.283	-.757**	.014	.017
PPT	-.763**	-.034	.320	.283	-.563**	.097	.047
Phonological-1	-.380	.563**	-.359	-.631**	.171	-.503*	.012
Phonological-2	-.371	.491*	-.317	-.513*	.124	-.443*	.032
Nonword reading	-.392	.219	-.063	-.232	-.184	-.330	-.121

Semantic-1 and Phonological-1 refer to the two factor scores extracted from the principal component analysis of the patients' scores that excluded picture naming (see Table 2).

Semantic-2 and Phonological-2 refer to the two factor scores extracted from the PCA that excluded picture naming, word reading and repetition (see Table 2).

PPT = Pyramids and Palm Trees test.

Phonol-Related = Phonologically-related.

* $p < .05$, ** $p < .01$

errors. Post hoc consideration leads to the following possible explanation. The rate/number of observed semantic errors may go down as phonology becomes more impaired simply because potential semantic errors may be corrupted by phonological deficits before they are uttered. In those patients with little or no phonological impairment, semantic errors can be produced without disruption.

DISCUSSION

The primary aim of the present study was to test the proposal that single-word speech production (as measured by naming to confrontation) is built primarily on two key systems: semantic memory/conceptual knowledge and phonology. If correct, such a demonstration would add to growing evidence that specific language activities, like naming, reading, and verb transformations, are best thought of as arising from the nature of, and interplay between, primary systems such as semantics, phonology, and visual processing (Joanisse & Seidenberg, 1999; Patterson & Lambon Ralph, 1999; Plaut & Shallice, 1993).

The present study investigated the nature of anomia predominantly due to CVA in a case-series of 21 aphasic patients. The degree of anomia varied considerably across the group from very poor accuracy on simple tests of confrontational naming through to mild levels of word-finding difficulties. Like previous studies (Gagnon et al., 1997; Goodglass et al., 1997), we identified concurrent deficits in phonological processing in all but the mildest patient—as shown by impaired reading and/or repetition of words and nonwords. Four tests of comprehension yielded evidence for additional impairments of semantic memory, albeit relatively mild deficits in most cases. There was no evidence of modality differences in comprehension for these patients and the scores for the four assessments were highly intercorrelated. This suggests that abnormal comprehension performance, where it was observed, was due to a deficit within a single, central semantic system. Clearly the major factor in the clinical presentation of these patients was their phonological deficits but it is entirely possible that even mild semantic impairments may

prove to be critical too. For example, Lambon Ralph et al. (2001) found that mild comprehension deficits in patients with semantic dementia led to considerable degrees of anomia.

Having established that the anomia in this case series was accompanied by phonological and mild semantic deficits, we went on to investigate evidence for a causal link. Three different methods were used to derive estimates of each patient's semantic and phonological deficits. Two methods produced a weighted average score for semantic and phonological integrity by using a principal components analysis (PCA) to group the different clinical test scores. The third method relied on two specific test results (the nonverbal version of the Pyramids and Palm Trees test for semantics and nonword reading for phonology). Semantic and phonological estimates were then used to predict naming accuracy across the case-series for the Boston Naming Test (BNT) and the 100-item naming test. These estimates explained a large portion of the variance in the patients' naming scores. The good fit produced by these regression models was confirmed by case-diagnostic analyses. Across all models, there was never more than 1/21 patients for whom the observed accuracy deviated significantly from the expected score. These results suggest that naming performance can be thought to reflect the processing of two "primary" systems—semantics and phonology—such that, allowing for fluctuations due to individual differences and test reliability/validity, picture naming can be predicted highly accurately with independent measures of semantic and phonological integrity. There was some limited evidence that the rate of different types of finding errors could also be predicted from the same measures.

This view of speech production would seem to question other models that include additional levels of abstract lexical representation. We shall highlight potential similarities and differences between these proposals later. It is first worth noting, however, that we did attempt to find direct evidence for this kind of abstract lexical representation in the present study, despite the difficulty of knowing how to measure the status of such representations without using the task of interest—i.e., to produce an independent estimate. If representations or processes are assumed to be dedicated to a specific process then they cannot, by definition, be probed independently. For example, if one found patients without phonological deficits whose anomia was greater than that expected by the degree of their semantic impairment (as seen in some cases of semantic dementia: e.g., Graham et al., 1995; Lambon Ralph et al., 2001) then this pattern could be used to argue in favour of a set of lexical representations dedicated to speech production. This leaves a potentially unfalsifiable position, however—if patients' degree of anomia is greater than that expected they are assumed to have an additional lexical deficit, if it is commensurate with their semantic deficit they are assumed not to have a lexical deficit. Such a circular position can only be broken by the use of independent measures of the processes assumed to be involved in an activity. In this study, we attempted to derive two independent estimates of lexical processing, although we note that they are dependent on specific theoretical assumptions about repetition and reading aloud (see earlier). If there are intermediate, abstract lexical representations that could be damaged in aphasia, then estimates of their integrity should significantly improve the prediction of naming performance over and above that already given by measures of semantic and phonological processing. The analyses containing these additional estimates failed to improve the prediction of naming accuracy significantly.

The remainder of the Discussion is split into four sections. In the first, we consider in more detail the nature of the case-series methodology adopted here and compare it with single-case and group investigations. In the next three sections we will contrast the

approach adopted here with two well-known speech production frameworks and then consider our proposal in a little more detail.

Case-series methodology

In this section we will stand back briefly from the theoretical issues and consider the nature of the methodology adopted here. One can think of the case-series methodology as falling in between single-case and group studies. Each of these techniques has benefits and disadvantages and we will rehearse some of them here (for a fuller and more complete discussion of these issues, we refer the reader to Shallice, 1988). Data arising from single-case studies have provided fascinating and important insights into mental performance including the structures that underpin language. There are certain questions, however, that are difficult or impossible to answer with results of individual cases. The one to be highlighted here revolves around the following question—how large does a performance difference between two sets of materials or two tasks have to be before we make the inference that the behaviour is underpinned by a separable cognitive process or representation? This is not a trivial issue and is complicated by at least three factors: impairment severity, measurement noise, and individual differences.

Lets imagine that we are studying a patient with an impairment to a cognitive system A (e.g., semantic memory) and relating this to behaviour Z (e.g., picture naming). In this patient we find that performance on Z is worse than accuracy on a range of other tasks designed to assess the integrity of system A. Does this mean that there is another system B (e.g., abstract lexical representations) that normally supports behaviour Z that is impaired in this patient? The answer to this question depends on a number of things but includes the following: What is the relationship between the severity of impairment to system A and the resultant behaviour Z? That is to say the patient's performance needs to be compared to some estimate of what we would expect it to have been. Of course, to plot out a function that relates impairment severity to behaviour requires a number of observations either in the form of multiple patients tested on the same assessments (i.e. a case series) or, possibly, repeated longitudinal investigations of a patient with a progressive disease. It is impossible to consider such a function by studying a single patient with a non-progressive disease.

Two other factors complicate the interpretation of the apparent difference between A and Z. The first is measurement noise—the results on all assessments have some amount of measurement error associated with them. Consequently for each level of impairment to A (e.g., semantic impairment) there will be a range of expected scores for behaviour Z (e.g., picture naming). With case-series or group data it is possible to estimate, as we have attempted here, both the expected level of performance and its associated confidence interval. Finally, there is the related issue of individual differences, which can either add to the measurement noise (e.g., subjects may vary in their premorbid abilities and/or their familiarity with the chosen stimuli) or alter the function that relates the underlying cognitive systems to the target behaviour (for a concrete example using a implemented computational model, see Plaut, 1997). Again, at least with respect to individual differences leading to additional measurement noise, the ability to compute expected scores and their confidence intervals provides a tangible method by which differences in performance can be categorised into those that fall within the expected range of scores versus those that require further explanation. For the latter cases it is possible that the difference either is due to the impairment of an additional system (e.g., abstract lexical representations) or reflects an individual difference that has changed the underlying

function. Both situations require further investigation to provide independent evidence for either (a) the status of the new, additional system (e.g., the integrity of abstract lexical representations), or (b) a formal description and measurement of the individual difference.

While we have concentrated so far on the benefits of case-series over single-case studies, it should be noted that they have some drawbacks too. Case-series, like group studies, need to have some method of selecting appropriate patients and thereby producing a group of cases that vary quantitatively (e.g., degree of semantic and phonological impairment) rather than qualitatively. One could argue that combining cases in this way leads to a situation in which evidence for very specific yet important aspects of the underlying cognitive machinery are missed (e.g., evidence for specific aspects of elaborate speech production models such as those described by Levelt and colleagues). If the patients described here did, in fact, have a wide variety of specific impairments to different parts of a speech production system—that is to say the group were very heterogeneous—it seems very unlikely that we would have found such strong evidence to relate their naming performance to two measures of generalised phonological and semantic integrity (the principal component analysis would fail to extract any meaningful factors and/or explain only a minimal amount of the variance in patient scores). An additional, practical drawback to case-series and group studies is in the simple fact that more than one case is required—something that can be hard to achieve with rare disorders (Shallice, 1988). Clearly, given the different advantages and drawbacks of single-case and case-series studies, there is an important role for both in neuropsychological research.

Single lexicon models of speech production

There are a number of proposals in which the activation of phonology from conceptual knowledge is mediated by a set of lexical representations (Caramazza, 1997; Dell et al., 1997; Morton, 1985). If one assumes that the principal role of these lexical representations is in some way to make available the sound representations of familiar words (i.e., segmental phonology, syllable structure, stress pattern, etc.) then they may be functionally equivalent to the lexemes in two-stage models (see later). It is also possible that there may be functional equivalence, at this level, with PDP models that do not include localist representations at all (Plaut & Kello, 1999; Plaut & Shallice, 1993). In effect all these models produce a “phonological space” (Butterworth, 1989, 1992) in which the positioning of representations is governed by the phonological similarity between word forms. This is achieved in models with localist representations, for example, by interactions between lexical nodes and phonemes (Dell, 1986). Not only are target lexical representations reinforced by this interaction (see later for a fuller description of this model) but phonologically similar forms will receive partial activation, which is a function of their phonological similarity with the target—i.e., there are cohort effects. Some models with distributed representations include a set of “clean up” units that are fully interconnected to the output layer that encodes patterns of phonological activation (Harm & Seidenberg, 1999; Plaut et al., 1996; Plaut & Shallice, 1993). In combination the two sets of units are able to represent high-order statistical properties of phonemic co-occurrences that include patterns of phonology corresponding to whole word forms. Again similar sounding words will have similar representations within this phonological space.

Before moving onto consideration of two-stage models of lexical access, we will compare our proposal with the model described by Dell and colleagues (Dell, 1986; Dell

& O'Seaghdha, 1992; Dell et al., 1997). The architecture of this implemented model is as follows: a concept is represented in terms of a collection of semantic features; when these features are turned on (either by a picture or an intended "message") activation spreads to a lexical layer. This causes a gradual rise in lexical activation for a number of semantically related words, although the target item will typically be the most prominent as it corresponds most closely to the pattern of semantic features. Partially activated lexical units send activation back to the semantic feature level (the model is interactive) and on to a phoneme layer (activation cascades). Excitation reverberates in the model until, after a set number of processing steps, the most activated lexical entry is selected. As Dell et al. (1997) note, this lexical selection process is equivalent to the first step of the two-stage models described later (in fact they call this part lemma access), although in this version but not the others, the model has cascading and interactive activation. Once a unique lexical unit is selected, a jolt of activation is applied to this word. Excitation passes up to and back from the word's associated semantic features and down to and up from the phonological layer. This second process (phonological access) continues until a certain number of processing cycles have occurred, at which time the most highly activated phoneme units are selected.

Dell et al. (1997) used this architecture in an impressive attempt to fit the individual naming data of a case-series of 23 aphasic patients. The model was first set up to match naming data collected from two groups of control subjects. Then the model's performance was varied via adjustment of two global parameters, p , the connection strength and, q , the decay rate (i.e., changes to these two key parameters were made throughout the model—the globality assumption). By using these two factors, Dell et al. were able to simulate very closely the performance of 21 of the 23 aphasic patients. We suspect that much of the data for the present case-series of 21 patients could also be fitted by this model, although omission errors were quite common in our group of patients—which proved to be problematic for 2/23 cases reported by Dell et al.

In our view, it is important to note that underlying motivations behind explanations of aphasic naming data adopted by Dell et al. (1997) and in our proposal seem to be very similar. The central notions are that there are major similarities across patients and detailed aspects of each individual patient's performance can be captured by variation in a small number of parameters. For Dell et al. these two parameters are connection strength between units and decay rate (how quickly activated units return to their resting level). In our proposal there are also two parameters—the degree of disruption to semantic and phonological systems. Perhaps the key difference is that we assume that the primary systems, which underpin language activities including speech production, can be separately damaged—i.e., we have not adopted the globality assumption. We note that in a recent paper, Foygel and Dell (2000) have also used a new two-parameter method that assumes damage to two specific elements of the speech production system (deficits in the semantic–lexical links or the lexical–phoneme connections). This is probably very similar to our own position.

Two-stage lexical access models

One of the most influential models of normal speech productions is that proposed by Levelt (1989, 1992). In two-stage lexical access models, speech production is split into a number of hierarchically arranged, discrete processing steps that enable the speaker to move efficiently from "intention" to "speech". We shall limit ourselves to considering the steps required to move from intention to a phonological form, although this

intermediate level of phonological representation then has to be encoded into phonetic units and eventually into motor articulation plans (see Levelt, 1989). We begin at the point at which the speaker (either for picture naming or for spontaneous speech) formulates a ‘message’, or in more standard neuropsychological terms, activates the appropriate semantic or conceptual representation (which are conceived of in terms of localist representations of each individual concept in the latest formulation of the model: Levelt et al., 1999). Semantics then act as input to a two-stage lexical access process. In the first part, the lemma that corresponds to the conceptual representation is retrieved. This represents one single word in abstract form that specifies syntactical information (e.g., lexical gender). At this stage, no sound information is available. This lemma, in turn, activates its corresponding lexeme, which ‘releases’ or ‘makes available’ the metric structure and stress pattern along with the appropriate phonological units. Note that, although stored within one unit, metric structure and phonological segments are released separately and have to be recombined later (segment-to-frame association). Having arrived at this stage in the sequence, the speaker then has a phonologically encoded word ready for the additional processing required for eventual articulation.

This framework has typically been applied to results from speech production experiments in normal subjects (e.g., Jescheniak & Levelt, 1994; Levelt et al., 1991), the literature on normal speech errors (for a review see Meyer, 1992), or the tip-of-the-tongue (ToT) phenomenon in normal subjects (e.g., Meyer & Bock, 1992). Two-stage models are yet to have the same impact on neuropsychology although there are some notable exceptions (Butterworth, 1989, 1992; Nickels, 1997). This tends to make it hard to relate aphasic naming data to these models, but it seems fairly likely that the 21 patients in the present case-series could be considered individually within this elaborate framework, and their data explained by one or more impairments within the multiple levels of representation. For example, those patients who make semantic errors would have damage to the conceptual layer or to the process involved in lemma selection, in which the target lemma is selected from a cohort of semantically similar items. Phonological errors could arise from failed segment-to-frame association and omission errors from insufficient activation within lemma or lexeme access.

Such a process would lead to the same problem noted earlier—that of establishing an independent way in which to verify the nature of the damage to one or more of these systems, or preferably in which to make falsifiable predictions about the patients’ performance. This might be possible if these models were extended from consideration of speech production alone to other language activities that require verbal output such as reading aloud or repetition (for a combination of repetition, spoken word comprehension and speech production, see Nickels, 1997, p.94). In the present study, we assumed that within this kind of framework, abstract lexical forms, or lemmas, could be critically involved in the reading of words with exceptional spelling-to-sound correspondences and/or repeating real words. The estimates derived, with these assumptions, for lexical integrity did not improve the prediction of naming across the case-series. There was, therefore, no positive evidence from the present study for such lemma representations, although we remain open to suggestions for other ways in which to measure or estimate lemma status.

The semantic-phonology hypothesis

In this final section, we shall sketch out a little more detail with respect to the proposal that single-word speech production is simply a reflection of the nature and interplay

between semantics and phonology. As noted in the Introduction, one of the motivations for the present study was to test the wider hypothesis that speaking, repeating, reading, and comprehending single words can all be understood with reference to a core set of “primary” systems such as semantics and phonology. This approach assumes that characteristics of particular tasks in normal and impaired populations reflect the underlying nature of the primary systems and the interaction between them. This approach has already been explored for reading aloud (Plaut et al., 1996; Seidenberg & McClelland, 1989) and for transforming a verb from present to past tense (Joanisse & Seidenberg, 1999; Patterson et al., 2001; Rumelhart & McClelland, 1986).

If we are to adopt this approach, however, there are a number of basic phenomena that require explanation:

- (a) The mapping between meaning and sound is arbitrary.
- (b) Lexical access is rapid.
- (c) The adult vocabulary is very large.
- (d) Speech production seems to involve an early stage of semantic activation followed by late phonological activation.
- (e) There are two broad categories of aphasic errors and normal “slips-of-the-tongue”.
- (f) Both normal and aphasic speakers sometimes have partial phonological knowledge about ToT words.
- (g) There must be a locus for the impact of syntactic knowledge in speech production (this locus is the lemma level in Levelt’s model).

A reason for positing at least one intermediate layer of localist word representations between semantics and phonology is that the mapping from meaning to sounds is an arbitrary one. Typically words of similar meaning have very different phonological forms (e.g., cat, dog), while those with similar sounds have disparate meanings (e.g., cat, mat). Most theorists (see, for example, Dell et al., 1997) seem to view this arbitrary mapping as an unfortunate computational problem of natural language. If arbitrariness is a bad thing, however, why have natural languages evolved in this fashion rather than to produce a closer alignment between sound and meaning? It is possible that this way of implementing the relationship between meaning and phonology may have useful functional/computational consequences.

Butterworth (1992) and Levelt (1992) both note that adults have a very large expressive vocabulary from which single items have to be retrieved exceedingly rapidly to produce the high speech rate observed in spontaneous conversation. A lexical layer containing a single unit for each word that is accessed in parallel would seem to provide a solution to all of points (a)–(c) just listed. However, rapid lexical access might also result from an arbitrary relationship encoded on direct but interactive connections between semantics and phonology. Consider the following example: if the intended target word is *cat*, semantic activation will gradually build up to form a representation of the meaning. This semantic representation will begin to activate the target phonology of *cat* and, to a lesser degree, the phonological representations of other semantically related items (*dog*, *fox*, *lion*, *tiger*, etc). That is to say a semantic cohort (to borrow a term from Levelt et al., 1991) will be partially activated within the phonological component of the system, and will interact with the semantic activation. As the relationship between meaning and sound is arbitrary, competitors of the target word will typically have no phonological overlap with the target. Thus, as soon as the activation of /kæt/ even begins to lead, the cohort of

potential words will be reduced rapidly to a very small number, most often to only one item. In our example, when the semantic representation of “a furry four-legged creature with claws” combines with /k/, there is only one item left. Compare this to the real problems that would be inherited if the system were set up so that meaning correlated more closely to sound (i.e., if *dog*, *fox*, *lion*, *tiger*, etc. were more like *cat*, *cap*, and *cab*). Having activated a semantic representation, it would be difficult to settle rapidly on the correct phonology of the desired target and thus to avoid a semantic error. This problem would also be found in reverse, i.e., for comprehension. Poor articulation, impaired hearing, or background noise would lead to a large number of phonological (and, if related, hence semantic) comprehension errors. So rather than treating the arbitrary relationship as an unfortunate mistake of human language, it is possible that we should view it as the key to another computational problem, that of rapid “lexical access”.

What else follows from this arbitrary relationship? In order to commence activation of phonology, the semantic input would have to be very “well-formed” both in terms of quality and quantity of input. This means that the time course in “lexical access” will be primarily semantic to begin with (while the meaning is formed). Although this gradual build up in semantic activation cascades to phonology, it has little impact until the activation is strong and precise. It is only at this point that the rapid cohort reduction and phonological “encoding” will begin in earnest—that is, significant phonological activation will tend to occur only when semantic activation begins to reach asymptote (NB: this seems functionally to be rather like a threshold of semantic activation which needs to be exceeded before a phonological representation is released: cf. the logogen model, Morton, 1985). This, of course, provides an alternative mechanism for the apparent “two-step” activation dynamic noted in point (d) described earlier.

In the same way, it seems likely that speech errors, point (e), would also appear to reflect a two-step time-course (just as they are proposed to do in the two-stage or interactive models described earlier), even if the process is in fact cascaded and gradual. Some will occur in the “semantic” phase (errors of meaning) while others will arise in the “phonological” stage (errors of sound). Tip-of-the-tongue phenomena, point (f), could reflect the information that is available when semantic input to phonology is insufficient for the rapid cohort reduction to occur (or to be completed). At this point a “clear” semantic representation may be formed (thus subjects “know” the intended word) but the sound information will only be partial and might include the approximate word-shape and/or the first sound (Lambon Ralph et al., 2000).

This leaves us with point (g)—if there are no intermediate word units (lemmas) between semantics and phonology, how are syntactic aspects of individual words encoded? This issue has been considered in light of other neuropsychological data by Caramazza (1997). In his “Independent Network” (IN) model, which also suggests that a set of lemma-like representations are unnecessary, Caramazza proposes a separate lexical-syntactical network that is linked to lexical-semantic, lexical-phonological, and lexical-orthographic systems. Each syntactic property will have a different pattern of connections with the semantic and phonological systems depending on whether meaning and/or sound predicts grammatical form (1997, pp. 194–95). In languages such as Italian and French in which names of inanimate objects have gender, for example, gender is generally unrelated to meaning but is highly (though not perfectly) predictable from phonology. This is reflected in the IN model as connections running from lexemes (“P-lexemes” in Caramazza’s terminology) to the gender nodes, but no link from gender to semantics. A similar approach can be taken from a PDP perspective. Grammatical and morphological aspects of language can be encoded within networks that learn about the

internal regularities of phonology and semantics/context (for some recent reviews of this approach, see Allen & Seidenberg, 1999; Seidenberg, 1997). Other structural aspects of speech that are normally encoded in terms of specific units can be simulated in models that learn about the regularities within phonology itself (Dell, Juliano, & Govindjee, 1993).

In conclusion, at this stage of model development, it is probably the case that almost all models of speech production can explain the data presented here. The main difference between the accounts is the level of explanation adopted. It is our claim that single-word speech production tasks like naming are best explained at a computational level by the interaction between two primary systems—semantics and phonology.

REFERENCES

- Allen, J., & Seidenberg, M.S. (1999). The emergence of grammatically in connectionist networks. In B. MacWhinney (Ed.), *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Behrmann, M., Nelson, J., & Sekuler, E. (1998a). Visual complexity in letter-by-letter reading: "Pure" alexia is not so pure. *Neuropsychologia*, 36, 1115–1132.
- Behrmann, M., Plaut, D.C., & Nelson, J. (1998b). A literature review and new data supporting an interactive account of letter-by-letter reading. *Cognitive Neuropsychology*, 15, 7–52.
- Bozeat, S., Lambon Ralph, M.A., Patterson, K., Garrard, P., & Hodges, J.R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38, 1207–1215.
- Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.
- Butterworth, B. (1992). Disorders of phonological encoding. *Cognition*, 42, 261–286.
- Butterworth, B., Howard, D., & McLoughlin, P. (1984). The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia*, 22, 409–426.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208.
- Chertkow, H., Bub, D., Deaoud, C., & Whitehead, V. (1997). On the status of object concepts in aphasia. *Brain and Language*, 58, 203–232.
- 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.
- Croot, K., Patterson, K., & Hodges, J.R. (1998). Single word production in nonfluent progressive aphasia. *Brain and Language*, 61, 226–273.
- Dell, G.S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G.S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149–195.
- Dell, G.S., & O'Seaghdha, P.G. (1992). Stages of lexical access in speech production. *Cognition*, 42, 287–314.
- Dell, G.S., Schwartz, M.F., Martin, N. Saffran, E.M., & Gagnon, D.A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–838.
- Farah, M.J., Stowe, R.M., & Levinson, K.L. (1996). Phonological dyslexia: Loss of a reading-specific component of the cognitive architecture? *Cognitive Neuropsychology*, 13, 849–868.
- Foygel, D., & Dell, G.S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, 43, 182–216.
- Franklin, S. (1989). Dissociations in auditory word comprehension: Evidence from 9 fluent aphasic patients. *Aphasiology*, 3, 189–207.
- Franklin, S., Turner, J., Lambon Ralph, M.A., Morris, J., & Bailey, P.J. (1996). A distinctive case of word meaning deafness? *Cognitive Neuropsychology*, 13, 1139–1162.
- Franklin, S., Turner, J.M., & Ellis, A.W. (1992). *The ADA comprehension battery*. London: Action for Dysphasic Adults.
- Friedman, R.B. (1996). Recovery from deep alexia to phonological alexia: Points on a continuum. *Brain and Language*, 52, 114–128.
- Gagnon, D.A., Schwartz, M.F., Martin, N., Dell, G.S., & Saffran, E.M. (1997). The origins of formal paraphasias in aphasics' picture naming. *Brain and Language*, 59, 450–472.

- Goodglass, H., Wingfield, A., Hyde, M.R., Berko Gleason, J., Bowles, N.L., & Gallagher, R.E. (1997). The importance of word-initial phonology: Error patterns in prolonged naming efforts by aphasic patients. *Journal of the International Neuropsychological Society*, 3, 128–138.
- Graham, K.S., Hodges, J.R., & Patterson, K.E. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia. *Neuropsychologia*, 32, 299–316.
- Graham, K.S., Patterson, K., & Hodges, J.R. (1995). Progressive pure anomia: Insufficient activation of phonology by meaning. *Neurocase*, 1, 25–38.
- Graham, N.L., Patterson, K., & Hodges, J.R. (2000). The impact of semantic memory impairment on spelling: Evidence from semantic dementia. *Neuropsychologia*, 38, 143–163.
- Harm, M., & Seidenberg, M. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, 106, 491–528.
- Hodges, J.R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115, 1783–1806.
- Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Bury St. Edmunds, UK: Thames Valley Test Company.
- 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.
- Joanisse, M.F., & Seidenberg, M.S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 7592–7597.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1976). *The Boston Naming Test*. Boston: Veteran's Administration.
- Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Lambon Ralph, M., McClelland, J., Patterson, K., Galton, C., & Hodges, J. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience*, 13, 341–356.
- Lambon Ralph, M.A. (1998). Distributed versus localist representations: Evidence from the study of item consistency in a case of classical anomia. *Brain and Language*, 64, 339–360.
- Lambon Ralph, M.A., Ellis, A.W., & Sage, K. (1998). Word meaning blindness revisited. *Cognitive Neuropsychology*, 15, 389–400.
- Lambon Ralph, M.A., Sage, K., & Ellis, A.W. (1996). Word meaning blindness: A new form of acquired dyslexia. *Cognitive Neuropsychology*, 13, 617.
- Lambon Ralph, M.A., Sage, K., & Roberts, J. (2000). Classical anomia: A neuropsychological perspective on speech production. *Neuropsychologia*, 38, 186–202.
- Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–45.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42, 1–22.
- Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122–142.
- Meyer, A.S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition*, 42, 181–211.
- Meyer, A.S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, 20, 715–726.
- Morton, J. (1985). Naming. In S. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia*. Edinburgh: Churchill Livingstone.
- Neville, H.J., & Bavelier, D. (1998). Neural organization and plasticity of language. *Current Opinion in Neurobiology*, 8, 254–258.
- Nickels, L. (1997). *Spoken word production and its breakdown in aphasia*. Hove, UK: Psychology Press.
- Patterson, K., & Hodges, J.R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia*, 30, 1025–1040.
- Patterson, K., & Lambon Ralph, M.A. (1999). Selective disorders of reading? *Current Opinion in Neurobiology*, 9, 235–239.
- Patterson, K., Lambon Ralph, M.A., Hodges, J.R., & McClelland, J.L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia*, 39, 709–724.
- Patterson, K., Suzuki, T., & Wydell, T.N. (1996). Interpreting a case of Japanese phonological alexia: The key is in phonology. *Cognitive Neuropsychology*, 13, 803–822.

- Pinker, S. (1999). *Words and rules: The ingredients of language*. London: Weidenfeld & Nicolson.
- Plaut, D.C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, 12, 765–805.
- Plaut, D.C. (1999). Systematicity and specialization in semantics. In D. Heinke, G.W. Humphreys, & A. Olsen (Eds.), *Connectionist models in cognitive neuroscience: Proceedings of the Fifth annual Neural Computation and Psychology Workshop*. New York: Springer-Verlag.
- Plaut, D.C., & Kello, C.T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- 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 of connectionist neuropsychology. *Cognitive Neuropsychology*, 10, 377–500.
- Riddoch, M.J., & Humphreys, G.W. (1992). *The Birmingham Object Recognition Battery (BORB)*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Rumelhart, D.E., & McClelland, J.L. (1986). On learning the past tense of English verbs. In J.L. McClelland, D.E. Rumelhart, & the PDP Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2, Psychological and Biological Models*. Cambridge, MA: MIT Press.
- Seidenberg, M., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Seidenberg, M.S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, 275, 1599–1603.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Snodgrass, J.G., & Vanderwart, M. (1980). A standardised set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Ullman, M.T., Corkin, S., Coppola, M., Hickok, G., Growdon, J.H., Koroshetz, W.J., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Language*, 9, 266–276.