

## **A Distinctive Case of Word Meaning Deafness?**

Sue Franklin

*University of Newcastle, UK*

Judy Turner

*University of Reading, UK*

Matthew A. Lambon Ralph

*University of York, UK*

Julie Morris

*Birkbeck College, London, UK*

Peter J. Bailey

*University of York, UK*

We report the case of a patient, DrO, who has impaired auditory comprehension but intact written comprehension. His ability to repeat words that he cannot comprehend, and his good performance on auditory lexical decision, suggest that DrO has an impairment of the mappings between the lexical representations of spoken words and their corresponding semantic representations; a word meaning deafness.

DrO's ability to understand heard words depends on their imageability and length, such that he is worse at comprehending words that are both short and abstract. This is interpreted in terms of phonological and semantic distinctiveness.

### INTRODUCTION

Word meaning deafness is a comprehension deficit specific to the auditory modality: Written comprehension is unimpaired. It is distinct from word sound deafness in that the ability to repeat is intact (Ziehl, 1896). Allport and Funnell

---

Requests for reprints should be addressed to Sue Franklin, Department of Speech, King George VI Building, University of Newcastle, Newcastle upon Tyne, NE1 7RU, UK.

We would like to thank Dinah Ledger and Felicity Lowe for referring DrO and are very grateful to DrO himself for tolerating such extensive testing. We greatly appreciate Andrew Ellis's comments on this paper. This research was supported by the Medical Research Council.

(1981) considered that word meaning deafness "is of great theoretical significance" because it suggests that comprehension of written words cannot be directly dependent on recoding to phonology. If this were the case, written comprehension would always be at least as impaired as auditory comprehension (assuming that a peripheral auditory impairment is ruled out because of intact repetition).

The existence of word meaning deafness is still a matter of controversy. Although the condition was described over a hundred years ago (Bramwell, 1897), studies of other patients with word meaning deafness are extremely rare, and generally not described in any detail. The term "word meaning deafness" has recently been used to describe a more specific deficit; a post-lexical problem in accessing semantics from speech (Franklin, Howard, & Patterson, 1994; Kohn & Friedman, 1986). This is of particular interest for contemporary theorising in that such a deficit is predicted by many of the functional architectures used in cognitive neuropsychology that specify a functionally encapsulated, lexical level of auditory processing (Patterson & Shewell, 1987).

The argument for the existence of this more narrowly defined type of word meaning deafness is based on the distinction between it and a specifically lexical deficit, termed "pre-access word meaning deafness" by Kohn and Friedman (1986) or "word form deafness" by Howard and Franklin (1988) and Franklin (1989). Kohn and Friedman contrasted two patients, HN and LL. Both patients, when unable to comprehend a heard word, could understand it once they had written it. HN was even able to write irregular words to dictation, such as "knee" and "thigh," when he couldn't understand them. There were, however, rather few examples of such instances given in the paper. Kohn and Friedman argued that HN must have had access to lexical knowledge in order to derive the correct spellings in these cases and that he was therefore a case of "post-access" word meaning deafness. LL tended to regularise irregular words in orally spelling heard words (e.g. "cough"  $\rightarrow$  COFF), which Kohn and Friedman argue suggests a "pre-access" word meaning deafness, since LL is using phonological rather than lexical knowledge to spell the words he hears. Unfortunately, Kohn and Friedman fail to rule out the possibility of LL's writing errors being due to a later processing deficit.

Howard and Franklin's description of patient MK provided a more detailed account of a "pre-access" or "word form" deafness patient (Howard & Franklin, 1988). Although able to discriminate heard phonemes, this patient was impaired at auditory lexical decision tasks, and appeared to "mis-hear" words as other words phonologically related to the target (e.g. "garden" for "pardon"). MK's ability to understand words was affected by phoneme length. This "reverse length" effect (i.e. that MK was more likely to understand a word if it were longer) was interpreted by Howard and Franklin as the consequence of an impairment of access to a phonologically ordered lexical system, since longer

words tend to be phonologically less confusable (for example "cat" has many more similar sounding words than does "crocodile").

By contrast, patient DRB (Franklin et al., 1994) was described as being word meaning deaf ("post-access" using Kohn and Friedman's terminology) on the basis of his good performance on tests of both auditory lexical access and written word comprehension. As well as performing within the normal range on tests of auditory lexical decision, DRB was able to reject phonologically related words as inappropriate for pictures (for example rejecting "iron" as an unsuitable name for a picture of a lion). By contrast, he made errors in the same task when semantically related distractors were present (for example "tiger" for the lion). In light of his good lexical decision performance, and the tendency for his errors in comprehension to be semantic rather than phonological, Franklin et al. argued that DRB's comprehension deficit was post-lexical. It was demonstrated, using a number of tests that required semantic processing, that he made errors when words were presented auditorily but was always within the normal range when the same words were presented visually. This pattern of results appeared to establish the deficit as post-lexical, but specific to auditory processing; a word *meaning* deafness. It was further demonstrated that DRB was less likely to comprehend auditorily presented words if they were abstract or of low imageability.

As his word repetition was also impaired, DRB could not be considered a case of word meaning deafness in the traditional sense. Franklin et al. suggested that his poor repetition was the result of an additional impairment, which made him unable to recode directly from input to output phonology (he was, for example, quite unable to repeat nonwords). It is, however, possible that his repetition deficit reflected some underlying early phonological impairment, which the minimal pair and lexical decision tests had been too insensitive to detect. If this were the case, DRB's apparent word meaning deafness could be seen rather as a mild form of word sound deafness (a more detailed consideration of this argument may be found in Franklin et al., 1994). It would still be of considerable interest, therefore, to find a case of word meaning deafness where the patient is reliably able to repeat the words he or she is unable to comprehend.

This paper describes such a patient, DrO, who, it is argued, has intact written comprehension and good auditory lexical access but impaired auditory comprehension of words. Unlike DRB, DrO is generally able to repeat the words he is unable to comprehend. Like the majority of patients who are "word deaf" and where comprehension of abstract words has been tested (Franklin, 1989), DrO's comprehension impairment is sensitive to word imageability. Although we will argue that DrO's word deafness constitutes a post-lexical impairment, unlike that of MK, who was earlier described as word form deaf, both DrO and MK were more impaired at comprehending short words.

## CASE HISTORY

DrO is a retired lecturer in sociology who was born in Norway in 1931. He moved to America at the age of 6 and since then has lived either in America or Britain. He has spoken Norwegian only occasionally in his adult years. In September 1991, at the age of 60, he had a CVA. A CT scan revealed a left middle cerebral infarct.

On initial assessment DrO was described as using jargon and showing severely impaired comprehension. When testing began at the end of 1992, DrO's speech was greatly improved, being fluent with only mild word-finding difficulties. Over the period of testing his naming continued to improve, although his auditory comprehension remained unchanged. He was able to read, but unable to write because of severely impaired spelling. He relied very heavily on contextual cues to understand what was being said to him, and was aware that auditory comprehension was his major problem. Pure tone audiometry indicated that DrO's hearing was normal. Although audiometric testing was carried out in a non-soundproofed room, the threshold for the speech frequencies was 0–5dB for the left ear and 0–15dB for the right ear (well within normal limits).

## SPOKEN AND WRITTEN COMPREHENSION

Three tests were administered to assess DrO's ability to access word meanings: an abstract word-to-picture matching test, a semantic association test, and a synonym judgements test. All tests were given on two separate occasions with several months intervening. On the first occasion the words were presented in written form. On the second occasion they were presented auditorily. If DrO had a central semantic deficit, the written and spoken versions of the tests should be equally impaired. If, however, he was able to comprehend the written but not the spoken words, he could not, by implication, have a semantic deficit, but must rather have an auditory comprehension deficit; that is, some form of word deafness.

### The Abstract Word-to-picture Matching Test (Shallice & McGill, unpublished)

This test required the patient to select the picture corresponding to the stimulus word from a choice of four pictures. Half the words used are concrete. For these words the match is quite straightforward, for example matching the word badge to a picture of a badge. For the abstract words, however, there is a rather more indirect relationship between the word and the picture (e.g. caution  $\square$  someone using a zebra crossing; democracy  $\square$  a group of people all with their hands raised).

### The Semantic Association Test (PALPA) (Kay, Lesser, & Coltheart, 1992)

In this test the patient heard or saw a word (e.g. "fog") and had to choose the word closest in meaning to it from a choice of four written words. The four written words were: a word closely related in meaning (= the correct response, e.g. "mist"); a word more distantly related in meaning (e.g. "steam") and two unrelated words (e.g. "bolt," "lock"). As there was a semantically related distractor for each item, this test requires quite specific semantic information. Half the items in this test used words that were highly imageable; the other half comprised words that have low imageability ratings.

### The Synonym Judgements Test (ADA Comprehension Battery) (Franklin, Turner, & Ellis, 1992)

This test required the patient to judge whether two words had the same meaning or not. This test compared performance on 40 high-imageability word pairs (e.g. "basin–bowl," "rug–stomach") with performance on 40 low-imageability word pairs (e.g. "genuine–authentic," "law–charlatan"), balanced for word frequency. The remaining pairs in the test compared 40 high-frequency word pairs with 40 low-frequency word pairs, balanced for word imageability. Half the pairs used were synonyms; all conditions were balanced for phoneme length. For the auditory version the stimuli were presented on DAT tape over headphones, using a female voice.

## Results

The results for all three tests are shown in Table 1. DrO made only two abstract word errors on the written version of the abstract word-to-picture matching test. This was well within normal performance on this test, given that Warrington

TABLE 1  
DrO's Performance on Tests of Word Comprehension

|                                            | <i>Spoken</i> | <i>Written</i> |
|--------------------------------------------|---------------|----------------|
| <i>Abstract Word-Picture Matching Test</i> |               |                |
| Concrete                                   | 28/30         | 30/30          |
| Abstract                                   | 21/30         | 28/30          |
| <i>The Semantic Association Test</i>       |               |                |
| High-imageability words                    | 13/15         | 13/15          |
| Low-imageability words                     | 7/15          | 13/15          |
| <i>Synonym Judgements</i>                  |               |                |
| High imageability                          | 35/40         | 40/40          |
| Low imageability                           | 25/40         | 39/40          |
| High frequency                             | 35/40         | 40/40          |
| Low frequency                              | 35/40         | 40/40          |

(1981) reported that the *mean* normal performance for abstract items was four errors. The two errors made were “inducement” (DrO selected a picture where a man is wielding a whip in front of a chain gang as opposed to a picture of a man with a carrot seated on a donkey) and “reminder” (DrO selected a picture of dentistry rather than a picture of a man in front of a gravestone). In the spoken word condition, DrO made two errors on the concrete word items but was significantly worse with the abstract words (Fisher Exact,  $z = 1.98$ ,  $P < .05$ ). The abstract auditory score was significantly worse than for the written version (McNemar test,  $P < .01$ ).

The same pattern was obtained for the semantic association test. When DrO was given this test in the written form, he made only four errors (two on each of the high- and low-imageability sets). Normal data collected by one of the authors found that 10 undergraduate students<sup>1</sup> made 0–2 errors on the high-imageability items and 0–5 errors on the low-imageability items. With spoken stimuli, he showed a trend towards worse performance on the low- than the high-imageability items, though this failed to reach significance (Fisher Exact Test,  $z = 1.57$ ), possibly because of the small number of items. Of the eight errors made on the auditorily presented low-imageability set, four were semantic foils (e.g. selecting PASSION rather than VALOUR as an associate for “courage”) and four were unrelated foils (e.g. selecting BELIEF as an associate for “career”). This test corroborated the previous finding; DrO's written word comprehension was unimpaired but he was poor at comprehending spoken words of low imageability.

For synonym judgements, with written presentation DrO's results were within the normal range for a group of 20 student controls (155–160). The only error DrO made was on the pair LISTEN–ATTEND, which he judged as having different meanings. Performance on the auditory version (130/160) was well outside the student range (152–160) and was significantly worse than his written score (McNemar Test,  $P < .02$ ). Significantly more errors occurred for low- than high-imageability words (Fisher Exact Test,  $z = 2.31$ ,  $P < .05$ ). However, there were equal numbers of errors for high- and low-frequency word pairs.

These tests of DrO's word comprehension conformed to a clear pattern: There was no evidence of any impairment in written word comprehension, but auditory comprehension was impaired, particularly for abstract or low-imageability words. Such a pattern of results suggests that DrO is “word deaf”; that is, he has a word comprehension deficit specific to the auditory modality. At what level of auditory comprehension processing is this impairment located?

---

<sup>1</sup>Given DrO is relatively young and is presumed to have a high premorbid verbal ability, we have chosen where possible to use our young normal controls, who were all undergraduate psychology students, to provide control data, rather than our elderly controls, who have a wider spread of educational experience and are, on average, older than DrO.

In order to distinguish between a word meaning deafness (a deficit of mapping between lexical and semantic forms) and an earlier auditory processing deficit, DrO's ability to process auditory information in other tasks was investigated.

## AUDITORY TESTS

Four tests were devised to assess aspects of early auditory processing important for speech recognition. In addition, a number of phoneme discrimination tests were administered to establish whether DrO was able to detect minimal feature differences between phonemes.

### Early Auditory Tests

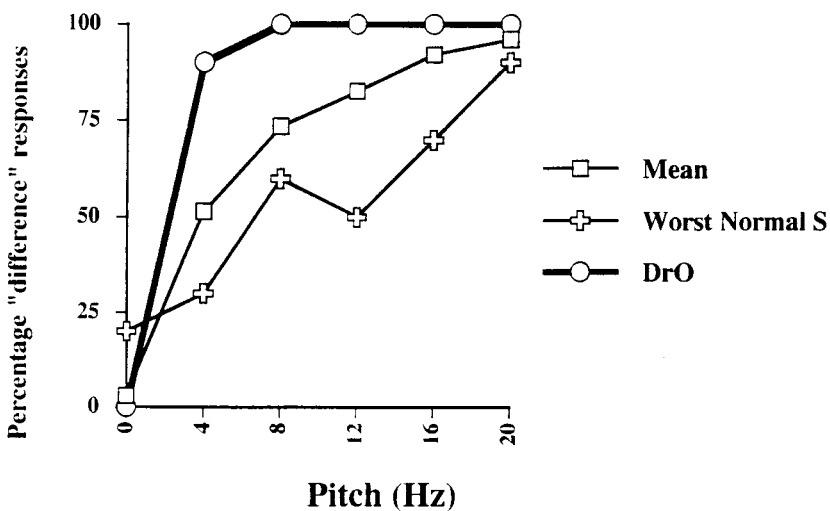
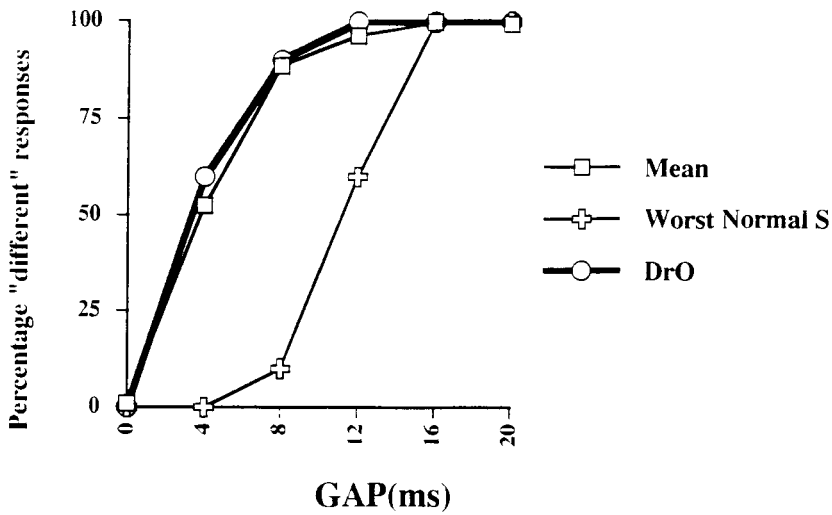
The "early auditory" tests involved measures of temporal resolution and of discrimination of formant and fundamental frequency. All the tests used a similar procedure; on each trial the listener was presented with a pair of stimuli and asked to decide whether the members of the pair were the same or different. Temporal resolution was assessed using a pair of noise bursts as stimuli; the noise bursts were either physically identical, or one burst incorporated a brief silent interval (or "gap") with duration up to 20msec. Discriminability of steady-state formant frequency was measured using pairs of stimuli each consisting of a single formant; the formant frequencies were either the same, or differed by up to 200Hz. The assessment of sensitivity to changes in formant frequency involved similar stimuli, but when the stimuli were different the formant frequency of one member of the pair was modulated smoothly in frequency (to a maximum frequency difference of 150Hz). The discriminability of fundamental frequency was measured using stimuli similar to those in the formant frequency tests, but in which, when present, the difference between members of the stimulus pairs was in fundamental frequency (up to 20Hz), perceived by listeners as a difference in pitch.

### *Results*

Figure 1 contains the results for all four "early auditory" tasks. Each graph shows the mean performance for a group of 20 elderly control subjects (mean age = 70.9, SD = 5.6) and the performance of the worst normal, as well as DrO's performance. The results are expressed as the number of times the pairs of stimuli were judged different. It can be seen that, in every case, DrO performed within the normal range. (Subsequent testing of 20 undergraduate students indicates that DrO also performs within the younger range.)

### Phoneme Discrimination Tests

To test DrO's ability to detect phonemes, two minimal pair tests from the ADA Comprehension Battery (Franklin et al., 1992) were used. Both tests comprised





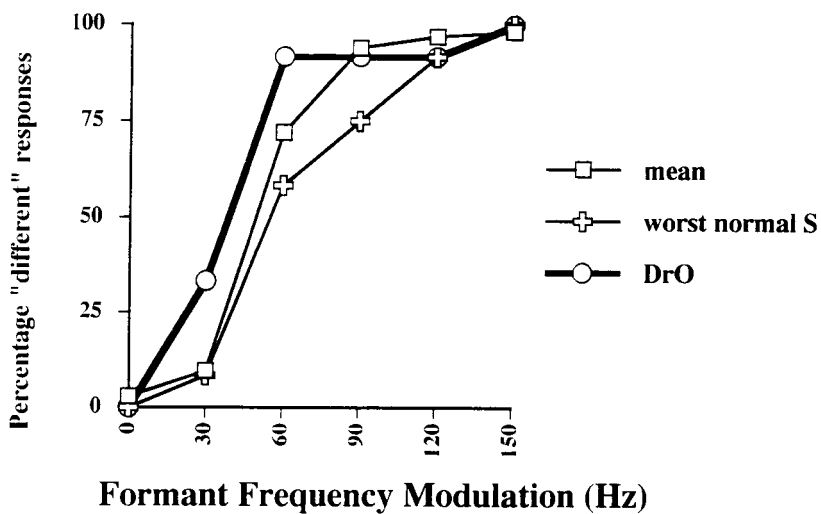
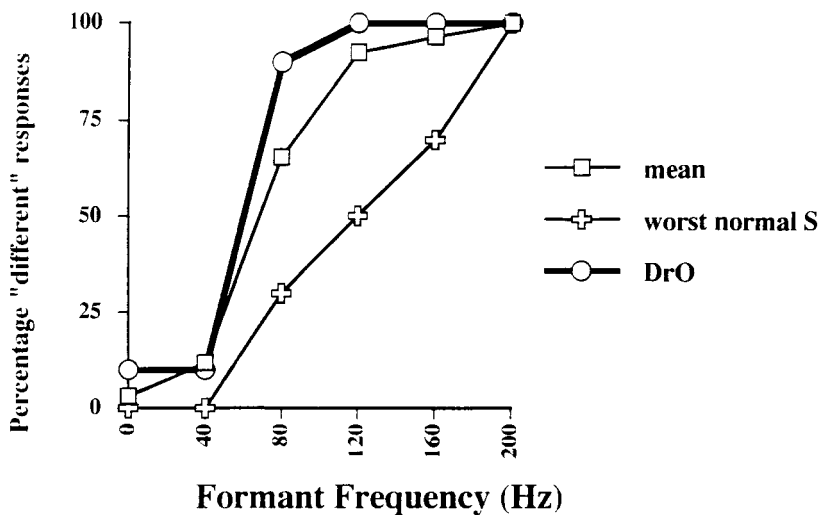


FIG. 1. Performance on the early auditory tests.

40 CVC pairs. For the different pairs, either the initial or final consonant differed by one or two distinctive features. One test comprised real words (for example “buck–bud”), the other novel strings (for example “peb–beb”).

### *Results*

DrO's performance on the CVC phoneme discrimination tests were within the normal range. This was true for both the word version (40/40 correct) and the nonword version (38/40 correct, student range 36–40) of the test.

### Picture/Word Decision Test (Howard & Franklin, 1988)

This test used 97 of the items from the 100 Picture Naming Test (excluding the 3 items that had no phonological real word neighbours). DrO was shown a picture, heard a word, and had to say whether the word was a correct name for the picture. There were four conditions, rotated around four presentations of the items spread over four weeks. The conditions were correct, semantically related words, phonologically related real words, and phonologically related nonwords (e.g. picture of a saddle: correct ↑ “saddle”; semantic related word ↑ “bridle”; phonologically related word ↑ “paddle”; phonologically related nonword ↓ /kæd@/).

If DrO's impairment was at a lexical or a prelexical level, then he should show a tendency to judge the phonologically related items to be appropriate names for the pictures.

### *Results*

The results of the Picture/Word Decision Test are shown in Table 2. The majority of errors were false positive errors, where the semantically related word was accepted as an appropriate name for the picture. Over a large number of items, DrO made very few false positive errors to phonologically related candidate names, irrespective of whether they were real words or nonwords.

TABLE 2  
DrO's Performance on the Picture/Word  
Decision Test

|                               |     |
|-------------------------------|-----|
| Correct (/388)                | .93 |
| Errors                        |     |
| Misses                        | .01 |
| Semantic errors               | .16 |
| Phonological real word errors | .06 |
| Phonological nonword errors   | .05 |

## REPETITION AND DEFINITION

To establish whether DrO was able to access the lexical forms of words he was unable to comprehend, he was given a list of words for repeating and defining. If the auditory comprehension deficit is prelexical, then repetition should be as impaired as comprehension. DrO heard each word, repeated it, and then gave a verbal definition. The experimenter said the word only once. The list comprised 100 high-imageability words and 100 low-imageability words (Howard & Franklin, 1988).

### Results

The numbers of correct repetitions and appropriate definitions in response to the heard word can be seen in Table 3. Since at this point DrO was still anomic, the definitions were scored according to fairly lax criteria; a partial definition was considered acceptable. These words were not given in a separate test of defining written words, but in every case where DrO was unable to produce a correct definition in response to the heard word, the word was written down for him. He was *always* able to give an appropriate verbal definition when he could see a word rather than hear it.

Although DrO was unable to give an appropriate definition for 55 words he was able to repeat all but 6 of them. As with the previous comprehension tests, there is a strong effect of imageability on definition (Fisher Exact Test,  $z = 5.90$ ,  $P < .001$ ). Table 4 contains examples of definitions, the first four being examples of acceptable (including joke) definitions.

Of the total corpus of 51 errors in definition, 43 were failures to give any definition. Only 3/51 errors corresponded to "mishearings" of phonologically similar words ("rate" defined as "radio," "harm" as "heart," and "trim" as "twin") and on each of these occasions the target word was repeated correctly. The four examples given in Table 4 were typical of the majority of errors in definition. In every case where DrO repeated a word but then failed to retrieve its meaning he was able to continue repeating it. It did not seem to be the case that the meaning was unavailable because the word was decaying abnormally quickly.

Was DrO's comprehension dependent on being able to repeat the word? That is, on the occasions when he was unable to repeat a word correctly, was he also

TABLE 3  
DrO's Repetition and Definition of 200 Heard  
Words

|                   | <i>Definition</i> | <i>Repetition</i> |
|-------------------|-------------------|-------------------|
| High imageability | .92               | 1.00              |
| Low imageability  | .57               | .94               |

TABLE 4  
Examples of DrO's Repetition and Definition

| <i>Target</i> | <i>Correct Responses</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| "debate"      | "debate, a debate is usually formally situation, like Oxford for example, where they have debating both two different people trying to figure out a different point of view."                                                                                                                                                                                                                                                                                                     |
| "spare"       | "a spare is something more than you actually need"                                                                                                                                                                                                                                                                                                                                                                                                                                |
| "science"     | "science—science is the disease of our generation."                                                                                                                                                                                                                                                                                                                                                                                                                               |
| "woman"       | "woman is the grown up of the male species"                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <i>Target</i> | <i>Incorrect Responses</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| "slow"        | "slow, slow, slow, slow, I know what it is but I can't get it, slow, slow—you'll have to write it down for me." [The written word was given at this point.] "Oh slow, well slow is the opposite of fast."                                                                                                                                                                                                                                                                         |
| "severe"      | "severe, severe, no I won't get that one." [nothing at all?] "severe, well I know what it means but you'll have to write it down for me." [you've no idea about what it means?] "not really—sometimes I can tell, but this I know I won't get." [not even part of it?] "no way." [The written word was given at this point.] "oh, severe, severe, severe is someone who is very hard, like a judge is very severe on giving you a, when you give somebody putting to, to prison." |
| "soul"        | "soul, soul, soul, I don't know what it is—I should know that." [The written word was given at this point.] "Oh soul, it's a religious concept, is the soul."                                                                                                                                                                                                                                                                                                                     |
| "mature"      | "mature, mature, mature, what's a mature? I don't know I missed it. [The written word was given at this point.] "oh mature, mature—that it somebody who is very wise and usually older."                                                                                                                                                                                                                                                                                          |

unable to define it? Because DrO made so few errors in repetition, the results were analysed from a larger corpus of words given for repetition and definition. The corpus included the 200-word list already described, with an additional 465 words that varied in abstractness, syllable length, and frequency. Out of the 665 items, 486 were both repeated and defined correctly. The majority of the errors (129) were in definition alone, as in the examples given in Table 4. On 29 occasions, DrO repeated a word incorrectly, and then gave a definition appropriate to the word he repeated (e.g. "same"  $\Pi$  "sane $\ddagger$  opposite of insane," "part" $\Pi$  "park $\ddagger$  what you do with a car"). However, on 21 occasions he repeated a word incorrectly but went on to give an appropriate definition (e.g. "medicine" $\Pi$  /menis@n/ $\ddagger$  something that's supposed to make you well," "agency"  $\Pi$  "aids $\ddagger$  something to do with the office"). The fact that DrO was on occasion able to define words he could not repeat is again incompatible with the notion that the semantic comprehension deficit is actually due to a mild phonological input impairment.

The relative preservation of repetition (along with his good performance on auditory and phoneme discrimination tests) suggests that DrO's word deafness is not due to some impairment of early auditory analysis; that is, he is not word sound deaf. Could it be that his impairment is in auditory lexical access? If this

were the case, then word repetition could only be achieved by repeating directly from input to output phonology. Since DrO's errors in word repetition are rare, and consist of single-phoneme changes, this route would have to be largely intact and nonword repetition would be preserved.

## WORD AND NONWORD REPETITION

DrO was given a list of 60 words and 100 nonwords to repeat. The real words comprised 20 each of lengths 3–4, 5–6, and 7–8 phonemes, matched for frequency and imageability. Sixty of the nonwords were made by changing a single phoneme in a real word. The remaining 40 ("non-word-like") nonwords were pronounceable strings of 5–6 and 7–8 phonemes, each of which differed from real words by at least 3 phonemes. The stimuli were presented on DAT tape over headphones.

### Results

The results can be seen in Table 5. DrO was able to repeat 48/60 real words correctly. Three of the errors to real words were spontaneously self-corrected and 8/9 errors differed by 1 phoneme from the target. The remaining error differed from the target word by two phonemes. There was a significant effect of phoneme length for the real word repetition (Jonckheere Trend Test,  $z = 2.16$ ,  $P < .05$ ) such that DrO was more likely to repeat shorter words correctly. In order to establish whether this effect of length was a consequence of the real words being presented mixed up with nonwords, the same words were presented on a later occasion, still on DAT tape, but without the nonwords. DrO

TABLE 5  
DrO's Word and Nonword Repetition

|                        | <i>Phonemes</i> |     |     | <i>Total</i> | <i>Control Range<sup>a</sup></i> |
|------------------------|-----------------|-----|-----|--------------|----------------------------------|
|                        | 3–4             | 4–6 | 4–8 |              |                                  |
| Words                  | 19              | 16  | 13  | 48/60        | 54–60                            |
| Word-like nonwords     | 4               | 1   | 2   | 7/60         | 40–53                            |
| Non-word-like nonwords | —               | 0   | 0   | 0/40         | 25–35                            |

| <i>Examples of Errors</i> |   |            |                               |   |          |
|---------------------------|---|------------|-------------------------------|---|----------|
| <i>Word-like Nonwords</i> |   |            | <i>Non-word-like Nonwords</i> |   |          |
| ætmoʃfəʊ                  | → | æsmiə      | vriðɔʃ                        | → | vəraz    |
| dræn                      | → | græd       | bimdelen                      | → | bɒlibɒbi |
| rains                     | → | nali       | briɔl                         | → | pɒliəpɔi |
| dæfədɪt                   | → | delɪdaɪ    | lɔpkæli                       | → | locally  |
| prəʊdesə                  | → | predestine | zɔfrəp                        | → | slowly   |

<sup>a</sup>Range taken from performance by 20 elderly controls (mean age = 70.9, SD = 5.6)

repeated 20/20 of the 3–4-phoneme words, 17/20 of the 5–6-phoneme words, and 14/20 of the 7–8-phoneme words. There was again a significant effect of phoneme length for real word repetition (Jonckheere Trend Test,  $z = 2.41$ ,  $P < .05$ ).

Performance on the word-like nonwords was markedly worse. DrO was only able to repeat 7/60 word-like nonwords correctly, and these tended to be the shortest strings. He was unable to repeat any of the “non-word-like” nonwords. Examples of his errors in repeating nonwords are shown in Table 5. It can be seen that the phonology of many of DrO’s nonword repetition errors differed markedly from their targets.

Why does DrO make any errors in word repetition, given that we have argued he has no lexical input impairment? Although not the focus of this paper, it is possible that his errors in repetition reflect an impairment at the level of speech production rather than input, since his performance on oral reading and repetition of both real words and nonwords are equivalent. Using a list of 80 words varying orthogonally in imageability and frequency (ADA Battery, Franklin et al., 1992), DrO repeated and read 67/80 words correctly (presented on different occasions). All errors were closely phonologically related to the target words. As reported earlier, when given nonwords for repetition DrO is able to repeat only the shortest phoneme strings and the errors are either phonologically related to the target or neologisms. When given a subset of the word-like nonwords for oral reading, he correctly read 9/20 of the 3–4-phoneme strings and 1/20 of the 5–6-phoneme strings; the errors were again a mixture of phonologically related strings and neologisms. It appears likely that such quantitatively and qualitatively similar performances denote a common underlying impairment, which must be at the level of speech output.

DrO does not appear to be word sound deaf; his few errors in word repetition are attributable to an impairment of output phonology. Is his word deafness pre- or post-lexical? If it is the latter, DrO’s ability to distinguish between auditorily presented words and nonwords should be preserved.

### LEXICAL DECISION

DrO was given a lexical decision test from the ADA Comprehension Battery (Franklin et al., 1992). Both the written and spoken versions of the test were administered. The 80 real word items in the test contrast high and low imageability, high and low frequency, and short and long phonological length (2–4 vs. 6–7 phonemes). The 80 nonwords were constructed by changing a single phoneme in each of the real words. The auditory version was presented on DAT tape. Items were presented in a random order and DrO was required to respond “yes” or “no” as appropriate.

## Results

DrO made no errors on the written version of the test and only 9 errors (94% correct) on the auditory version, which was within the range of the 20 student controls (147–160 correct). Out of the nine errors, seven were false positive errors on long nonwords, which also conforms to the normal pattern. There was no evidence to suggest that auditory lexical decision was impaired. Good lexical decision and the mild impairment of repetition contrasted markedly with DrO's poor performance on any test that necessitated the word's *meaning* to be accessed from the heard word. The word deafness appeared to be a highly specific impairment of the mappings between lexical form and meaning, that is, word meaning deafness. It has already been demonstrated that success in comprehension is predictable on the basis of the imageability of the word; are there any other factors that affect comprehension?

### IMAGEABILITY AND PHONEME LENGTH

All the tests used so far have contrasted a set of high-imageability words with a set of low-imageability words. Is there a critical level of imageability, with words above that level being comprehensible to DrO, or is it simply the case that, for DrO, as words become more imageable there is a higher probability of them being correct? DrO was asked to repeat and define words that were spoken to him; they had four different levels of imageability, matched on frequency (Franklin et al., 1994). There were 40 words in each of the 4 levels of imageability (2.5  $\Pi$  3.5, 3.5  $\Pi$  4.5, 4.5  $\Pi$  5.5, and 5.5  $\Pi$  6.5; Paivio, Yuille, & Madigan, 1968). For every set of 40 words, half were low-frequency words (0–5 occurrences per million) and half were high-frequency (16–150 occurrences per million). The words were matched for phoneme length across the high- and low-frequency sets, but there was a slight tendency for the more imageable words to be shorter (the means across the different sets varied between 4.2 and 5 phonemes per word)<sup>2</sup>.

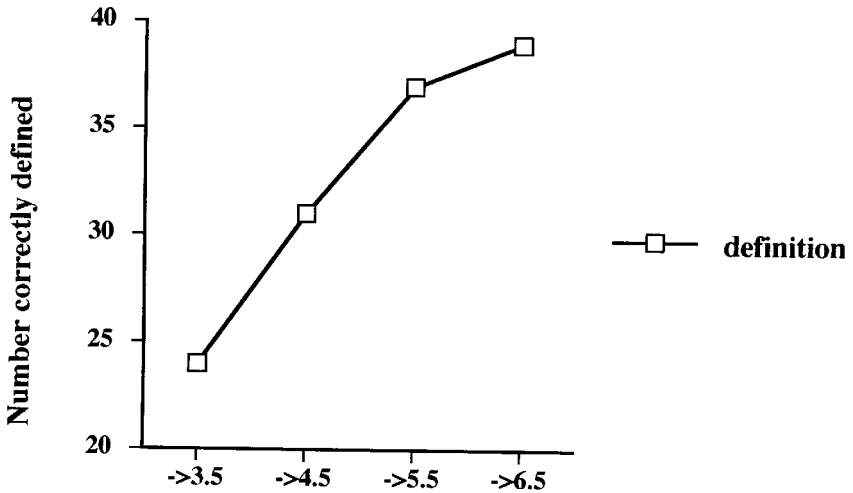
To establish whether DrO's comprehension was affected by the length of the word, an unpublished list devised by Howard was administered. The list consisted of 30 high-imageability and 30 low-imageability words for each of 3 lengths (1, 2, and 3 syllables), matched for frequency. The same procedure was used; DrO heard the word, had to repeat it and then give a definition.

## Results

The word definitions were scored as for the previous list, using fairly lax criteria. The results of defining the four levels of imageability list can be seen in Fig. 2. It is clear that there was a gradual improvement in performance as

---

<sup>2</sup>Since DrO finds longer words easier to understand, this bias in the word sets merely serves to underestimate the effect of imageability across the levels.



## Imageability

FIG. 2. Number of correct definitions given by DrO to words of different levels of imageability.

imageability increased (Jonckheere Trend Test,  $z = 4.57$ ,  $P < .001$ ). Word frequency did not affect performance; DrO defined 63/80 high-frequency words and 67/80 low-frequency words.

The results of the syllable length  $\times$  imageability list are shown in Table 6. As expected, there is a significant effect of imageability in definition (Fisher Exact Test,  $z = 4.92$ ,  $P < .0001$ ). Although DrO makes more errors in repeating low- than high-imageability words, this difference fails to reach significance ( $z = 1.21$ ,  $P > .05$ ). For definition, as the words got longer, DrO was more likely to be able to define them. This effect was significant both for the high-imageability words (Jonckheere Trend Test,  $z = 2.52$ ,  $P < .01$ ) and for low-imageability words (Jonckheere Trend Test,  $z = 2.53$ ,  $P < .01$ ). Repetition, however, tended to get *worse* as the words got longer, although this trend failed to reach significance (Jonckheere Trend Test,  $z = 1.61$ , ns).

To establish that word length and imageability were the two critical variables that predicted performance, a final definition test was given. This consisted of 132 words taken from the Gilhooly and Logie (1980) corpus, which had measures for imageability and phoneme length. Word familiarity ratings and word frequency measures were taken from the MRC Database (Quinlan, 1992). The reverse length effect shown by DrO could be explained in terms of longer words being more phonologically distinct, in that they have fewer phonological neighbours. In order to test this, the number of phonological neighbours was



TABLE 6  
DrO's Repetition and Definition of Words of 1,  
2, and 3 Syllables (Syll.)

|                   | 1 Syll. | 2 Syll. | 3 Syll. |
|-------------------|---------|---------|---------|
| <i>Definition</i> |         |         |         |
| High imageability | .87     | .97     | 1.00    |
| Low imageability  | .47     | .67     | .77     |
| <i>Repetition</i> |         |         |         |
| High imageability | .90     | .93     | .80     |
| Low imageability  | .87     | .80     | .73     |

TABLE 7  
Mean Values and Ranges for Five Word Variables  
for Corpora for Repetition and Definition

| Variable     | Mean  | Minimum | Maximum |
|--------------|-------|---------|---------|
| Familiarity  | 4.78  | 2.56    | 6.57    |
| Frequency    | 37.00 | 1.00    | 660.00  |
| Imageability | 4.96  | 2.65    | 6.39    |
| Phon. length | 4.96  | 2.00    | 10.00   |
| Neighbours   | 8.88  | 0.00    | 51.00   |

calculated for all the words (D. Howard, personal communication). This "N" value was the sum of the number of words obtained by substituting, adding, or deleting a single phoneme. The mean values and ranges for each variable for the corpora selected are shown in Table 7.

As before, DrO heard the word, repeated it, and had to give an appropriate definition. Since this test was administered later in DrO's recovery, his anomia had to some extent resolved and it was deemed appropriate to use a more rigorous scoring system for his definitions. These definitions were then marked independently by 10 raters. A correct score was given if a majority of the raters scored the definition as correct; a minority led to an incorrect score.

The definition scores were used as the dependent variable in a number of simultaneous logistic regression analyses. Various lexical variables were used in the regression equation in order to predict the dichotomous definition scores. Word familiarity was used in the regression rather than word frequency, but neither even approached significance in the simple correlational analyses (correlation log frequency and definition success = .065,  $P = .460$ ; correlation familiarity and definition success = .01,  $P = .922$ ). When entering word familiarity, imageability, and phonological neighbours into the equation, the definition scores were best predicted by imageability (Wald = 9.25,  $P = .002$ ) and phonological neighbours (Wald = 3.10,  $P = .08$ ), whereas familiarity was a poor predictor (Wald = 0.93,  $P = .33$ ). When the neighbourhood count was replaced with the number of phonemes, a very similar pattern of results was

found (imageability: Wald = 9.28,  $P = .002$ ; phonemes: Wald = 3.05,  $P = .08$ ; familiarity: Wald = 0.86,  $P = .36$ ). However, due to a very high intercorrelation ( $r = -.77$ ), the logistic regression technique is unable to resolve any differences between the predictive power of length and phonological neighbours: When phonological neighbours and phoneme length were entered together along with imageability and familiarity, only imageability was found to be a significant predictor of DrO's definition score.

Although the reverse length effect only approached significance in both analyses ( $P = .08$ ), the results do replicate those found in the orthogonal manipulation of length and imageability. DrO was better at accessing words that were more imageable and words that were phonologically distinct (whether this is measured in terms of length or number of neighbours), even when the word's familiarity was taken into account.

By comparison, when DrO's repetition scores were used as the dependent variable in simultaneous regressions, only imageability reached significance (Wald = 4.94,  $P = .03$ ), whereas familiarity and phoneme length were relatively poor predictors (familiarity: Wald = 1.32,  $P = .25$ ; phonemes: Wald = 2.54,  $P = .11$ ). It was interesting to note that although the length effect in his repetition did not reach significance, the trend was in the reverse direction to that found in word definition. The simple correlation between definition and phoneme length was  $r = +.16$ , and between repetition and phoneme length was  $r = -.13$ . DrO was better at defining long words but better at repeating short words.

## DISCUSSION

The dissociation between written and spoken comprehension indicated that word meanings were available to DrO, but he was often unable to *access* them using auditory information. Although it was clear that there was a significant degree of impairment in DrO's auditory comprehension, there was no indication of *any* impairment in the written modality; in Shallice's (1988) terminology a "classical" dissociation. We can be confident in categorising DrO's comprehension impairment as a form of "word deafness." But at what level of processing in the auditory modality should the deficit be located?

The traditional characterisation of "word meaning deafness" is of poor comprehension and intact repetition. Was this true for DrO? There was a clear dissociation between DrO's auditory comprehension and his ability to repeat. This was not a "classical" dissociation, however; although repetition was markedly better than comprehension, DrO's word repetition was impaired relative to normal performance and, like his comprehension, was somewhat affected by word imageability. However, there are several reasons why his repetition and comprehension impairments are unlikely to have the same underlying cause. DrO is often able to repeat a word correctly that he is unable

to comprehend. Equally there are other, although fewer, occasions where he repeats a word incorrectly but is still able to give an acceptable definition. Whereas DrO is better at comprehending words if they are long, he is better at repeating short words. Finally, DrO was often able to continue repeating words he was unable to understand, suggesting that the impairment is not one of rapidly decaying phonological information (Martin, Dell, Saffran, & Schwartz, 1994).

An alternative hypothesis for the qualitative dissociation between repetition and comprehension is that the two systems might reveal different characteristics following a common underlying impairment. Specifically, a single auditory analysis impairment could result in repetition being more successful for short words and comprehension for long words. It would in principle be possible to account for the dissociation found in accuracy between repetition and comprehension in similar terms. However, for such alternative explanations to stand there must be independent evidence of an early auditory analysis impairment. DrO's ability to perform well on a number of tests requiring auditory processing (early auditory tests, phoneme discrimination, picture/word decision) are inconsistent with a deficit at this level. Neither is there any clear account of how such a deficit would result in activation that was sufficient to produce a highly detailed phonological output, but insufficient to access semantic information.

DrO's comprehension deficit is not therefore accounted for in terms of a prelexical, phonological type "word sound deafness." Could it be a deficit at a lexical level (word form deafness, or pre-access word deafness)? Since DrO is profoundly impaired when asked to repeat nonwords, his relatively good word repetition suggests he is able to use lexical knowledge for this task; however it is at least plausible that output lexical knowledge is being utilised for repetition. The notion that DrO's comprehension deficit is post-lexical is supported by his ability to distinguish reliably between words and nonwords for lexical decision, even for a list that contains those words most difficult for him to comprehend: short, low-imageability words. Further support comes from his ability to reject phonologically related words in the Picture/Word Decision Test.

To summarise, we argue that DrO's comprehension is not explicable in terms of a phonological decay or a lexical deficit because of the following:

1. Normal performance on the early auditory tests.
2. Performance within the normal range on word and nonword phoneme discrimination tests.
3. Performance within the normal range on auditory lexical decision.
4. An ability to reject phonological errors in comprehension.
5. A strong dissociation between word and nonword repetition.
6. A strong dissociation between repetition and comprehension.
7. A significant advantage for longer words in comprehension + a trend towards better repetition of shorter words.

8. An ability to continue to repeat a word correctly without comprehending it.

The deficit appears to be one of mappings between lexical and semantic information in the auditory modality; a word meaning deafness. As pointed out by Allport and Funnell (1981), this ability to comprehend a word when written but not spoken is problematic for models where written words are understood using a single comprehension system via phonological recoding. DrO appears to be a particularly convincing case of word meaning deafness; what are the specific characteristics of such a deficit?

There is a strong and reliable effect of imageability (or abstractness) in DrO's auditory comprehension. Where it has been tested, this has been demonstrated in all patients who have a lexical or semantic deficit specific to auditory processing, and whose performance is above chance (Franklin, 1989; Howard & Franklin, 1988; Katz & Goodglass, 1990; Martin & Saffran, 1992).

An inability to comprehend words of low imageability is a common finding in acquired dyslexia as well as in aphasia (Coltheart, Patterson, & Marshall, 1980). Plaut and Shallice (1993) have recently implemented a connectionist model that contained features representing meanings for both abstract and concrete words. It assumed that abstract words are represented by a smaller number of features. The model was "lesioned" at different input levels as well as within the semantic level itself. Irrespective of type or site of "lesion," more errors were produced for abstract than for concrete words (with the exception of lesioning the semantic level clean-up units, which produced a slight advantage for abstract words). This model of reading accords with the pattern of results, in terms of imageability, described here for auditory comprehension. Particularly interesting in this regard is the finding that the effect of imageability in DrO's comprehension appears to be an effect of "graceful degradation," which is more in accord with the Plaut and Shallice account of how abstract words are represented than with an account that treats them as a special word category (see Franklin, Howard, & Patterson, 1995, for a fuller discussion). As the words increase in abstractness, their corresponding patterns of activation are represented across a small number of features. So patterns for abstract words are confined to a semantic space with fewer dimensions than are patterns for concrete words. This is equivalent to there being less space in which to locate abstract patterns, giving rise to a greater degree of potential confusability.

It is interesting to note that DrO's ability to comprehend a word does not appear to depend on the frequency of the word's occurrence. Word frequency was manipulated in both the synonym judgements test and the test of defining words of different levels of imageability. In neither case were the higher-frequency words better understood. In the logistic regression analyses, neither frequency nor rated word familiarity were found to be significant predictors of correct definition (the simple correlation between log frequency and ability to

define was in the order of .065). This is consistent with other patients' lexical and post-lexical auditory comprehension deficits reported in the literature; word frequency does not appear to be an important determinant of accuracy. This inability to demonstrate a word frequency effect is problematic for any connectionist model that employs a learning algorithm (such as the delta rule) in which connection strength depends on frequency of presentation.

DrO's ability to comprehend a word is determined by its phoneme length. It is not clear whether this is actually an effect of phoneme length or number of phonological neighbours. Since this is a greater impairment for shorter words, it may be termed a "reverse" length effect (Howard & Franklin, 1988). Clearly length and number of phonological neighbours are highly correlated. Best (1995) describes a patient, CGJ, who was better at naming long words than short words, and this was consistent with a neighbourhood effect. Best was able to disambiguate the two effects. When syllable length was held constant, CGJ actually named words better if they had more neighbours. Unfortunately, because DrO only makes significant errors in comprehending very short words, it was not possible to carry out a similar orthogonal comparison.

Is this reverse length effect as ubiquitous as the imageability effect? This effect has not been reported as often as the imageability effect, but this is perhaps unsurprising since it has rarely been tested. MK (Howard & Franklin, 1988) was found to show a reverse length effect in both comprehension and repetition. Since MK's pattern of impairment was consistent with poor ability to access auditory lexical information, it seemed unsurprising that long words, which are more resistant to degraded stimuli because they have fewer phonological neighbours, should be easier for MK to process (Norris, 1994). We have argued that DrO has the same reverse length effect in comprehension, but that the deficit causing this effect is at a post-lexical stage of the model. How does this accord with different models of word comprehension?

There are three different broad classes of model of lexical processing currently in use; logogen type models (Morton & Patterson, 1980), interactive activation (IAC) models (McClelland & Elman, 1987), and connectionist models with distributed representations (Plaut & Shallice, 1993). The logogen model has highly discrete levels of processing, where only one representation will reach threshold at any one time. Such a model would not be able to account for a variable, which is presumed to have its effect at an earlier stage of processing (i.e. either phoneme length prelexically or number of neighbours at a lexical level), being affected by damage at a later stage. Although length could affect the time-course or accuracy of the logogen reaching threshold, once threshold is achieved semantic information is only activated from that particular "logogen" or lexical item. DrO appears to be able to access lexical information irrespective of the item's length, and the output to semantic information should not, therefore, be affected by this variable.

It is possible to predict that IAC models, on the other hand, could account for the data quite well, irrespective of whether the effect is one of syllable/phoneme length or number of phonological neighbours. The more distinct a lexical representation, the more quickly and the more strongly it will activate a semantic representation. If the activation process itself is damaged, then the more distinct words‡ those with fewer neighbours‡ are more likely to be activated sufficiently. However, it should be noted that it is possible to obtain the opposite effect; for the TRACE model there is a general advantage for short over long words, at least in the early stages of processing the word (McClelland & Elman, 1987).

It is necessary to be more speculative in the case of what might happen to words of different phoneme/syllable length, since most existing word recognition IAC models only code words with very short phoneme strings. If one assumed that each phoneme in a word produced some activation to the lexicon, then longer words would produce more activation. Although arising from a different locus, this would ultimately have the same effect on the impaired semantic access as would the number of neighbours. Distributed connectionist models are equally likely to provide an account of the reverse length effect. In their model of reading, Plaut and Shallice (1993) only use words of 3 or 4 letters in length, and only 40 words in total. However, they do find that a visual error is significantly more likely when the stimulus word shares a high degree of visual similarity with the given set of words. If one extrapolates the similarity effect to errors of comprehension, one might predict that visually similar words would be harder to differentiate within the semantic space when the input to that system has been damaged. This will be a natural consequence of the model itself, where visually similar words are represented by similar vectors that point initially to closely aligned points in the semantic space, prior to their adjustment by the clean-up units. Vectors representing visually distinct strings will point to relatively spacious areas of the semantic space and therefore will be less vulnerable to noise applied to the pointing vector. Applied to auditory rather than orthographic processing, this could result in the reverse phoneme length effect. As for the concrete/abstract dimension coded in the model, this also seems to be explicable in terms of "potential confusability." The reverse length effect and the imageability effect may be a consequence of the same type of distinctiveness effect but realised in different codes. Furthermore, if this characterisation is correct, the effects of reverse length and imageability should be apparent in all "pure" cases of lexical or semantic access impairment specific to the auditory modality.

It would in principle be possible to model the effects of imageability and reverse phoneme length using either an IAC or a distributed system; in neither case would feedback connections be required. As with the modelling of word comprehension, tests used for investigating aphasic comprehension rarely manipulate phoneme length as a variable of interest. It is at least plausible that

phoneme length will have a widespread effect in impaired lexical recognition. Will other patients, given the appropriate testing, show reverse phoneme length effects in auditory comprehension? Might some acquired dyslexic subjects demonstrate corresponding reverse letter length effects in written comprehension?

Manuscript received 28 November 1995

Revised manuscript received 20 August 1996

Manuscript accepted 26 August 1996

## REFERENCES

- Allport, D.A., & Funnell, E. (1981). Components of the mental lexicon. *Transactions of the Royal Society of London*, B295, 397–410.
- Best, W. (1995). A reverse length effect in aphasic naming: When elephant is easier than ant. *Cortex*, 31, 637–652.
- Bramwell, B. (1897). Illustrative cases of aphasia. *Lancet*, 1, 1256–1259. Reprinted (1984) as: A case of word meaning deafness. *Cognitive Neuropsychology*, 1, 249–258.
- Coltheart, M., Patterson, K.E., & Marshall, J.C. (1990). *Deep dyslexia*. London: Routledge & Kegan Paul.
- Franklin, S. (1989). Dissociations in auditory word comprehension; evidence from nine fluent aphasic patients. *Aphasiology*, 3, 189–207.
- Franklin, S., Howard, D., & Patterson, K. (1994). Abstract word meaning deafness. *Cognitive Neuropsychology*, 11, 1–34.
- Franklin, S., Howard, D., & Patterson, K. (1995). Abstract word anomia. *Cognitive Neuropsychology*, 12, 549–566.
- Franklin, S., Turner, J., & Ellis, A.W. (1992). *The ADA Comprehension Battery*. London: Action for Dysphasic Adults.
- Gilhooly, K.J., & Logie, R.H. (1980). Methods & designs: Age of acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behaviour Research Methods and Instrumentation*, 12, 395–427.
- Howard, D., & Franklin, S. (1988). *Missing the meaning?* Cambridge, MA: MIT Press.
- Katz, R.B., & Goodglass, H. (1990). Deep dysphasia: An analysis of a rare form of repetition disorder. *Brain and Language*, 39, 153–185.
- Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Kohn, S., & Friedman, R. (1986). Word-meaning deafness: A phonological–semantic dissociation. *Cognitive Neuropsychology*, 3, 291–308.
- 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, N., & Saffran, E.M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language*, 43, 240–274.
- McClelland, J.L., & Elman, J.L. (1987). Interactive processes in speech perception: The TRACE model. In J.L. McClelland, D.E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*. Cambridge, MA: MIT Press.
- Morton, J., & Patterson, K.E. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K.E. Patterson, & J.C. Marshall (Eds.), *Deep dyslexia*. London: Routledge & Kegan Paul.

- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189–234.
- Paivio, A., Yuille, J.D., & Madigan, S.A. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology: Monograph Supplement*, 76 (1, part 2), 1–25.
- Patterson, K.E., & Shewell, C. (1987). Speak and spell: Dissociations and word-class effects. In M. Coltheart, R. Job, & G. Sartori (Eds.), *The cognitive neuropsychology of language*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Plaut, D.C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, 10, 377–501.
- Quinlan, P.T. (1992). *The MRC Psycholinguistic Database*. Oxford: Oxford University Press.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Warrington, E.K. (1981). Concrete word dyslexia. *British Journal of Psychology*, 72, 175–196.
- Ziehl, F. (1896). Ueber einen Fall von Worttaubheit und das Lichtheim'sche Krankheitsbild der subcorticalen sensorischen Aphasie. *Deutsche Zeitschrift für Nervenheilkunde*, 8, 259–307.