Semantic Loss without Surface Dyslexia

Matthew A. Lambon Ralph, Andrew W. Ellis and Sue Franklin
University of York, Heslington, York, UK

Abstract

Recent models of reading, including some computational models, suggest that error-free performance in reading aloud may be reliant, at least in part, on support from the semantic system. Such models predict that, following semantic impairment, a pattern of acquired dyslexia known as 'surface dyslexia' will be an inevitable consequence. We present data from a patient with dementia of Alzheimer type who shows a severe semantic impairment but near error-free performance on reading aloud real words (including low-frequency, irregular, abstract words) and non-words. The data are discussed in light of various models of reading and in relation to previously reported cases.

Introduction

In a series of recent papers, Patterson, Hodges and their colleagues have reported a number of patients who show a combination of semantic impairment and a particular form of reading disorder known as 'surface dyslexia' (e.g. Patterson and Hodges, 1992; Graham et al., 1994; Patterson et al., 1994). All these patients have some form of progressive dementia, either a generalized atrophy (e.g. DAT) or a more circumscribed atrophy [e.g. progressive fluent aphasia or semantic dementia (Snowden et al., 1989)]. Within these patient groups, there are a number of converging indicators of semantic impairment. For example, the patients show poor word comprehension, both spoken and written. In tasks such as matching words to pictures, they make semantic errors; for example, pointing to a picture of a knife in response to the word 'fork'. They also make semantic errors in object naming. They show a reduced ability to generate exemplars from categories such as types of animal, and have an impoverished fund of general knowledge. In comparison, other components of language, for example syntax and phonology, may be relatively spared. Unlike DAT patients, the semantic dementia patients also show preserved perceptual skills, non-verbal problem-solving abilities and episodic memory.

Surface dyslexia is a pattern of reading disorder in which the patient reads words with regular, consistent spellings (e.g. mint) better than words with irregular, exception spellings (e.g. pint). Their errors tend to take the form of misreading irregular words as if they were regular; for example, misreading colonel as 'kollonel' or trough as 'trow' (Patterson et al., 1985). Such errors suggest a

Correspondence to: M. A. Lambon Ralph, Department of Psychology, University of York, Heslington, York YO1 5DD, UK.
There may be potential problems in the way that DRN’s definitions were scored. Cipolotti and Warrington described their scoring criterion as ‘lenient … taking into account the patient’s word retrieval problems’. They scored definitions in which only superordinate information was provided (e.g. dog→‘an animal’) as incorrect. Patterson and Hodges are not explicit about how much semantic support is required to enable a low-frequency exception word to be read correctly. It may be that superordinate information would suffice. If so, then some of the words classed as being misdefined by DRN may have activated enough semantic information for correct reading aloud.

Thirdly, DRN’s definitions were better for abstract than concrete nouns. This is an unusual pattern: semantic impairment is normally reflected in better performance on concrete than abstract words (e.g. Plaut and Shallice, 1993; Franklin et al., 1993). It is not clear from the theory of Patterson and Hodges how the imageability of words might affect accuracy in impaired reading. With regard to normal readers, Strain and colleagues (1995) argue that semantic support for the reading of low-frequency exception words will be stronger for words with concrete meanings than for words with abstract meanings. Hence the words most prone to error in surface dyslexics with semantic impairment should be low-frequency, exceptional, abstract words such as suave, caste and guise. This pattern of impaired reading has been reported before (e.g. Howard and Franklin, 1988; Franklin et al., 1995). Furthermore, Strain and colleagues (1995) have shown that normal readers have particular problems in reading aloud words of this type when under time pressure. Thus, on the one hand, the model may claim that semantic representations are important in reading all low-frequency, exceptional words. Since abstract semantic information appears to be generally prone to damage then one might expect that patients with impaired semantics will find low-frequency, exceptional, abstract words the most difficult to read aloud. If this position is correct then a patient with intact reading but relatively poor semantics for abstract words would be problematic for the model. On the other hand, the model of normal reading may be conceived in terms of semantic support being greater for concrete words. If this is the case then semantic damage may critically affect the reading of concrete rather than abstract words. Consequently, the pattern of semantic impairment shown by DRN would be a problem for this latter formulation of the Patterson and Hodges model. Without a clear description of the Patterson and Hodges model in this respect (or the inclusion of imageability within the simulated semantic route of the Plaut et al. connectionist model), it is hard to adjudicate between these two positions. However, a patient with semantic damage greater for abstract items but intact reading aloud, taken with DRN’s opposite pattern of semantic breakdown, must cause problems for whichever interpretation of the model is correct.

We suggest that the Patterson–Hodges theory would be falsified by a patient who shows a severe semantic...
impairment favouring concrete over abstract items but who can nevertheless read low-frequency, exceptional, abstract words normally. We present such a patient here.

**Patient history**

DC was born in 1910. She attended school until the age of 14 years. She went into domestic service until 1947 when she married. During her married life she maintained the house and family. In 1993, DC presented with memory problems. At home, she had been frequently leaving the gas on, burning pans, losing her keys, etc. Direct questioning revealed memory problems — she was only able to give her name and the town in which she lived; and was unable to give her age, the time, day of week, month, season, year, the names of family members, etc. Medical examination revealed mildly high blood pressure, borderline hypothyroidism and dementia-type symptoms. A diagnosis of dementia of Alzheimer’s type was made.

When the present tests were conducted in 1995 her spontaneous speech was well structured and fluent, although relatively simple in content with occasional word-finding difficulties. In line with her original medical assessment, her autobiographical memory was extremely poor (she only recalled her own name and the town in which she lived), as was her episodic memory in general.

**Semantic tests**

Semantic impairment was evident in comprehension and naming.

**(a) Word–picture matching**

DC was given the word–picture matching task from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA) battery (Kay *et al.*, 1992). This required her to point to one of five object pictures in response either to a written word or to a word spoken by the experimenter. The five pictures depict the target word (e.g. button), a close semantic distractor (zip), a more distant semantic distractor (bow), a visually similar object (a coin) and an unrelated object (a banknote). DC scored 26/40 correct in the written version, making seven semantic (six close semantic, one distant) and seven ‘don’t know’ errors, and 25/40 in the spoken version where she made nine semantic (seven close semantic, two distant) and six ‘don’t know’ errors. With reference to the no response errors, it should be noted that DC could not be encouraged to pick a specific picture from the test array but rather was adamant that she ‘didn’t know’ what the word meant.

**(b) Object naming**

DC was presented with three different sets of object pictures to name on different occasions. She named 29/40 of the items from the PALPA naming test correctly, 24/40 of the target items from the word–picture matching test, and 75/140 of a third set of object pictures. Overall, 46 of her 92 naming errors (50%) were semantic in nature (e.g. axe→hammer; bear→ ‘dog or cat or cow’). The remaining errors were 16 (17%) circumlocutions, 28 (30%) ‘don’t knows’ and three (3%) visual.

**(c) Category naming**

Patients with semantic impairment have difficulty generating exemplars from semantic categories. DC was quite unable to do this task, but her comments indicated that she understood what was required. Thus, when given the category ‘breeds of dog’ she said, ‘We have always had a dog, but I can’t remember any of the different types’.

**Reading aloud versus defining**

On separate occasions DC was asked to read aloud and to define 40 object names from the PALPA picture naming test. Half of these have regular spellings and half have irregular spellings. She also read and defined on separate occasions 120 words from Shallice *et al.* (1983) which have regular, mildly irregular or very irregular spellings.

DC was also given the low-frequency items from the ‘Surface list’ of words from Patterson and Hodges (1992) which contains 42 regular and 42 matched exception (irregular) words. For these she was asked to read each word first, then to give its definition immediately afterwards.

In order to maximize the possibility of retrieving any semantic information when giving definitions, DC was always encouraged to give all the information she could on each item by repetition of the question ‘What does it mean?’; together with probe questions such as ‘What kind of animal is it?’; ‘What do you use it for?’; ‘Which part of your body is it?’ and so on.

Each definition was scored on two different criteria. The first was a lax criterion similar to that used by Cipolotti and Warrington (1995). The definition was scored as correct if DC gave any relevant information about the word (particularly in view of her word-finding difficulties), but was scored incorrect if, despite the probe questions, she only gave superordinate information, put the word into a sentence without revealing any additional information about the meaning of the word, gave an inappropriate definition, or responded with ‘don’t know’.

A second, very lax criterion was used, under which each item was scored as correct if DC gave any appropriate information about the item whatsoever, including superordinate information or any additional word given in her utterance that was semantically or associatively related to the target item. An incorrect score was given if she put the word into a sentence without revealing any further information, gave an inappropriate definition, or responded...
with 'don't know'. Examples of DC’s definitions are given below with their appropriate scores.

1. Scored as correct with both lax and very lax criteria:
   (a) sandal: ‘Something to put on feet’
   (b) blunder: ‘Made a mistake or something’

2. Scored as incorrect under lax criterion but correct under very lax:
   (a) trout: ‘A fish’
   (b) worm: ‘I don’t like worms . . . [E: where do you find them?] . . . underground’
   (c) mince: ‘Eating mince pies . . .’

3. Scored as incorrect with both criteria:
   (a) hoop: ‘Like a hoop’
   (b) swear: ‘Not to swear. Something I don’t do’
   (c) heart: ‘Don’t know’

In order to establish that the scoring adhered to the criteria, the definitions were scored separately by the first author and by a second rater who was unaware of the purpose of the study. Overall, the inter-rater agreement was extremely high (96% for the lax criterion and 96% for the very lax criterion). Discrepancies were settled by agreement.

The results of DC’s reading and defining of the word sets are presented in Table 1. As can be seen, DC’s reading aloud was very accurate. She made no errors on the PALPA object names, just four errors on the lists of Shallice et al., and two on the Surface list. The Shallice et al. lists were also given to four age-matched, healthy control subjects to read aloud. They made two, three, six and six errors each, placing DC’s score firmly within the normal range. In contrast, DC’s defining was very poor at around 20–25% on the lax criterion and 50% on the very lax criterion.

**Further investigation of DC’s reading and defining**

As DC seemed to have greater problems in comprehending abstract words, a stringent test of her reading would be abstract, exceptional words of low frequency. List 1 from Strain et al. (1995) manipulates frequency, regularity and imageability across 96 items, while list 2 manipulates regularity and imageability for a set of 64 low-frequency words. DC was asked to read-then-define these lists. The results are shown in Table 2. Overall, DC read 94/96 of list 1 and 60/64 of list 2 correctly. The four control subjects made zero, one, one and two errors on list 1, and one, one, two and five errors on list 2. Thus DC is within the normal range on both lists for reading aloud.

In contrast, she was again severely impaired at defining the meanings of these words.

### Table 1. DC’s reading and defining of three word lists (number correct)

<table>
<thead>
<tr>
<th>Test</th>
<th>Word type</th>
<th>Read aloud</th>
<th>Definition: lax criterion</th>
<th>Definition: very lax criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object labels from PALPA</td>
<td>Irregular</td>
<td>20/20</td>
<td>5/20</td>
<td>12/20</td>
</tr>
<tr>
<td></td>
<td>Regular</td>
<td>20/20</td>
<td>2/20</td>
<td>11/20</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>40/40</td>
<td>7/40</td>
<td>23/40</td>
</tr>
<tr>
<td>Shallice et al. (1983)</td>
<td>Very irregular</td>
<td>36/40</td>
<td>13/40</td>
<td>21/40</td>
</tr>
<tr>
<td>Levels of regularity</td>
<td>Mildly irregular</td>
<td>40/40</td>
<td>10/40</td>
<td>18/40</td>
</tr>
<tr>
<td></td>
<td>Regular</td>
<td>39/40</td>
<td>3/40</td>
<td>14/40</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>115/120</td>
<td>26/120</td>
<td>53/120</td>
</tr>
<tr>
<td>Patterson and Hodges (1992)</td>
<td>Exceptional</td>
<td>40/42</td>
<td>13/42</td>
<td>24/42</td>
</tr>
<tr>
<td>Surface list</td>
<td>Regular</td>
<td>42/42</td>
<td>11/42</td>
<td>19/42</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>82/84</td>
<td>24/84</td>
<td>43/84</td>
</tr>
</tbody>
</table>

### Table 2. DC’s reading and defining of word sets from Strain et al. (1995) (number correct)

<table>
<thead>
<tr>
<th>Test</th>
<th>Word type</th>
<th>Read aloud</th>
<th>Definition: lax criterion</th>
<th>Definition: very lax criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High I</td>
<td>48/48</td>
<td>16/48</td>
<td>27/48</td>
</tr>
<tr>
<td>List 1</td>
<td>Low frequency</td>
<td>46/48</td>
<td>12/48</td>
<td>23/48</td>
</tr>
<tr>
<td>(Regularity × Frequency × Imageability)</td>
<td>High frequency</td>
<td>48/48</td>
<td>14/48</td>
<td>14/48</td>
</tr>
<tr>
<td></td>
<td>Exc</td>
<td>47/48</td>
<td>13/48</td>
<td>22/48</td>
</tr>
<tr>
<td></td>
<td>Reg</td>
<td>47/48</td>
<td>13/48</td>
<td>19/48</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>94/96</td>
<td>26/96</td>
<td>41/96</td>
</tr>
<tr>
<td>Strain et al. (1995)</td>
<td>Low I, Exc</td>
<td>13/16</td>
<td>3/16</td>
<td>4/16</td>
</tr>
<tr>
<td></td>
<td>Low I, Reg</td>
<td>16/16</td>
<td>3/16</td>
<td>6/16</td>
</tr>
<tr>
<td>List 2</td>
<td>High I, Exc</td>
<td>15/16</td>
<td>1/16</td>
<td>5/16</td>
</tr>
<tr>
<td>(Regularity × Imageability for low-frequency words)</td>
<td>High I, Reg</td>
<td>16/16</td>
<td>6/16</td>
<td>10/16</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>60/64</td>
<td>13/64</td>
<td>25/64</td>
</tr>
</tbody>
</table>

I. imageability; Exc, exceptional; Reg, regular.
definitions on the items from the Strain et al. (1995) list 1 that her performance was better for the high imageability items than for the low imageability items (significant for both the lax criterion: binomial test, \( P=0.02 \); and the very lax criterion: binomial test, \( P=0.0001 \)). There are no significant differences if the items are split by frequency or regularity, nor is there an interaction between imageability and frequency (lax criterion \( \chi^2 = 3.77 \), n.s.; very lax criterion, \( \chi^2 = 1.01 \), n.s.).

Other reading tests

(a) Lexical decision

Three lexical decision tests were given to DC, all taken from the PALPA battery. In each she was shown randomly interleaved words and non-words and was asked to say whether each one was ‘Real’ or ‘Made up’.

The legality lexical decision task contains 30 words and 30 non-words which are made up from illegal letter strings (e.g. Ibao). The Frequency × Imageability lexical decision task varies the frequency of occurrence and the imageability of 60 word targets. The 60 non-words in this task are all legal and word-like. Finally, a lexical decision task was included which has 30 words varying on spelling–sound regularity and 30 non-words, with half the non-words being pronounced like real words (pseudohomophones, e.g. brume), while the other half are simple non-words (e.g. durl).

DC made no errors on the legality lexical decision task. She made three errors on the Frequency × Imageability task – all false positive responses to non-words. The mean score for 26 control subjects on these non-words (PALPA norms) is 59.88 (s.d. = 0.45). She made four errors on the third lexical decision task (two false positive errors to pseudohomophones and two to ordinary non-words). Control subjects score a mean of 14.52 (s.d. = 0.75) on the pseudohomophones and 14.89 (s.d. = 0.32) on the non-words (PALPA norms). Thus, DC shows a very high level of performance in lexical decision though she may lie just outside the normal range in her tendency to accept a few non-words as real words.

(b) Non-word reading

DC was asked to read aloud 24 non-words of three to six letters in length (PALPA test of non-word reading) and the 30 non-words from the third of the above-mentioned lexical decision tasks. She read 23/24 and 29/30 correctly.

Discussion

DC’s reading aloud of words was very good, indeed at normal levels. Even on the theoretically critical set of low-frequency, abstract exception words she performed again within the range of control subjects. Overall she made just 12 errors, of which nine were regularizations (e.g. chasm→‘chazm’, suave→‘swayve’) and three visual errors (stingy→‘sting’, yore→‘york’, mow→‘meow’). Normal control subjects made similar sorts of errors including regularizations and visual errors.

DC was also good at lexical decision and non-word reading, but her definitions of words were extremely poor. Thus she could provide no substantive semantic information for half of the words which she nevertheless succeeded in reading correctly. If DC is compared with Cipolotti and Warrington’s (1995) case DRN, the two patients show similar levels of accuracy at reading aloud. DC’s overall ability at defining is, if anything, worse than DRN’s (overall figures: DC defined 24% of words scored on the lax criterion; DRN defined 49%), but DC shows the more common pattern of better performance on concrete than abstract words.

Another difference between the present study and that of Cipolotti and Warrington (1995) is that converging evidence for a semantic impairment is available for DC in the form of impaired naming and word–picture matching with semantic errors predominating in both tasks. She also understood but was unable to perform the category fluency task. We do not believe that DC’s poor defining was due to a failure to understand the task or its general cognitive demands. She was continually prompted by the tester to give as much information as she could when defining, and some of her definitions were of good quality (e.g. mattress→‘something you sleep on’; iron→‘to smooth your dresses or clothes’). DC could also cope perfectly well with the demands of the novel metalinguistic task of lexical decision.

DC’s pattern of intact single word reading with very poor semantics, like that of WLP (Schwartz et al., 1979, 1980) and DRN (Cipolotti and Warrington, 1995), effectively undermines the Patterson and Hodges (1992) theory that surface dyslexia is an inevitable consequence of semantic breakdown. It also undermines the Plaut et al. (1995) theory that semantic support is required for the reading of (at least some) exceptional words by normal readers.

These patients establish that semantic impairment can reach a severe level without a patient necessarily becoming surface dyslexic. But what kind of reading model can account for their reading performance in the context of semantic impairment? There are at least two possible candidates. First, a dual-route model of reading allows for the possibility of semantic impairment without surface dyslexia: intact reading can proceed in this model via the sub-lexical and lexical, non-semantic routes without the need to refer to central semantic representations (see Coltheart and Funnell, 1987; and Ellis and Young, 1988, for further discussion). Alternatively, a single-route model might be adopted (for example, see simulations 1–3; Plaut et al., 1995) in which correct reading aloud may proceed via this single route alone, again without the need of semantic support. However, Plaut and colleagues note that the current single-route simulations do not adequately
capture the full range of reading performance found across the range of patients with acquired surface dyslexia.

Are there alternative explanations of the patient data that allow the present Patterson and Hodges theory (and simulation 4 of Plaut et al., 1995) to escape relatively unscathed? Two possible alternative explanations have been put forward. First, Patterson et al. (1995) suggest that ‘another hypothesis which may explain this apparent discrepancy is that the nature of the semantic memory impairment in AD differs from that in lobar atrophy in some yet unidentified way that is critical to the reading process’. This explanation seems relatively unlikely for two reasons. Unlike the DAT patients described, who do show performance like DC (Schwartz et al., 1980; Cummings et al., 1986; Friedman et al., 1992; Sasanuma et al., 1992; MB – Raymer and Berndt, 1994), other DAT patients show a correlation between surface dyslexia and severity of semantic impairment (Patterson et al., 1994). Secondly, Cipolotti and Warrington’s (1995) case DRN has a progressive lobar atrophy but, as described in the Introduction, he has intact reading despite a semantic impairment. Thus, there seem to be both lobar atrophy and DAT cases described, all with semantic impairments, either with accompanying surface dyslexia or with intact reading.

The second explanation involves the notion of individual differences in reading (Patterson et al., 1995; Plaut et al., 1995). If these individual differences are viewed in terms of a continuum, some readers may rely more heavily on the support from semantic representations, while at the opposite extreme, others may develop good oral reading without any support from semantics (of which children with hyperlexia may be an extreme example). Consequently, following damage to semantic representations, the first group would show a correlation between severity of semantic impairment and degree of surface dyslexia, while the latter group may show little or no surface dyslexia. Although this explanation does seem post hoc, it is a possibility. However, the Patterson and Hodges model originally stemmed from the notion that the link found between surface dyslexia and damage to the semantic route for reading was not just an associative relationship, it was a causal one. The notion of individual differences is invoked to predict that the association is not inevitable, yet it is this association on which the original model was based. Therefore, to keep this approach viable, it seems necessary to specify an independent psychological criterion for quantifying individual differences in patients’ premorbid systems for reading aloud, and hence, whether they should show evidence of surface dyslexia following semantic impairment.

Acknowledgements

We are grateful to Karalyn Patterson and an anonymous reviewer for useful suggestions for revisions. This study was conducted while M.A.L.R. was in receipt of a studentship from the Engineering and Physical Sciences Research Council. We would like to thank Dr Linda Brown for her kind assistance, the staff at the day service that DC attended, and in particular we thank DC for her patience in completing our study with her.

References


Semantic loss without surface dyslexia

Matthew A. Lambon Ralph, Andrew W. Ellis and Sue Franklin

Abstract
Recent models of reading, including some computational models, suggest that error-free performance in reading aloud may be reliant, at least in part, on support from the semantic system. Such models predict that, following semantic impairment, a pattern of acquired dyslexia known as ‘surface dyslexia’ will be an inevitable consequence. We present data from a patient with dementia of Alzheimer type who shows a severe semantic impairment but near error-free performance on reading aloud real words (including low-frequency, irregular and abstract words) and non-words. The data are discussed in light of various models of reading and in relation to previously reported cases.

Journal
Neurocase 1995; 1: 363–69

Neurocase Reference Number:
O22

Primary diagnosis of interest
Dementia of Alzheimer’s type

Author’s designation of case
DC

Key theoretical issue
● The relationship between semantic impairment and surface dyslexia

Key words: surface dyslexia, semantic impairment

Scan, EEG, and related measures
None

Standardized assessment
None

Other assessment
Detailed assessment of reading aloud and comprehension. Written lexical decision and non-word reading results reported (PALPA battery).

Lesion location
● Generalized atrophy presumed

Lesion type
Progressive atrophy presumed

Language
English