CORTEX 45 (2009) 784–790

available at www.sciencedirect.com

ScienceDirect

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

Note

Are the same phoneme and lexical layers used in speech production and comprehension? A case-series test of Foygel and Dell's (2000) model of aphasic speech production

J. Richard Hanley[a,*] *and Lyndsey Nickels*[b]

[a]*Department of Psychology, University of Essex, Colchester, UK*
[b]*Macquarie University, Sydney, Australia*

ARTICLE INFO

Article history:
Received 30 November 2007
Reviewed 10 March 2008
Revised 12 June 2008
Accepted 26 September 2008
Action editor Roberto Cubelli
Published online 14 November 2008

Keywords:
Lexical layer
Phonological layer
Speech comprehension
Speech production

ABSTRACT

In this paper, we investigate the claim that although the same lexical units are involved in speech production and comprehension, there are separate input and output phoneme layers (Foygel and Dell, 2000). Data from a case series of aphasic patients are used to test this claim by examining the relationship between performance on a test of picture naming and performance on tests of phonological input. Estimates of each patient's semantic-lexical and phonological impairments in speech production were derived from Foygel and Dell's computational model of picture naming. It was found that the strength of the semantic-lexical impairments in speech production was significantly correlated with performance on auditory comprehension tests. This finding is consistent with the claim that the same lexical units are involved in speech comprehension and production. Conversely, the correlations between the strength of the phonological lesions in speech production and performance on tests of phonological input were non-significant, consistent with Foygel and Dell's claim that there are distinct input and output phoneme layers.

1. Aims of this study

One enduring controversy in the cognitive neuropsychology of language is the extent to which the processing mechanisms that are involved in speech production are also used in speech perception and comprehension. According to some theorists (e.g., Allport, 1984), the same lexical systems are employed in processing speech input and output. Others such as Morton (1979) and Monsell (1985) have argued that different systems are involved. The debate between these two competing theoretical viewpoints has never been resolved. In the present study, we investigate the validity of an alternative possibility, derived from Foygel and Dell's (2000) computational model of speech production, that can be seen as falling somewhere between these two extreme positions. In Foygel and Dell's model, phoneme units and lexical units (sometimes referred to as 'lemmas') constitute separate levels of representation within the speech production system. According to Foygel and Dell, comprehension of spoken language involves a separate set of phoneme units from those that are used in speech production. Their model therefore assumes the existence of distinct input and output phoneme layers. However, Dell and his colleagues have never proposed that different sets of *lexical* units are required for speech input and output. In this

* *Corresponding author*. Department of Psychology, University of Essex, Colchester CO4 3SQ, UK.
E-mail address: rhanley@essex.ac.uk (J.R. Hanley).
0010-9452/$ – see front matter © 2008 Elsevier Srl. All rights reserved.
doi:10.1016/j.cortex.2008.09.005

paper, therefore, we investigate the claim that the same lexical units are involved in speech production and perception, but there are separate input and output phoneme layers.

We will test this claim by examining the relationship between the errors that aphasic patients make in naming pictures and their performance on tests of phonological input. Previous studies have investigated whether there is any relationship between input processing and the number of semantic errors and phonological errors that patients make in speech production (Martin and Saffran, 2002; Nickels and Howard, 1995). Phonological errors comprised both real word and nonword errors in Nickels and Howard's study, whereas Martin and Saffran's phonological errors were all nonwords. What is different about our investigation is that we will obtain a more precise measure of each patient's difficulties in picture naming by estimating the strength of their phonological and semantic-lexical impairments directly from Foygel and Dell's (2000) computational model of speech production. We will then investigate whether there is any correlation between the severity of each patient's phonological and semantic-lexical lesions in picture naming and their performance on a variety of different phonological input tests.

2. Foygel and Dell's model of speech production

Foygel and Dell (2000) put forward an interactive activation and competition model of the processes involved in normal adult speech production. At the initial stage of spoken picture naming, the semantic features that are associated with the target picture ('cat' in the example in Fig. 1) are activated. As a consequence, lexical units that are linked to these features gain in activation in line with the strength of the associative connections ('weights') between the semantic and lexical units. Under normal circumstances, the lexical unit that is associated with the largest number of activated semantic features will reach threshold and become selected.

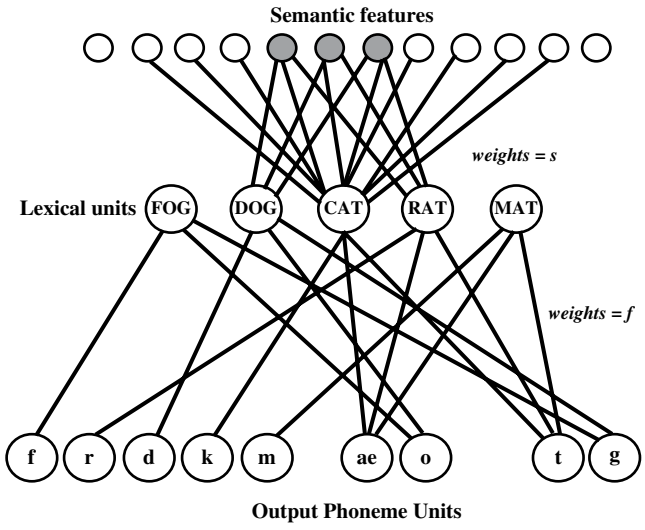

Fig. 1 – Foygel and Dell's (2000) phonological-semantic model of picture naming. It is assumed that a picture of *a cat* is to be named.

Nevertheless, other lexical units share semantic features with the target (in Fig. 1, 'dog' and 'rat' share the shaded semantic features with 'cat') and will also receive some activation. As soon as the lexical units start to become activated, the activation level of the phonological units to which they are connected in the output phoneme layer increases until the appropriate set of phonemes has been generated and a word can be produced. The strength of the associative connections between the lexical and phonological layers determines the amount of activation that individual phoneme units receive. Activation at the phoneme level also provides interactive feedback to phonologically related words at the lexical level. Such activation reduces the chance that a phonologically unrelated lexical unit will reach threshold, but increases the chance that a phonologically related word will be erroneously selected.

When the weights within the system are lesioned by reducing them below those which simulate accurate word production (as found in normal speakers), the model can be used to account for impaired speech production in aphasia for both spoken picture naming (Schwartz et al., 2006) and the auditory repetition of spoken words (Dell et al., 2007; Hanley et al., 2004; Baron et al., 2008). Foygel and Dell (2000) and Schwartz et al. (2006) have shown that it is possible to simulate accurately the number and type of errors that most aphasic patients make in spoken picture naming by lesioning to different degrees the weights between the semantic features and the lexical units (semantic-lexical lesion) and the weights between the lexical units and the output phoneme layer (phonological lesion). For each patient, they simulated overall level of performance and the five types of errors that are observed most frequently in aphasic speech: semantic (e.g., *table* > *chair*), mixed (both semantic and phonological, e.g., *pear* > *peach*), formal (e.g., *chair* > *chart*), nonword (e.g., *anchor* > *antala*), and unrelated errors (e.g., *fork* > *bell*). Patterns of errors were simulated by reducing independently the connection weights in the model between the semantic and lexical level (parameter *s*), and between the lexical and phonological level (parameter *f*). Each patient's naming performance was characterised by the value of *s* and the value of *f* between .001 (most severe lesion) and .1 (unimpaired) that produced the best fit to their naming data. For every patient, the decay rate *q* of all units was set at .6, which is the standard decay rate that the model assumes in an unimpaired system. The best fit was determined by minimising the difference between the obtained and predicted proportions of correct responses and error types, as measured by chi-squared. (See Dell et al., 2004, for the method of searching for the best-fitting parameters).

When the semantic-lexical connections (*s*) are lesioned, there is an increased probability that one of the target word's competitors at the lexical level will be selected. Unless there is further damage to the connections between the lexical and phonological levels (*f*), then phonological encoding will take place normally and a real word will be produced. Semantic-lexical lesions will therefore produce a relatively high incidence of errors in which the response is a familiar word (semantic, formal, mixed, and unrelated word errors). If *f* is lesioned, the probability of incorrect phonemes being selected increases. This will result in a high proportion of nonword