Proc. 14th. Annual Conf. Cognitive Science Soc., Bloomington IN, 1992, pp. 378-383.

Modelling Paraphasias in Normal and Aphasic Speech

Trevor A. Harley & Siobhan B. G. MacAndrew
Department of Psychology
University of Warwick
Coventry CV4 7AL
England
EMAIL: psrds@csv.warwick.ac.uk

Abstract

We model word substitution errors made by normal and aphasic speakers with an interactive activation model of lexicalization. This comprises a three-layer architecture of semantic, lexical, and phonological units. We test four hypotheses about the origin of aphasic word substitutions: that they result from pathological decay, loss of within-level inhibitory connections, increased initial random noise, or reduced flow of activation from the semantic to the lexical level. We conclude that a version of the final hypothesis best explains the aphasic data, but with random fluctuations in connection strength rather than a uniform decrement. This model accounts for aspects of recovery in aphasia, and frequency and imageability effects in paraphasias. Pathological lexical access is related to transient lexical access difficulties in normal speakers to provide an account of normal word substitution errors. We argue that similar constraints operate in each case. This model predicts imageability and frequency effects which are verified by analysis of our normal speech error data.

Introduction

Paraphasias are the erroneous substitution of one word for another in speech. They occur as errors in normal speech (e.g. "warm" -> *cold*), and as word substitutions which arise in the acquired speech deficit of jargon aphasia (see Butterworth, 1985). The speech of these patients is copious but characterized by gross word finding difficulties. We outline how aphasic paraphasias might be explicable within the context of a model of normal speech production. We "lesion" a connectionist model of lexicalization so that it produces paraphasias similar to those of aphasic speakers. Finally, we argue that similar mechanisms are involved in the production of errors in normal speech.

The model of lexical access upon which our simulation is based is close to that of Harley (1990) and Stemberger (1985), and uses an interactive activation architecture (McClelland & Rumelhart, 1981). It accounts for much that is known about normal speech production. It shares some features with that of Dell (1986), but differs importantly in that we use a different type of semantic representation. This enables us to explore semantic word substitutions in a more plausible way. Our model also has intra-level inhibitory connections that have the computational consequence of increasing the effects of within-level competition, and inter-level inhibitory connections that speed up processing by more quickly suppressing inappropriate competitors. These are motivated by findings of inhibitory priming in lexical access in a naming task (Wheeldon, 1989), and phonological blocking in the tip-of-the-tongue state (Jones, 1989).

Units in our model are organized into semantic, lexical, and phonological levels. There is converging evidence in the literature for two stages in speech production, with semantic representations first mapped into abstract lexical forms (our lexical level), followed by the retrieval of phonological forms (e.g. Levelt et al., 1991a). Motivated by speech error data such as phonological facilitation and lexical bias (Harley, 1984; Stemberger, 1985), our model postulates interaction between these stages. There is currently debate about the extent to which these stages are modular (Levelt et al., 1991b). However, interactive models can be shown to be consistent with naming data which at first sight support the modular hypothesis (Dell & O'Seaghdha, 1991).

Each unit in the model is connected to every unit in the following layer. Appropriate between-level connections (such as the lexical unit **dog** to the phonological unit /d/) are excitatory, whereas inappropriate connections (such as **dog** to /k/) are inhibitory. Units within the lexical and phonological levels are completely inter-connected by inhibitory connections. There are feedback connections between the phonological and lexical levels. As usual the net input net_u to a unit u is the sum of the products of all inputs a_j from j units with the weights of the appropriate connections w_{ju}, $\Sigma a_j w_{ju}$. In each cycle the change in activation of unit u is given by the equations:

This research was supported by a grant from the UK Tri-Council Initiative on Cognitive Science/HCI, No. SPG 9018232.

$$\Delta a_u \quad = \quad (max - a_u)net_u - decay(a_u - rest)$$

if $net_u > 0$; otherwise,

$$\Delta a_u \quad = \quad (a_u - min)net_u - decay(a_u - rest)$$

Max and *min* are the maximum and minimum levels of activation, and *rest* a resting level of activation dependent for lexical units upon the frequency of those words and close to zero for other units. Separate parameters control the decay of activation at each level. The degree to which the resting levels of lexical units vary around a mean of zero is determined by the value of a parameter *freqgain* whereby:

$$rest \quad = \quad freqgain * (\log_e(\text{item frequency}) - \text{mean} \log_e(\text{item frequency}))$$

Units also possess a variable amount of normally distributed random noise at the beginning of each processing epoch. The standard deviation of this distribution is determined by a parameter. Our simulated lexicon contains 70 lexical units, which receive input from 26 semantic feature units and send output to 21 phonological units for each of five positions in a serial order phonological output frame. For simplicity, each input feature is a simple on-off binary unit, a semantic representation similar to that of Hinton and Shallice (1991).

The model's architecture is shown in Figure 1. Figure 2 illustrates the normal time course of lexicalization of the word **cow**, the target in all subsequent examples. It shows the activation level of units plotted against processing cycle (time) when the semantic features corresponding to the semantic representation of **cow** are activated. For illustrative purposes, the activations of a semantically and phonologically competing lexical neighbour (**calf**) and of an unrelated control word (**dart**) are also shown, as are the activation values of the target's initial phoneme (/k/) and a non-target phoneme (/d/). Such simulations produce an accurate account of normal lexicalization. The model can also account for findings such as data on the time course of lexicalization in picture naming and facilitation in speech error data.

Lesioning the network

We wish to show that under certain conditions paraphasias are produced by the model when it is nevertheless given the target semantic input. Although it is our goal to produce all types of error, at present we are concentrating upon failures of lexical access. This is a vital first step towards explaining a further phenomenon of jargon aphasia, the production of neologisms (non-words). If there is a clear competitor to the lexical target, a word substitution is likely to occur. If there is either no competitor or a number of equally activated competitors, then the conditions for the generation of neologisms have been met. Finally, we contend that pathological paraphasias are very similar in key respects to normal paraphasias.

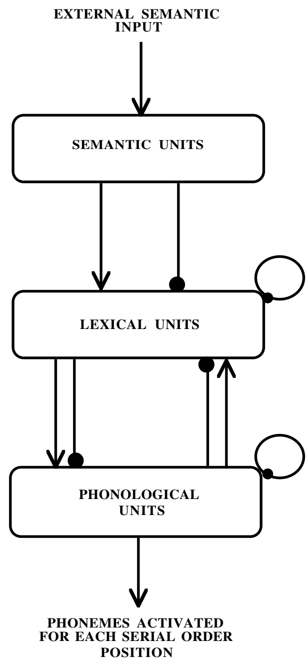

Fig. 1. General architecture of our lexicalization model, with excitatory connections shown by an arrow, inhibitory connections by a filled circle.

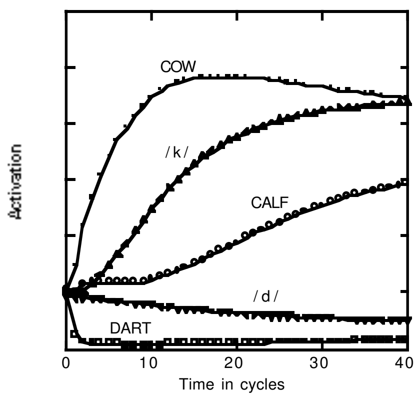

Fig. 2. The time course of a simulation of the normal lexicalization of the word "cow".

To produce substitutions we must somehow disrupt the flow of activation from the input semantic units to the output phonological units. We mimic lesioning by manipulating the parameters of our model. These include both connection strengths, and the control parameters of the network (e.g. the rate of decay of activation, the amount of random noise, the effect of lexical frequency, and the time external semantic input is received by the semantic units). Rather than trying all possibilities, we specifically test predictions derived from the aphasia literature on the origin of paraphasias.

Increase in the Rate of Decay. Martin and Saffran (1991) describe a deep dysphasic patient NC whose speech output includes a large number of paraphasias and neologisms. NC also has a severely restricted phonological short-term memory. They argue that his symptoms arise from a pathological increase in the decay rate of the target lexical nodes. This increases the probability that phonologically or semantically related lexical words will replace the target. If this hypothesis is correct then increasing the parameter controlling the rate of decay of the lexical units in our simulation should increase the probability of paraphasias occurring. In particular, the activation level of the target should decrease as the rate of lexical decay increases, while those of its competitors should increase. The results of these simulations are shown in Figure 3. Although increasing lexical decay causes the activation of the target to fall, it is still considerably higher than those of its close competitors, and remains well above its resting level, even at exceptionally high levels of decay (0.99). Furthermore, with increasing decay, the activation levels of the competitors level off at a low value.

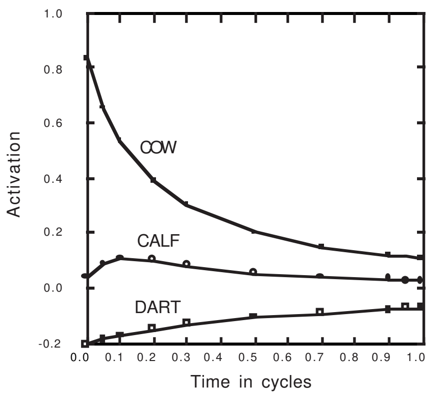

Fig. 3. The effect of increasing the rate of decay of lexical units upon the activation level of the target lexical unit **cow** after 15 processing cycles.

The effect of combining pathological decay with curtailing the time semantic units receive external activation that can in turn be passed on to lexical units is shown in Figure 4. Here the time external input is given to the semantic units is reduced to only 3 cycles. We do not think this is a plausible account of the generation of paraphasias and neologisms for two reasons. First, it requires two simultaneous deficits. Second, though the activation level of the target unit is reduced to near zero, so are those of its competitors. We propose therefore that a pathological increase in lexical decay is unlikely to produce paraphasias.

The Loss of Intra-level Inhibitory Connections. Harley (1990) proposed that the paragrammatisms often associated with neologistic jargon result from excessive blending of syntactic fragments as a consequence of the loss of within-level inhibitory connections. Can a similar mechanism also account for the presence of word substitutions? If so, then decreasing the value of the parameter that controls the degree of intra-lexical inhibition, *gammall*, should decrease the activation of the target unit and increase those of its competitors. Figure 5 shows the effect upon the target activation value for different levels of *gammall*.

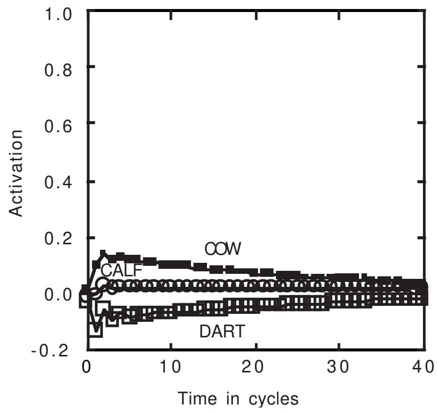

Fig. 4. The effect of combining a pathological increase in the rate of lexical decay (to 0.95) with reducing the amount of time semantic units receive an external input.

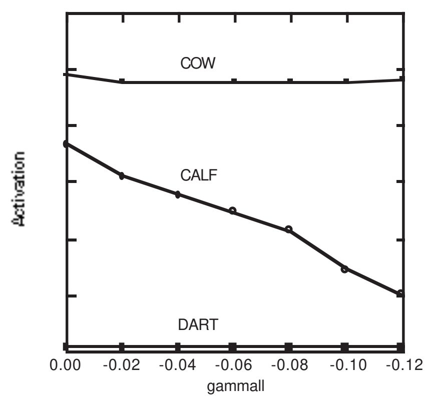

Fig. 5. The effect of decreasing the amount of intra-lexical inhibition upon the activation level of the target unit **cow** after 15 cycles.

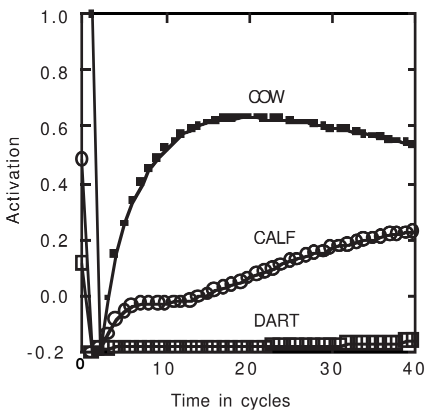

Fig. 6. The time course of activation for a number of competing units with high initial random noise.

Although decreasing the strength of these connections does increase the activation levels of the competitors, it also slightly increases the activation of the target. Even after 25 cycles the target unit is still the most activated, although semantic competitors also have high levels of activation. Such a mechanism then is unlikely to be able to account for the jargon data.

Increased Initial Random Noise. Another possibility is that increasing the amount of initial random noise will increase the probability of paraphasias occurring. Simulations were run with exceptionally high levels of initial lexical noise. Figure 6 shows the effect of increasing the lexical noise level a hundred fold. The target unit quickly recovers and then progresses normally. We can rule out high initial random noise as a causal factor in jargon paraphasias.

Weak Lexical Activation. Our final hypothesis is based on Miller and Ellis (1987). They argue that the impairments found in their patient RD can be explained by difficulty in activating lexical units in the speech output lexicon. They propose that the flow of activation from the semantic level to the lexical level in neologistic jargon aphasics is reduced to a trickle. As units at the lexical level have received insufficient activation, they cannot in turn properly activate the target phonemes. Other phonemes, which have high activation levels due to random noise, are usually accessed in preference.

If this hypothesis is correct then it should be possible to generate substitutions by reducing the value of the parameter that governs the rate of spread of activation between the semantic and lexical levels, *alphasl*. The results are shown in Figure 7. Manipulating *alphasl* does not behave exactly as predicted by the weak lexical activation hypothesis. Although a decrease in *alphasl* does decrease the activation of the target, over part of the range the activation levels of the semantically competing items decrease even more rapidly until very low levels of *alphasl* are reached. Hence the weak lexical activation hypothesis makes an additional prediction: if lower levels of *alphasl* are reflected in increasingly severe symptoms, more severe cases of jargon aphasia should show a lower level of semantic paraphasias relative to other types of word substitutions. We know of little data that address this issue, although Kertesz and Benson (1970) provide evidence from the evolution of aphasia that supports this prediction. They show that during recovery, there is a general progression from neologistic jargon to semantic jargon and then to circumlocutory anomic speech.

At a very low level of *alphasl* (0.0001) the activation level of the **cow** lexical unit has reached 0.662 by only the hundredth processing cycle, and is clearly distinguished from other lexical candidates. This suggests that if jargon aphasics had sufficient time, they would eventually retrieve the correct target. Clearly this is not the case, as lexicalization attempts do not improve over time and do not converge upon the target (Miller & Ellis, 1987). It is necessary to make the further assumption that the semantic units are unable to send activation to the lexical units for more than a fixed time. (Note that this time is not pathologically low, as in Figure 4.) This is consistent with data on the time course of lexicalization in picture naming (Levelt et al., 1991a). Even then, if semantic units send activation to lexical units for only 10 processing cycles, **cow** still reaches an activation level of 0.25 after 100 processing cycles, and the associated /k/ phoneme reaches a level of 0.55.

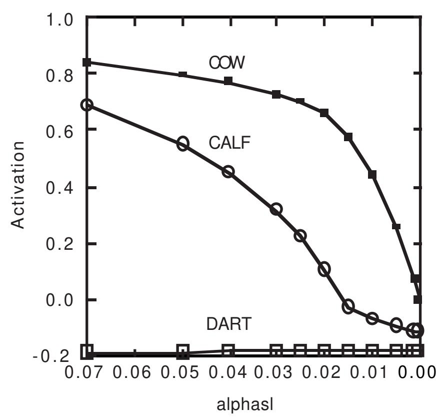

Fig. 7. The effect of reducing the strengths of the excitatory connections between the semantic and lexical levels, *alphasl*.

This hypothesis further predicts an interaction with frequency such that reducing *alphasl* has more effect upon low frequency items. More frequent words should be more robust because they have a higher resting level. They should hence be relatively well preserved at low levels of semantic-lexical facilitation compared with low frequency items. With aphasics, a major determinant of success for a patient attempting to produce a word is its frequency (Ellis, Miller & Sin, 1983). We compute the ratio of the activation level of a lexical unit at a high level of semantic-lexical facilitation (*alphasl* = 0.03) to its activation level at a low level of facilitation (*alphasl* = 0.005) after 15 processing cycles. We call this ratio the *sensitivity ratio* for a particular item as it reflects a lexical unit's sensitivity to different levels of *alphasl*. High frequency items should have lower sensitivity ratios than low frequency items. The sensitivity ratio was computed across a range of lexical frequencies. Figure 8 shows the result of these simulations. Although other factors are clearly operating, frequency does behave as predicted by the weak lexical activation hypothesis. Inspection suggests that the main origin of the residuals in a regression of frequency onto sensitivity ratio is the number of the semantic units that are "on" for any particular lexical item. That is, the effect of lowering the strength of the semantic-lexical connections is moderated not only by frequency, but also by the richness of the underlying semantic representation for each item. We take this to be reflected in the *imageability* of words, in the same way as Plaut and Shallice (1991). Further simulations teased out the differing contributions of lexical frequency and imageability. Two types of simulations were run with artificial lexical items. In the first, the effect of varying the frequency of the target lexical units was investigated while the semantic representation was held constant. In the second, the number of "on" semantic units in the input was varied while the frequency was held constant. In both cases near linear relationships are found between the sensitivity ratio and pure frequency and pure imageability. This further predicts that high imageability words should also be preferentially preserved in jargon aphasia independent of frequency. Again, we know of no data that directly address this issue, though deep dyslexics perform better on more imageable words (Coltheart, 1980).

Although this gives a more satisfying distribution of lexical activations than the other accounts, it still fails to satisfy the criterion that, on some occasions, the activation of competitors should be above that of the target. To achieve this, it is necessary to introduce some random variation into the weakening of the semantic-to-lexical connections. Hence the excitatory semantic-to-lexical connections were randomly lesioned. This was achieved by adding an amount of normally distributed random noise to each connection. The severity of lesioning is mimicked by increasing the standard deviation of the noise distribution. Lexical units then behave as desired (Figure 9). Random lesioning of *alphasl* affects the target lexical unit such that the greater the severity of the lesioning, the lower the probability of the target unit being highly activated. Further, the greater the lesioning, the higher the probability of other lexical units being highly activated. Of course, because this manipulation of *alphasl* is random, actual results vary from trial to trial, and this variation increases as the amount of lesioning increases.

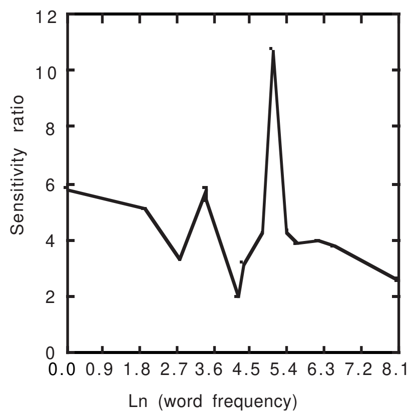

Fig. 8. The sensitivity ratio plotted against $\log_e$ (lexical item frequency).

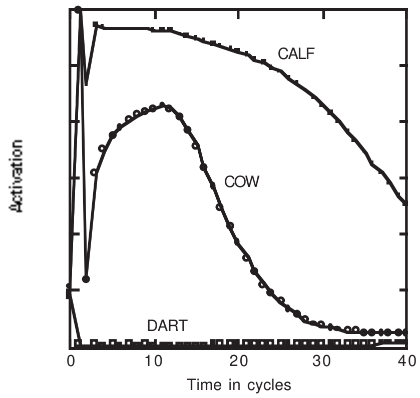

Fig. 9. The effects of randomly lesioning semantic-to-lexical connections upon lexical units. Moderate damage (standard deviation =0.05) has been applied.

Implications for Normal Speech Production

If normal and pathological paraphasias form a continuum, differing only in the amount of random noise that is added to the between-level connections, then they should share many characteristics. In particular, we can make two predictions about frequency and imageability. Those words that are most robust under noisy conditions are going to be the more frequent and imageable words in the language. Hence the words upon which errors occur should be of lower frequency than average, because these are just those items that are particularly susceptible to disruption. We also predict that when normal speakers make a spontaneous word substitution, the target word should be replaced by one more frequent and imageable. It is possible to test these predictions against our corpus of 5468 naturally occurring speech errors. We looked at completed content word substitutions where there was either a semantic or phonological relationship between the target and error words. Both semantic ($t[798] = 4.01$, $p < 0.001$, all results two-tailed) and phonological ($t[448] = 3.62$, $p < 0.001$,) targets were significantly lower in frequency than control words in the corpus. Semantic word substitutions resulted in more imageable words replacing the target ($t[201] = 2.42$, $p < 0.02$), although there was no effect for pure phonological cases ($t[25] = 0.57$, $p > 0.5$). This final result perhaps says no more than that semantic and phonological word substitutions arise at different loci, and that the latter are less affected by semantic constraints. It is a reminder that our model only adequately addresses semantic substitutions. Finally, it has been argued that jargon paraphasias and normal tip-of-the-tongue states share many properties (Miller & Ellis, 1987). Our current simulations suggest that they arise from weakened lexical-to-phonological connections.

Conclusions

The model described here has two important limitations. First, there are no phonotactic constraints: any string of phonemes is permissible. A related problem is that the slot-and-filler mechanism used to implement the serial ordering of phonemes is primitive and inconsistent with the connectionist, non-explicitly rule-based foundations of the model. However, as Miller and Ellis (1987) point out, because phoneme substitutions are random and within-word phoneme exchanges occur no more than would be expected by chance, for RD at least it is not necessary to postulate an additional phoneme ordering mechanism deficit. Another limitation of our work so far is that it is limited to monosyllabic, morphologically simple content words. Nevertheless, lesioning this model by adding noise to the semantic-to-lexical connections can account for a number of important characteristics of jargon aphasic speech.

We would like to conclude by pointing out that connectionist explanations of this type are not inconsistent with earlier hypotheses concerning the origins of jargon, but explain what is happening at a lower level of explanation. Earlier models point to a failure of lexical access; we hypothesize how that failure occurs.

References

Butterworth, B. (1985). Jargon aphasia: Processes and strategies. In S. Newman & R. Epstein eds., *Current perspectives in dysphasia*. Edinburgh: Churchill Livingstone.

Coltheart, M. (1980). Deep dyslexia: A review of the syndrome. In M. Coltheart, K. Patterson, & J. Marshall eds., *Deep dyslexia*. London: Routledge & Kegan Paul.

Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93:283-321.

Dell, G., & O'Seaghdha, P. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt *et al*. (1991). *Psychological Review* 98:604-614.

Ellis, A., Miller, D., and Sin, G. (1983). Wernicke's aphasia and normal language processing: A case study in cognitive neuropsychology. *Cognition* 15:111-144.

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

Harley, T. (1990). Paragrammatisms: Syntactic disturbance or failure of control? *Cognition* 34:85-91.

Hinton, G., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review* 98:74-95.

Jones, G. (1989). Back to Woodworth: Role of interlopers in the tip-of-the-tongue phenomenon. *Memory and Cognition* 17:69-76.

Kertesz, A., & Benson, D. (1970). Neologistic jargon: A clinicopathological study. *Cortex* 6:362-386.

Levelt, W., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T., & Havinga, J. (1991a). The time course of lexical access in speech production: A study of picture naming. *Psychological Review* 98:122-142.

Levelt, W., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T., & Havinga, J. (1991b). Normal and deviant lexical processing: A reply to Dell and O'Seaghdha. *Psychological Review* 98:615-618.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of the basic findings. *Psychological Review* 88:375-407.

Martin, N., & Saffran, E. (1991). A connectionist account of deep dysphasia: Evidence from a single case study. Forthcoming.

Miller, D., & Ellis, A. (1987). Speech and writing errors in "neologistic jargonaphasia": A lexical activation hypothesis. In M. Coltheart, G. Sartori, and R. Job eds. *The cognitive neuropsychology of language*. London: Erlbaum.

Plaut, D., & Shallice, T. (1991). Effects of word abstractness in a connectionist model of deep dyslexia. In Proceedings of the 13th Annual Conference of the Cognitive Science Society, 73-78. Chicago, IL.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis ed. *Progress in the psychology of language* (Vol. 1). London: Erlbaum.

Wheeldon, L. (1989). Competition between semantically related words in speech production. Paper presented to the Experimental Psychology Society, London.