LANGUAGE AND COGNITIVE PROCESSES, 1995, *10* (1), 1–12

A Note on Parallelism Effects in Processing Deep and Surface Verb-phrase Anaphora

Gail Mauner, Michael K. Tanenhaus
Department of Psychology, University of Rochester, Rochester, New York, USA

Greg N. Carlson
Department of Linguistics, University of Rochester, Rochester, New York, USA

Tanenhaus and Carlson (1990, experiment 1) reported that *surface* anaphors were more difficult to comprehend than *deep* anaphors when the antecedent for the anaphor did not form a syntactic constituent, i.e. when it was presented in a passive construction as compared to an active construction and the anaphor was in active voice. However, parallelism of the antecedent did affect the time it took readers to comprehend deep anaphors. This paper reports a reanalysis of that data and two experiments based on the reanalysis. The parallelism effects for deep anaphors were completely eliminated following short passives but not full passives. The results are interpreted as support for the claim that deep and surface anaphors access different types of representations. We also suggest that previous findings of parallelism effects for deep anaphors may be due to factors other than structural parallelism.

INTRODUCTION

Hankamer and Sag (1976) proposed that anaphoric expressions can be divided into two classes: *deep* anaphors and *surface* anaphors. Sag and Hankamer (1984) subsequently claimed that deep anaphors and surface

Requests for reprints should be addressed to Gail Mauner, Department of Psychology, University of Rochester, Rochester, NY 14627, USA. Email: phooph@psych.rochester.edu.

This work was supported by NIH Grant HD27206 awarded to M.K.T. and G.N.C. Preliminary versions of this report have appeared as a poster at the 1991 CUNY Sentence Processing Conference, held at the University of Rochester, Rochester, NY and in the Proceedings of the 1991 Meeting of the Northeastern Linguistics Society.

anaphors differ in the level of representation which must be accessed in order for comprehenders to find their antecedents. They argued that comprehenders find antecedents for deep anaphors at a conceptual level of representation within a discourse model or mental model (Garnham, 1981; Johnson-Laird, 1983; Webber, 1981), whereas antecedents for surface anaphors must be found initially at a linguistic level of representation.

The deep–surface distinction unifies two observations about anaphoric expressions. The first is that some anaphors require the presence of an antecedent that has been introduced linguistically. The second is that there seems to be a clear constraint on the form a linguistically expressed antecedent may take for just those anaphors that *require* a linguistic antecedent. This paper will focus on the processing consequences of the second claim.[1]

Hankamer and Sag (1976) referred to the restriction on the linguistic form of antecedents of surface anaphors as a *parallelism* constraint. To a first approximation, this means that an antecedent must be a linguistic constituent that could felicitously replace or complete an anaphoric expression (allowing for modifications required for morphological agreement). This parallelism requirement is illustrated in (1) and (2).

1a. Someone has to ***transcribe this tape***. (Parallel antecedent)
b. ***This tape*** has to be ***transcribed***. (Non-parallel antecedent)
c. I'm sure John won't *want to* ***do it***. (Deep anaphor)
d. I'm sure John won't *want to* _____. (Surface anaphor)

The interpretation of the surface anaphor *want to* _____ and the deep anaphor *want to* ***do it*** is roughly *I'm sure John won't want to* ***transcribe this tape***. In sentence (1a), the antecedent for the anaphor—the verb phrase *transcribe this tape*—is a linguistic constituent. However, in (1b) there is no corresponding single linguistic constituent. Intuition suggests that the deep anaphor in (1c) is fully felicitous with either antecedent context. In contrast, the surface anaphor in (1d) is fully felicitous only when it is preceded by sentence (1a), which contains a parallel antecedent. When it is preceded by the non-parallel antecedent sentence (1b), the surface anaphor (1d) seems less felicitous and more difficult to comprehend.

[1]There are some exceptions to the generalisation that a surface VP anaphor (VP ellipsis) cannot be pragmatically controlled (Chao, 1987; Hankamer & Sag, 1976; Hardt, 1993; Schacter, 1977). Hardt (1993) presents a few examples that are particularly compelling. However, these examples are restricted to models which Hankamer (1978) argues are not generally productive and therefore are not, like VP ellipsis, syntactically derived.

A number of experiments have tested the Sag and Hankamer hypothesis by manipulating the amount of material that intervenes between the sentence containing the antecedent of a subsequent anaphor and the anaphor itself (Garnham & Oakhill, 1987; Murphy, 1985a; 1990; Tanenhaus, Carlson, & Seidenberg, 1985). The logic underlying this manipulation is the well-established finding that memory for linguistic form, except under unusual circumstances, is fleeting, whereas memory for content is quite stable (Anderson, 1974; Dooling & Christiaansen, 1977; Garnham, 1982; Johnson-Laird & Stevenson, 1970; Sachs, 1967). Since the interpretation of a surface anaphor is dependent on accessing a linguistic antecedent, interpretation of a surface anaphor should become more difficult relative to a deep anaphor as the distance between the antecedent and the anaphoric expression increases. In general, these studies have found distance effects for both surface and deep anaphors. However, the results are somewhat difficult to interpret because of the presence of large parallelism effects for deep anaphors, which suggest that the material intervening between the antecedent and the anaphor may have introduced focus or topic shifts that affected the accessibility of their antecedents (Anderson, Garrod, & Sanford, 1983; Garnham, 1987; Lesgold, Roth, & Curtis, 1979). The one study that used stative sentences as intervening material in an attempt to control for these factors found distance effects for surface anaphors but not for deep anaphors (Tanenhaus et al., 1985).

A more straightforward interpretation of Sag and Hankamer's hypothesis is that a surface anaphor ought to be more difficult to comprehend than a deep anaphor that has approximately the same interpretation, when the antecedent is not a linguistic constituent. Tanenhaus and Carlson (1990, experiment 1) reported results that, for the most part, supported this prediction. In the parallel conditions, the antecedent was the verb phrase of an active sentence and in the non-parallel conditions the antecedent was introduced in a passive construction. Sentences containing an anaphoric expression were in active voice. Subjects were timed as they decided whether the sentence containing the anaphoric expression made sense in the context of the sentence containing the antecedent. Subjects were more likely to judge a surface anaphor to make sense when its antecedent was parallel than when it was non-parallel, whereas parallelism did not affect judgements to sentences containing deep anaphors. However, for those sentences judged to make sense, parallelism affected judgement times for both surface *and* deep anaphors. There was a 705 msec parallelism effect for the surface anaphors and a 149 msec parallelism effect for the deep anaphors. Despite the fact that the parallelism effects were numerically larger for the surface anaphors, the interaction was not

statistically significant. Moreover, the simple effect of parallelism was reliable for deep anaphors.[2]

In this paper, we report a reanalysis of Tanenhaus and Carlson's data and two follow-up experiments based on that reanalysis. To anticipate, the results demonstrate that parallelism effects for deep anaphors were completely eliminated with short passives but not full passives. This suggests that factors other than syntactic parallelism may have been responsible for the "parallelism" effect with deep anaphors in this study and, perhaps, in other studies reported in the literature as well. We will return to this issue briefly after we present the results of the reanalysis and the follow-up experiments.

A REANALYSIS

The passive sentences used by Tanenhaus and Carlson (1990) contained a mix of full passives with agent *by*-phrases and agentless short passives. While there are no obvious structural differences between short and full passives with respect to the syntactic parallelism of the antecedent they provide for an active voice VP anaphor, many of the sentences with full passives were somewhat awkward. This prompted us to conduct an item analysis in which we divided the items into those with full passives (9 items) and those with short passives (10 items). The mean judgement times for item groups are given in Table 1. When response times for deep and surface anaphors in the short passive group were compared, a reliable interaction obtained [$F_2(1, 17) = 7.72$, $P < 0.05$]. There was only a 22 msec difference between judgement times to deep anaphors with parallel antecedents and deep anaphors with non-parallel antecedents, whereas surface anaphors took 823 msec longer to judge when they followed non-parallel antecedents compared with when they followed parallel antecedents. In contrast, when judgement times to deep and surface anaphors in the full passive group were compared, neither the type of anaphor × parallelism of antecedent interaction nor the main effect of type of anaphor were significant ($F_2 < 1$). The effect of parallelism, however, was significant [$F_2(1, 17) = 5.03$, $P < 0.05$]. While judgements to both deep and surface anaphors were longest following non-parallel antecedents, they did not differ significantly from each other ($F_2 < 1$).

[2]Murphy (1985a) also tested the effects of syntactic parallelism, however parallelism was manipulated in interaction with length of antecedent and amount of material intervening between an antecedent and an anaphor. Although Murphy found only inconsistent effects of parallelism, when one looks at the clearest comparison, namely short antecedents with no intervening material, the effect of parallelism was numerically larger for surface anaphors (for details, see Murphy, 1982; Tanenhaus & Carlson, 1990). Moreover, while materials for Murphy (1985a) were not available, the example set of materials provided suggests that at least some of the passives used for non-parallel antecedents contained agent *by*-phrases.

TABLE 1
Mean Judgement Times (msec) to Surface and Deep Anaphors Following Parallel and Short Passive and Full Passive Non-parallel Antecedents

	Type of Antecedent			
	Short Passive		*Full Passive*	
Type of Anaphor	*Parallel*	*Non-parallel*	*Parallel*	*Non-parallel*
Deep	2134	2112	2158	2478
Surface	2063	2891	2159	2741

Note: Judgement latencies are reported only for those sentences judged to make sense.

These analyses suggest that the variability introduced by differences between short and full passives may have masked an interaction between anaphor type and parallelism. More importantly, the short passive data seem to provide clear support for Sag and Hankamer's prediction that deep anaphors are not subject to structural constraints. However, these analyses are obviously *post hoc* and there were too few items across conditions for reliable analyses by subjects. Moreover, the heterogeneous nature of the stimuli may have been an additional source of variability in that Tanenhaus and Carlson's materials employed a variety of linguistic devices to form both deep and surface anaphors.[3] We therefore tried to establish independently, with a homogeneous set of materials, whether the deep–surface parallelism prediction holds for short passives but not full passives. Experiments 1 and 2 were conducted for this purpose. In Experiment 1, we used short agentless passives to determine whether, as the reanalysis suggested, parallelism effects would be found only for surface anaphors. In Experiment 2, we used full passives, expecting to see parallelism effects for deep anaphors as well as for surface anaphors.

EXPERIMENT 1

Method

Subjects. Thirty-two native English-speaking undergraduates at the University of Rochester received either partial credit or a minimal sum for their participation.

[3]Murphy (1985b) has argued that employing non-equivalent forms of deep and surface anaphors may lead to interpretive differences, and recent evidence suggests that deep anaphors formed by adding *do it* to a sentence formed by sluicing (one of the anaphor types used by Tanenhaus and Carlson) introduces parallelism confounds (Shenkman, 1993).

Materials. The experimental materials consisted of pairs of context and target sentences. Sentence pairs were drawn from 28 sets of sentences like those given in (2). Each set contained two context sentences, an active (parallel) sentence with the word *someone* as subject and a short passive (non-parallel) sentence, and two target sentences, one ending in a surface anaphor formed by VP ellipsis, and one ending in a deep anaphor formed by adding *do it* to the VP ellipsis form. Both target sentences were in active voice and expressed the same content regardless of the type of anaphor used. For each stimulus set, each context sentence was paired with each target sentence to form four sentence pairs which were counterbalanced across four presentation lists. The 28 experimental items in each list were intermixed with 56 filler sentence pairs, 25% of which were constructed so that the second sentence would not make sense given the first sentence.

2a. Someone needs to ***feed the kitten***. (Parallel antecedent)
b. ***The kitten*** needs to be ***fed***. (Non-parallel antecedent)
c. Joey *forgot to* ____ again. (Surface anaphor)
d. Joey *forgot to do it* again. (Deep anaphor)

Procedure. The subjects performed a timed make-sense judgement task. They were directed to respond as quickly as possible without sacrificing comprehension. A button press revealed two rows of dashes with each dash corresponding to a non-white space character in a sentence pair. The subjects pressed a button to reveal a context sentence containing either a syntactically parallel or non-parallel antecedent on the first line. After reading the context sentence, another button press replaced the first sentence with dashes and revealed the target sentence containing either a surface or a deep anaphor on the second line. The subjects were instructed to press a "yes" button if they thought that the target sentence made sense given the context sentence, and a "no" button if it did not.

Results and Discussion

Separate analyses of variance were conducted on the proportion of sentences judged to make sense (judgement data), and on the response latencies for sentences judged to be sensible (latency data). The means for the judgement and latency data are presented in Table 2.

Judgements. Both the main effect of type of anaphor [$F_1(1, 28) = 14.85$, $P < 0.01$; $F_2(1, 24) = 16.40$, $P < 0.01$] and parallelism of antecedent [$F_1(1, 28) = 5.61$, $P < 0.05$; $F_2(1, 24) = 8.79$, $P < 0.01$] were significant. The crucial interaction between anaphor type and parallelism of antecedent was also significant in analyses by subjects [$F_1(1, 28) = 10.63$, $P < 0.01$] and by items [$F_2(1, 24) = 10.23$, $P < 0.01$]. Moreover, the simple effect of parallelism was not reliable for deep anaphors (F's < 1).

TABLE 2
Percentage of Deep and Surface Anaphors Judged to Make Sense and Judgement Latencies (msec) as a Function of the Parallelism of the Antecedent when Non-parallel VP Antecedents are Short Passives

	Type of Antecedent			
	Parallel		*Non-parallel*	
Type of Anaphor	*% Judged Sensible*	*Latency (msec)*	*% Judged Sensible*	*Latency (msec)*
Deep	96	3075	96	3149
Surface	94	2979	82	3531

Note: Judgement latencies are reported only for those sentences judged to make sense.

Latencies. Five scores which were over 10 sec were not used to calculate the means for individual subjects and items. The effect of parallelism of antecedent was significant [$F_1(1, 28) = 15.45$, $P < 0.01$; $F_2(1, 24) = 11.91$, $P < 0.01$], but no effect of anaphor type was found (F's < 1). Crucially, the effect of primary interest, the interaction between anaphor type and parallelism of antecedent, was significant [$F_1(1, 28) = 7.55$, $P < 0.01$; $F_2(1, 24) = 5.03$, $P < 0.05$]. Planned comparisons revealed no effect of parallelism for deep anaphors (F's < 1).

To summarise, surface anaphors were judged to make sense more often when their antecedents were syntactically parallel. In contrast, judgements to deep anaphors were unaffected by the syntactic form of their antecedents. Moreover, in contrast to previous studies, we found a clear interaction between anaphor type and parallelism of antecedent for both judgements *and* latencies. Additionally, there was no suggestion of a parallelism effect in either the judgement or the latency data.

EXPERIMENT 2

In this experiment, the materials from Experiment 1 were modified to form full passives to see whether parallelism effects would be seen for deep anaphors.

Method

Subjects. Thirty-two native English-speaking undergraduates attending the University of Rochester received partial course credit or a minimal sum for their participation.

Materials and Procedure. Twenty sentence sets from the homogeneous set of experimental materials used in Experiment 1 were modified so that all non-parallel context sentences contained an agent *by*-phrase. This modification was the only way in which these materials differed from those in Experiment 1. The phrase *by someone* was used so that the interpretation would closely match the interpretation of the parallel context sentences and the non-parallel agentless passives in Experiment 1. A sample set of materials is provided in (3). The procedure in Experiment 2 was identical to that of Experiment 1.

3a. Someone needs to ***feed the kitten***. (Parallel antecedent)
b. ***The kitten*** needs to be ***fed*** by someone. (Non-parallel antecedent)
c. Joey *forgot to* _____ again. (Surface anaphor)
d. Joey *forgot to do it* again. (Deep anaphor)

Results and Discussion

As with Experiment 1, separate analyses of variance were conducted on the proportion of sentences and response latencies for sentences judged to make sense. The means for the judgement and latency data are presented in Table 3.

Judgements. The only effect to reach significance was parallelism of antecedent [$F_1(1, 28) = 9.03$, $P < 0.01$; $F_2(1, 16) = 8.11$, $P < 0.01$]. Planned comparisons revealed an effect of parallelism for surface anaphors [$F_1(1, 28) = 5.44$, $P < 0.05$; $F_2(1, 16) = 7.40$, $P < 0.05$], with a non-significant trend in the same direction for deep anaphors [$F_1(1, 28) = 2.93$, $P < 0.10$; $F_2(1, 16) = 2.12$, $P < 0.16$]. Thus, both surface and deep anaphors were judged to make sense less often following a non-parallel antecedent.

TABLE 3
Percentage of Deep and Surface Anaphors Judged to Make Sense and Judgement Latencies (msec) as a Function of the Parallelism of the Antecedent when Non-parallel VP Antecedents are Full Passives

	Type of Antecedent			
	Parallel		*Non-parallel*	
Type of Anaphor	*% Judged Sensible*	*Latency (msec)*	*% Judged Sensible*	*Latency (msec)*
Deep	96	2674	92	2884
Surface	95	2701	86	2916

Note: Judgement latencies are reported only for those sentences judged to make sense.

Latencies. Prior to statistical treatment, the responses of one subject were eliminated because there were too few remaining "yes" responses (three or less) from which to form a reliable condition mean for that individual. Additionally, three scores which were greater than 10 sec were eliminated from the calculations for the subject and item means. Only the effect of parallelism reached significance. It was reliable when subjects were random and just missed significance in the analysis by items [$F_1(1, 27) = 14.67$, $P < 0.01$; $F_2(1, 16) = 4.21$, $P = 0.06$]. Judgements were longer to deep anaphors following non-parallel antecedents in analyses by subjects and by items [$F_1(1, 27) = 5.71$, $P < 0.05$; $F_2(1, 16) = 4.31$, $P < 0.05$], while judgements to surface anaphors were reliably longer following non-parallel antecedents only in the analysis by subjects [$F_1(1, 27) = 7.22$, $P < 0.01$; $F_2(1, 16) = 2.47$, $P = 0.14$].

The judgement results are similar to, although somewhat weaker than, those of Tanenhaus and Carlson (1990). Likewise, the latency results are also quite similar to those obtained by Tanenhaus and Carlson (1990) and Murphy (1985a). Tanenhaus and Carlson's (1990) results are, in fact, intermediate with respect to the results obtained in Experiments 1 and 2. Presumably, this is because only half of their non-parallel antecedents contained an agent *by*-phrase and half were short passives. There was one puzzling aspect of the data. The parallelism effects for the surface anaphors in both judgements and judgement times were smaller than those found in Experiment 1. One possibility is that because the full passives with *by someone* were somewhat awkward, the subjects may have adopted a looser criterion and thus were more likely to respond "yes" to surface anaphors with non-parallel antecedents.

GENERAL DISCUSSION

These experiments and the reanalysis of Tanenhaus and Carlson's (1990) data suggest that, at least under one set of conditions, syntactic parallelism affects the comprehension of surface anaphors but not deep anaphors. They also suggest that previous findings of parallelism effects for deep anaphors may have been due to factors unrelated to syntactic parallelism. Tanenhaus et al. (1985) suggested that longer reading or judgement times to deep anaphors following distant antecedents found in previous studies were due to focus or topic shifts introduced by intervening material. When these factors were controlled for, they observed distance effects for surface anaphors only.

Tanenhaus and Carlson (1990) extended a variant of the focus/topic shift explanation to account for why the parallelism of an antecedent was found to affect both surface and deep anaphors. They suggested that in passives, in contrast to their active counterparts, information is structured to give special prominence to the subject noun phrase. This results in an interpretation that is less eventive. Because the antecedent of a verb-phrase anaphor in a discourse model is an event, this shift in focus has the effect of making the antecedent less accessible. However, the results of Experiment 1 cast doubt on this explanation. If an antecedent event is made less accessible in a discourse model when it is presented in a passive construction, then this effect should have been observed for short passives as well.

We have demonstrated that with short passives, only surface anaphors display clear parallelism effects. We have not explored possible reasons for why the presence of *by*-phrases should lead to longer reading and judgement times to deep anaphors other than to note that it seems unlikely that syntactic parallelism was the relevant factor. Some of the full passive sentences used by Tanenhaus and Carlson (1990) are clearly awkward (e.g. *The valuable antique vase which belonged to Mrs. Jones was broken by John* and *The clothes were washed by Mary every weekend*). Likewise, the full passives in Experiment 2 are also somewhat unnatural because the phrase *by someone* is awkward except in quite restricted contexts. However, it seems unlikely that the awkwardness of full passives has to do with the construction itself. Consider the examples in (4):

4a. Our dishwasher will have to be repaired by an expert.
b. An expert will have to repair our dishwasher.
c. Mom thinks we should ask the Maytag repairman to do it.

The full passive in (4a) appears to be as natural as its active counterpart in (4b). Thus we would expect that the processing of a deep anaphor such as (4c) would not differ following a full passive such as (4a) as opposed to an active such as (4b). Under such circumstances, it is also unlikely that the type of passive would make any difference to how a deep anaphor such as (4c) is processed.

In sum, the work reported here demonstrates clear effects of syntactic parallelism for surface but not deep anaphors. While the results support predictions derived from Sag and Hankamer's hypothesis, more definitive research on parallelism effects is clearly needed to evaluate fully their processing claims. In this work, it will be crucial to separate effects that are likely to be due to pragmatic/stylistic factors from those that are structural.

Manuscript received June 1993
Revised manuscript received December 1993

REFERENCES

Anderson, A., Garrod, S.C., & Sanford, A.J. (1983). The accessibility of prenominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology*, *35A*, 427–440.

Anderson, J. (1974). Verbatim and propositional representation of sentences in immediate and long-term memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 149–162.

Chao, W. (1987). *On ellipsis*. PhD dissertation, UMASS Amherst. Cambridge, MA: Graduate Linguistics Student Association.

Dooling, D.J., & Christiaansen, R.E. (1977). Levels of encoding and retention of prose. In G.H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 11. New York: Academic Press.

Garnham, A. (1981). Mental models as representations of text. *Memory and Cognition*, *9*, 560–565.

Garnham, A. (1982). Testing psychological theories about inference-making. *Memory and Cognition*, *10*, 341–349.

Garnham, A. (1987). Understanding anaphora. In A.W. Ellis (Ed.), *Progress in the psychology of language*, Vol. 3. Hove: Lawrence Erlbaum Associates Ltd.

Garnham, A., & Oakhill, J. (1987). Interpreting elliptical verb phrases. *Quarterly Journal of Experimental Psychology*, *39A*, 611–627.

Hankamer, J. (1978). On the non-transformational nature of some null VP anaphors. *Linguistic Inquiry*, *9*, 66–74.

Hankamer, J., & Sag, I.A. (1976). Deep and surface anaphora. *Linguistic Inquiry*, *7*, 391–426.

Hardt, D. (1993). *Verb phrase ellipsis: Form, meaning, and processing*. Unpublished PhD dissertation, University of Pennsylvania.

Johnson-Laird, P. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P., & Stevenson, R. (1970). Memory for syntax. *Nature*, *227*, 412.

Lesgold, A., Roth, S., & Curtis, M.E. (1979). Foregrounding effects in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior*, *18*, 291–308.

Murphy, G.L. (1982). *Understanding anaphora*. Unpublished doctoral dissertation, Stanford University.

Murphy, G.L. (1985a). Processes of understanding anaphora. *Journal of Memory and Language*, *24*, 290–303.

Murphy, G.L. (1985b). Psychological explanations of deep and surface anaphora. *Journal of Pragmatics*, *9*, 785–813.

Murphy, G.L. (1990). Interpretation of verb phrase anaphora: Influences of task and syntactic context. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *42A*, 675–692.

Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, *2*, 437–442.

Sag, I.A., & Hankamer, J. (1984). Towards a theory of anaphoric processing. *Linguistics and Philosophy*, *7*, 325–345.

Schacter, P. (1977). Does she or doesn't she? *Linguistic Inquiry*, *7*, 664–671.

Shenkman, K. (1993). *Structure sensitivity and language processing in adult learners of English*. Unpublished PhD dissertation, University of Rochester.

Tanenhaus, M.K., & Carlson, G.N. (1990). Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes*, *5*, 257–280.

Tanenhaus, M.K., Carlson, G.N., & Seidenberg, M.S. (1985). Do listeners compute linguistic representations? In D. Dowty, L. Kartunnen, & A. Zwicky (Eds), *Natural language*

parsing: Psychological, computational and theoretical perspectives. Cambridge: Cambridge University Press.

Webber, B.L. (1981). Discourse model synthesis: Preliminaries to reference. In A.K. Joshi, B.L. Webber, & I.A. Sag (Eds), *Elements of discourse understanding*. Cambridge: Cambridge University Press.

LANGUAGE AND COGNITIVE PROCESSES, 1995, *10* (1), 13–45

Getting it Right? Using Aphasic Naming Errors to Evaluate Theoretical Models of Spoken Word Recognition

Lyndsey Nickels

Department of Psychology, Birkbeck College, University of London, London, UK

Different models of spoken word production make different predictions regarding the extent of effects of certain word properties on the output of that model. These predictions are examined with regard to the effect of these variables on the production of semantic and phonological errors by aphasic subjects. Thus the production of semantic errors is found to be affected by imageability and the production of phonological errors by word length, but not vice versa. It is argued that this pattern of variables affecting the production of semantic and phonological errors is better explained by models which require strictly sequential stage-by-stage processing (e.g. Levelt et al., 1991a; Morton, 1970; 1979) than by multi-layer perceptron (e.g. Plaut & Shallice, 1991; 1993) or interactive activation models (e.g. Dell, 1986; 1989).

INTRODUCTION

Current models of spoken word production may be classified into two main types: (1) strictly sequential stage models, where semantic and phonological aspects of lexical access are represented by separate, temporally distinct stages (e.g. Garrett, 1980; 1984; Levelt, 1983; Morton, 1970; 1979), and (2) connectionist or spreading activation models (e.g. Dell, 1986; 1989; Stemberger, 1985), where any activation spreads continuously through the network and, unlike in stage models, processing need not be complete at one level before activation is transmitted to the next. This latter type of

Requests for reprints should be addressed to Lyndsey Nickels, Department of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK.

I would like to thank Tim Shallice and David Howard for helpful discussions during the preparation of this paper. Gerry Altmann and two anonymous reviewers gave useful comments on an earlier draft of the manuscript. The work was carried out while the author was supported by MRC Project Grant No. G8903890N.