Authors' Response

Feedback on feedback on feedback: It's feedforward

Dennis Norris,[3] James M. McQueen," and Anne Cutler*
^*Medical Research Council Cognition and Brain Sciences Unit, Cambridge, CB22EF, United Kingdom;*[b]*Max-Planck-Institute for Psycholinguistics, 6525 XD Nijmegen, The Netherlands;* **dennis.norris@mrc-cbu.cam.ac.uk www.mrc-cbu.cam.ac.uk** **james.mcqueen; anne.cutler@mpi.nl www.mpi.nl**

Abstract: The central thesis of our target article is that feedback is never necessary in spoken word recognition. In this response we begin by clarifying some terminological issues that have led to a number of misunderstandings. We provide some new arguments that the feedforward model Merge is indeed more parsimonious than the interactive alternatives, and that it provides a more convincing account of the data than alternative models. Finally, we extend the arguments to deal with new issues raised by the commentators such as infant speech perception and neural architecture.

R1. Definitions

Many commentators' points rest on particular interpretations of the terms "top-down" and "interaction." In several cases, commentators have used these terms quite differently from the way we used them in the target article and, in some cases, quite differently from the way they have used

them in their own previous writings. When we introduced these concepts we made it clear that we were using them in the senses which most closely correspond to the notion of feedback. Remember, feedback is the central issue, not only because it is the focus of the target article, but because it is the focus of the debate in the literature.

R1.1. Interaction. In the target article, we used the term interaction as synonymous with feedback. Two stages which interact are linked by feedback as well as feedforward connections, that is, each can influence the other.

Although "interaction" is most commonly used to characterise the information flow between processes, interaction is sometimes used instead to make statements about how different kinds of information are used or combined. So, if lexical and phonemic knowledge are combined in making phonemic decisions, one might want to say that lexical and phonemic knowledge interact. We can call these two senses of interaction "process interaction" and "information interaction" (Norris 1980). Information interaction does not imply process interaction. For example, one might (like **Luce et al.**) make no distinction between lexical and phonemic processes, but still characterise lexical and phonemic information as different kinds of knowledge. In Merge, lexical and phonemic knowledge are combined in the decision nodes, but no processes in the model interact with one another. Merge has no process interaction and no feedback. We have therefore not adopted a narrow definition of interaction as **Pitt** and **Slowiaczek** suggest, but we have tried not to conflate the two quite distinct senses of interaction.

R1.2. Top-down. The sense of top-down which predominates in the psychological literature refers to the direction of information flow within the system. In this architectural sense, flow of information from one process back to previous processes in the chain is referred to as top-down. Merge is not top-down. Lexical units give output only to decision units which are themselves output units and are not part of the processing chain delivering input to lexical units. Note that top-down does not refer to the direction of lines on the page. If it did, classification of models would depend how one drew the lines, and possibly on how one held the page! We leave it as an exercise for the reader to redraw Figure 1 so that all lines from the input phoneme and lexical levels point upwards.

Note that although this sense of top-down gets close to the concept of feedback, and is generally used synonymously with feedback in the literature, it is not identical. Nonspecific top-down flow of information, such as might be involved in generalised attentional activation, would not in any way be the same as specific feedback from particular lexical items which altered the processing of specific phonemes. The target article concerns itself with specific feedback, and not with nonspecific top-down effects, such as attention, which are not part of a lexicon-phoneme feedback loop.

Top-down is also sometimes used in a less well-defined sense that appears to be a close synonym of information interaction. In this second sense, top-down is used to mean that information at one level of analysis is brought to bear in processing information specifiable at a more fine-grained level of description. So if lexical knowledge is used in any way to influence decisions about phonemes, this is evidence

that lexical and phonemic information are combined in a top-down fashion. As the target article demonstrates, this is quite independent of the issue of feedback, or even direction of information flow. Strictly feedforward models like the Race model, Merge, and FLMP are, in this sense, top-down. (Thus although both Shortlist and Merge are feedforward and bottom-up in terms of information flow, they do use the lexical constraints that **Benki** wants them to use.)

In concluding this section we emphasise that our choice of terminology and definitions is not arbitrary. It reflects the core issues in the "interaction" debate that has been pursued in the literature for over 20 years. This has been part of a more general debate about the architecture of the language processing system, perhaps most clearly set out by Forster (1979). During this time there has been no disagreement about whether lexical information can influence phoneme identification. The debate has been about process interaction and feedback (Cutler et al. 1987; Elman & McClelland 1988; Frauenfelder et al. 1990; Massaro & Cohen 1991; McClelland 1991; Pitt 1995; Pitt & McQueen 1998; Samuel 1997).

R2. Theory

None of the commentaries has explained why feedback might be necessary. **Tanenhaus et al., Montant, Shillcock,** and **Stevens** all express their conviction that it really should (under certain circumstances) be helpful. But without specific reasons why our arguments might not hold under such circumstances, pleas like "feedback is surely helpful" (Tanenhaus et al.) remain wishful thinking. In the following sections we discuss the general points of theory that were raised. The majority of the commentaries have concentrated on issues concerning the Merge model itself, raising three main concerns: that Merge might not be "ecologically valid"; that, contrary to our characterisation, Merge might really be a top-down or interactive model after all; and that Merge might not really be simpler than interactive models.

R2.1. Ecological validity. In order to make the case for a feedforward model of speech perception we must be able to explain data from laboratory tasks that have been presented as evidence for feedback. Merge was designed to explain these data in a manner consistent with the principles of Shortlist, which is concerned with modelling word recognition in continuous speech. Some commentators question the ecological validity of Merge (**Appelbaum, Benki, Vroomen & de Gelder).** After all, Merge has been used to explain behaviour in laboratory tasks involving metalinguistic judgements. In part this is true. None of us is primarily concerned with explaining laboratory data rather than naturalistic processing. Psycholinguists have to live with the fact that the experimental tasks they use do not directly reveal the inner workings of the speech perception system. These tasks do, however, give us some very good clues, whereas naturalistic observation of speech perception tells us nothing at all about processing architecture. To make the best use of these clues, models like Merge must attempt to explain both normal processing and performance in laboratory tasks. The data that Merge explains have on occasion been taken as evidence for feedback, so we cannot ignore

these data. The commentators who criticise the ecological validity of Merge present no alternative.

R2.2. Terminological confusion. Some commentators seem to be in a state of terminological confusion. This worries us because it indicates that there is confusion over the use of some fundamental terms in the literature. More worrying still is the fact that some commentators (**Pitt, Samuel**) who have published papers using terms like "top-down" in the standard sense of direction of information flow used in the target article, use the terms in a quite different sense in their commentaries.

Appelbaum and **Samuel** suggest that the interaction debate has not been about processing interaction and information flow (see sect. R1, Definitions) and that we should now call feedforward models like Merge and FLMP interactive. **Pitt** believes that we have narrowed the meaning of "interactivity" by restricting it to cover only top-down feedback. Interestingly enough, if we look at recent papers on interaction written by Pitt and Samuel, we see that the opening paragraph of each of these papers defines both the terms and the issues very clearly (Pitt 1995; Pitt & Mc-Queen 1998; Pitt & Samuel 1993; Samuel 1997). We quote here from Samuel (1997, p. 97) (although the clearest definition is to be found in Pitt & McQueen), "Some models hypothesize strictly bottom-up connections between the lower level (phonemic) and higher (lexical), while others posit bidirectional information flow." The fact that bidirectional information flow really is the issue is confirmed in an earlier paper by McClelland (1991, p. 3), which makes it clear that the debate is whether "perception involves a bidirectional flow of information," a point endorsed by Massaro and Cohen (1991) who cite the same quotation from McClelland. It is not surprising that our own papers contain many similar quotations (e.g., Cutler et al. 1987; McQueen 1991; McQueen et al. 1999a; Norris 1992).

Why does **Samuel** now think that Merge is interactive and nonautonomous? Given that he has adopted the standard conventions in the past, it is hard to know why he adopts different interpretations here. Part of Samuels problem may be attributable to the fact that he wrongly equates phoneme and decision nodes in Merge with the feature and phoneme nodes of TRACE. In TRACE, features feed into phonemes, which in turn feed into words. In Merge only the input phonemes feed into words. Decision units cannot be equated with phoneme nodes in TRACE as they do not feed into lexical units. But Samuel has chosen to call the connections from lexical to decision nodes "top-down." He then states that "Norris et al. offer several reasons for including such top-down connections, and they are exactly correct: Top-down lexical influences are necessary." It is important that decision nodes are influenced by the lexicon, but this influence does not involve top-down flow of information. Information in these connections passes from input to output.

Possibly **Samuel** believes that any information flow from lexical to phonemic representations is "top-down": "if lexical nodes influence the activation of phonemic codes, a model is *not* autonomous." Note that the effect of this would be to redefine "top-down" so that any demonstration of lexical effects on phoneme identification (which must surely be based on phonemic codes) is "top-down." All researchers in the field have been in agreement about the existence of lexical effects on phoneme identification for

more than 20 years (see Cutler & Norris 1979 for review). Furthermore, lexical nodes have always influenced phonemic codes in bottom-up models. In the Race mode, lexical access makes the lexically based phonological code of the word available. In Samuel's terms, lexical nodes activate phonemic codes. If we were to adopt the terminology of Samuels commentary everybody would accept that the data argue for "top-down" processing and all of the models would be "top-down" too. Have all of us who have worked on this question (including Samuel) been wasting our time? No. We have all been addressing the much more interesting question of whether there is top-down feedback. Furthermore, Samuel himself has made some rather ingenious contributions to this debate (e.g., Samuel 1997). We only hope that his terminological *volte face* is just a temporary aberration and not an attempt to rewrite history and pretend that he believed in what we are now proposing all along. It is not that we do not want him to agree wiyh us. But we think he should agree on our terms.

Appelbaum suggests that we have reinterpreted the interactive/autonomy distinction. But in fact it is Appelbaum who seems to have interpreted the distinction incorrectly. In an earlier paper, Appelbaum (1998) assumed that lexical effects on phonemic processing (e.g., Ganong 1980) were evidence of "top-down information flow" (Appelbaum 1998, p. 321) and hence evidence against a modular stage of phonetic perception. The Race model (Cuder & Norris 1979) had long ago shown that lexical effects are entirely consistent with a modular stage of phonetic perception, and Merge maintains modular prelexical processes. Remember, decision nodes are not prelexical. Appelbaum's attempts to use lexical effects on phoneme decisions as evidence against modularity are therefore flawed; and her criticisms of our terminology may stem from a misreading of die literature.

Appelbaum, Pitt, and **Samuel** also seem confused by our application of the term "autonomous" to Merge. As we pointed out, autonomy is properly applied to stages rather than models, and Merge "preserves the essential feature of autonomous models - independence of prelexical processing from direct higher-level influence" (sect. 5.1, para. 7). Prelexical processing, and that is what the debate is about, is autonomous in Merge. The appropriate classification of the decision units is less straightforward. The decision units are flexible and configurable according to task demands, so they certainly do not constitute a Fodorian (Fodor 1983) module. Once configured for the task, however, they take input from two sources (lexical and prelexical) and then produce an output without interference or feedback from subsequent processes. This justifies the label "autonomous."

Finally in this section we should respond to the claim of **Tanenhaus et al.** that there is feedback from lexical to decision nodes. Where there is no feedforward (decision to lexical) there can be no feedback. The lexical-to-decision connections are feedforward.

R2.3. Parsimony. The question of parsimony rests largely on the issue of whether the decision nodes in Merge are an added extra that interactive models can do without (see **Doeleman et al., Gow, Murray, Pitt, Slowiaczek,** and **Whalen**). For example, there are no explicit decision nodes in TRACE so, although TRACE has interaction, it has no counterpart of Merge s decision nodes. How then can we claim that Merge is simpler than TRACE? There are two

parts to our answer. As we explained in the target article, one is that even if TRACE is as simple as Merge, it cannot account for the data (e.g., Pitt & McQueen 1998). We will remind readers of the details of this argument when discussing comparisons between Merge and TRACE in a later section. The second is that all models need some form of decision mechanism. Merge only appears more complex because it makes that mechanism explicit.

R2.3.1. Decision processes in Merge. Most psychological theories give a less than complete account of how a model might be configured to perform various experimental tasks. For example, TRACE and Merge must be able to perform either lexical decision or phoneme identification depending on the requirements of the task. In early phoneme-monitoring studies, listeners were typically required to monitor only for word-initial phonemes (Foss 1969). By definition, this demands that positional information from the lexicon is combined with information about phoneme identity. Neither Race nor TRACE ever specified a mechanism for performing this part of the task. This is unsurprising because there is practically no limit to the complexity of the experimental tasks we might ask our subjects to perform. Listeners could no doubt be trained to monitor for word-initial phonemes in animal words when a signal light turned green. Correct responding would require combining phonemic, lexical, semantic, and cross-modal information. But this does not mean that we have hard-wired {initial /p/, animal, green} nodes just sitting there in case someone dreams up precisely such an experiment. It certainly does not mean that we should conclude that the processes of colour perception, semantic processing, and phoneme perception all interact in normal speech recognition. A far more likely explanation is that a number of simple non-interacting processes deliver output to a system that can monitor and merge those outputs to produce a response. This system has to have enough flexibility to cope with all manner of bizarre tasks that experimenters, and die world in general, can throw at it. In Merge we have finessed the issue of how this system configures itself, and assumed that we can represent the process of combining different sources of information by a set *of* decision nodes. Merge does one extra thing. Although we can devise phoneme identification tasks that necessarily take account of lexical information, in die simplest phoneme identification tasks listeners could, in principle, ignore the output of the lexicon (and in fact often appear to do so; Cutler et al. 1987). In Merge we assume that listeners sometimes monitor the phonemic and lexical levels even when this is not explicitly required by die task, and that this is the source of lexical effects in phoneme identification.

Additional evidence that we need something more than just the phoneme nodes of TRACE to perform phoneme identification was reviewed in section 7 of the target article. The ability to perform phoneme identification is not an automatic consequence of being able to recognise spoken words. For instance, it is greatly facilitated by having learned to read an alphabetic script (Read et al. 1986). Furthermore, neuroimaging work reveal different patterns of brain activity in tasks involving explicit phonological decisions from those involving passive listening to speech (Demonet et al. 1994; Zatorre et al. 1992; see Norris & Wise, 1999, for review).

In conclusion then, the decision nodes in Merge do not

undermine its parsimony compared to other models. All models must make allowance for the facts that diere is a flexible and configurable decision mechanism, that listeners have to learn to interpret the workings of prelexical processes, and that explicit phonological decisions appear to activate parts of the brain not activated during normal speech recognition. The important point is that the decision process is not an optional extra. Without some such process listeners could not perform the experimental tasks we give them. The decision process is not something Merge has but other models can do without. All models need a decision process. Our claim is that when that decision process is taken into account we see that it is probably responsible for lexical effects in phoneme identification, leaving normal speech perception as a feedforward process.

R2.3.2. Rewiring decision nodes. The decision process has to be very flexible. Our suggestion that the connections in Merge might be rewired on the fly is the subject of criticism by both **Grainger** and **Grossberg.** Grossberg's worry about the plausibility of "rewiring" seems to apply to the very literal rewiring that might be done by a neural electrician. Our intention is to capture the process of reconfiguring network connectivity as in the Programmable Blackboard model of McClelland (1986). As we have argued above, all models must be able to configure themselves according to task demands. Grossberg's ART model must find a way of doing this too.

Grainger suggests that rewiring is implausible and unimplementable. The original suggestion for wiring on the fly, as proposed for Merge and Shortlist, rests on the assumption that it is worth adding an extra mechanism in order to save the need to have vast (possibly astronomical) numbers of permanent connections. The issue of rewiring is quite orthogonal to the question of feedback. However, it should be clear that if two different representations (say lexical and decision) are to be wired dynamically, then there must be some way to identify pairs of representations that are to be wired together. Lexical representations should therefore not be considered to be single unstructured nodes. They must contain the form-based lexical representation which can be dynamically associated with an appropriate decision node. It has always been part of Shortlist and the Race model that the lexicon explicitly represents phonological form. Grainger's assumption that a dynamically rewirable version of Merge would have no lexical representation of phonological form is bizarre.

Note that if we set aside the issue of rewiring on the fly, Merge simply does not have the problems **Grainger** supposes. In the simulations we presented, the decision nodes are effectively the output representations. Activating a word activates its phonological form on the decision nodes.

For some reason, **Grainger** believes that the problem of merging lexical and phonemic information presents a problem for Merge which is not faced by his own DROM model (Grainger & Jacobs 1994) simply because the DROM can combine letter and spelling information "at leisure." The speed of the process does not alter the logic of the connectivity. It is fortunate that Merge does not have this problem as DROM would have exactly the same problem.

R2.3.3. Feedback consistency. A further issue of parsimony is raised by **Ziegler & Van Orden** who believe that models with feedback have been able to generate important

theoretical predictions such as the feedback consistency effect in reading which "would never have been predicted by exclusively feedforward models." Interesting to note, Norris (submitted) demonstrates that the reported data on feedback consistency effects in reading can be well explained by the feedforward multiple-levels model (Norris 1994a) without any modification whatsoever. The reason that a feedforward model can simulate a "feedback consistency" effect is that the effect is not actually due to feedback at all, but to the type frequency of body-rime correspondences. Other things being equal we might expect most rimes to appear in a roughly equal number of words. If those rimes are always spelled in the same way, then the type frequency of each body-rime correspondence will be roughly equal. But, for rimes that are feedback inconsistent (i.e., spelled in more than one way), the major body-rime correspondence will tend to have a lower type frequency than in feedforward consistent words. Feedback consistency has an effect on naming because it tends to alter the type frequency of the correspondence. Feedforward models like the multiple-levels model are sensitive to type frequency. Feedforward models predict the data and correctly explain it as an effect of type frequency which has nothing to do with feedback from phonological to orthographic processing.

R3. Comparison of Merge with other models

R3.1. **Merge versus TRACE.** Throughout the target article, we claim that Merge is more parsimonious than interactive models. It is quite possible that Merge could be theoretically sound, but actually less parsimonious than interactive models. If models with and without feedback were otherwise equal, and the trade-off were simply between having the phoneme decision units units required by Merge and having feedback, it is hard to see which would be more parsimonious. This is essentially the point raised by **Murray, Pitt,** and **Tanenhaus et al.** How do we set about choosing between similar models? As we pointed out in section R2..3.1 above, all models need some form of decision process. Merge incorporates that explicitly, other models do not. So, comparing Merge and TRACE for parsimony is not actually comparing like with like. TRACE has extra hidden complexity, even though it may have fewer free parameters (Pitt). But most importantly, Merge still satisfies Occam's precept better than TRACE does. Occam's razor favours the most parsimonious theory consistent with the data; TRACE (the original or our modified version) is inconsistent with the data from Pitt and McQueen (1998). TRACE is also inconsistent with the development of phonological awareness with literacy without adding something akin to decision units; and finally TRACE is unable to account for detection of mispronunciations. We did our best to show that TRACE could be modified to account for the subcategorical mismatch data, but that is not enough.

In discussing the Merge simulations, **Tanenhaus et al.** state that we would like to conclude that the superior performance of Merge over the interactive model simulation is "compelling evidence for Merge and against TRACE." This is incorrect. As we point out (sect. 6.1), "With Merge-like dynamics, an interactive model could approximate the correct data pattern." The importance of the simulations is to demonstrate that a feedforward model can account for the subcategorical mismatch data and to show how a model like

TRACE might be modified to simulate that data too. The compelling evidence against TRACE comes from the data from Pitt and McQueen and the fact that TRACE fails to account for the bigger picture.

Tanenhaus et al. believe that we have made "questionable linking assumptions between the data and the models" (without saying why they believe this), and they seem to take exception to our assumption that positive lexical decision responses should be made when any lexical node exceeds a threshold. Note that we make exactly the same assumptions about response thresholds for both Merge and the interactive model. There is convincing neurophysiological evidence that reaction times are determined by neural activation thresholds in eye movement control (Hanes & Schall 1996). Both **Tanenhaus et al.** and **Caskell** remark that the threshold in Merge needs to be precisely set to simulate the correct pattern of data. This is true, but follows from the need to match the models performance to that of subjects. Subjects in these experiments make few errors. To respond correctly, they must place their decision criterion high enough not to be exceeded by nonword activation and low enough to always be exceeded by word activation. In Merge this requirement ties the criterion down to a range of 0.06 activation units and in the interactive model about 0.1 units. In both models a high criterion within this range leads to equally fast responses to N3W1 and W2W1, whereas the lowest possible criterion would lead to slightiy faster N3W1 responses. With the lowest criterion, the N3W1 advantage is twice as large for the interactive model as for Merge. Contrary to what Tanenhaus et al. claim, we would not expect to see evidence of fast lexical decision responses based on early activation of W2 if subjects are responding accurately. Also, contrary to their claims, the RT distributions of our data are clearly unimodal and not bimodal. Because W2W1 and W2N1 follow the same trajectory until cycle 8 there is no way that subjects could possibly make fast "Yes" responses to W2W1 based on early W2 activation without also making erroneous "Yes" responses to W2N1. This is not surprising because the final phoneme is not fully presented until cycle 9. Note that the error rate to W2N1 items is only 3%.

R3.2. Merge versus FLMP: FLMP is running a different race. In terms of the central argument about feedforward processing there is no fundamental conflict between Merge and FLMP. But **Massaro's** and **Oden's** commentaries now make us think that, in processing terms, FLMP must be much more different from Merge than we had originally thought.

Both **Oden** and **Massaro** criticise us for having misrepresented FLMP when discussing their account of the Ganong effect, where we say that "the support for a word has nothing to do with the perceptual evidence for that word" (sect. 6.3, para. 6). Oden points out that when they say "support for the voiced alternative given by the following context" (Massaro & Oden 1995, p. 1054) they are not saying that *gift* supports /g/, but that *ift* supports /g/. The evidence for *ift* is independent of the evidence for /g/ whereas the evidence for *gift* would not be. But why is the probability of responding /g/ dependent on the evidence for *ift?* The sequence *ift* does not support /g/ any more than it supports any other phoneme. The word *gift* might support /g/, but there is simply no reason why the sequence *ift* should support any onset phoneme in the absence of infor-

mation about lexical forms. Oden's claim that *ift* supports /g/ only makes sense if the relevant information is derived from the lexicon. So, when making a phonetic decision, the listener must decompose the input into the phoneme of interest and the residue. Then the residue *-ift* can be fed into the lexicon to determine that /g/ is a possible word onset in this context and /k/ is not. This way the context is independent of the support that /g/ provides for the word *gift*. Of course, at the same time *gift* is also being fed into the lexicon so that the word can be recognised. All of this is to avoid violating independence by feeding only *gift* into the lexicon and allowing lexical information to bias interpretation of /g/. Perhaps Massaro and Oden will think that our attempt to discover die processes behind the FLMP equations has led us to misrepresent them again. But in fact this is the heart of our criticism of FLMP. Although the FLMP equations are simple, they do not specify a process model, and it is far from clear what the underlying process model should be (for similar criticisms see Grossberg et al. 1997). Also, within the broad scope of the FLMP equations, there seems to be just too much room for manoeuver in how they are used to explain any particular piece of data.

This flexibility is apparent in **Oden's** commentary. In response to our criticism of the FLMP account of compensation for coarticulation, Oden offers a new explanation of sequential effects in FLMP terms of decisions about the "candidate identity of the sequence of words." The implication of this statement is that compensation for coarticulation takes place not at a prelexical, or even a lexical level, but at a new level representing sequences of words. The one straightforward thing we can say about this explanation is that it is wrong. As we will show later in section R4.3, there is abundant evidence that compensation for coarticulation is prelexical. Compensation applies even to nonword stimuli. Therefore, as we originally argued, FLMP still has no plausible account of the Pitt and McQueen data.

Oden suggests that the inhibition in Merge might produce all-or-none decisions. This tends not to be true given the levels of inhibition employed at the decision stage. As we pointed out, adding noise would also stop the model being deterministic. However, there is no doubt that there is work to be done in developing the model to account for both response probability with ambiguous input and speed of responding with unambiguous input (see Carpenter 1999; Ratcliff et al. 1999; Usher & McClelland 1995).

Both **Massaro** and **Meyer & Levelt** criticise us for concentrating too much on modeling activation levels. Massaro assumes we believe that activations are somehow better than response probabilities; Meyer & Levelt suggest that it is preferable to use the Luce choice rule than to allow inhibitory effects on activation. However, the Luce rule is not simply an alternative to inhibition, because models still need a mechanism whereby the rule can be implemented. Any complete model needs an account of how response probabilities are computed. Network models have the ability to suggest mechanisms which show how differences in activation can be translated into differences in response probabilities and latencies (Carpenter 1999; Page 2000).

Massaro criticises models with hidden units as being untestable. Contrary to his claim, however, there is no connection between the fact that networks with hidden units can approximate any measurable function (Hornik et al. 1989) and their testability. Nothing in this work implies that a network trained on a given data set (such as speech input)

will then behave as people do (in some experimental task on which it was not explicitly trained). A clear example of this comes from Seidenberg and McClelland's (1989) model of reading aloud. The model is trained to translate orthography to phonology. The model succeeds or fails (it actually does a bit of both) depending on its ability to simulate human reading behaviour, something it was never explicitly trained to do. A model trained directly to reproduce the correct RTs and error rates might not be testable, but then it would just be a redescription of the data. Massaro s criticism of models with hidden units is fallacious, and so, in consequence, is his attempt to extend such criticism to network models in general.

R3.3. Merge versus the distributed cohort model. The commentary by **Gaskell** shows that the Distributed Cohort Model (DCM, referred to in the target article as the postlexical model) can be modified to overcome the technical criticisms we made in the target article and to simulate the subcategorical mismatch data. This suggests that the subcategorical mismatch data might not be as diagnostic as we originally thought, and that at least one other bottom-up model can account for the data. Although the model still cannot explain the variability in the effect, for exactly the reasons we originally suggested, **Marslen-Wilson** suggests that the DCM probably needs a decision mechanism that can shift attention from the normal phonological output to a lower-level auditory output less influenced by lexical factors. This is essentially the same explanation as in Merge where attention can be shifted between levels. However, the recurrent net architecture still fails as a model of continuous speech recognition for the reasons pointed out by Norris (1994b). Other shortcomings of recurrent networks in speech recognition are highlighted by Nearey. Page (2000) presents a more general critique of models relying on distributed representations. One problem that DCM faces is that it is not clear how lexical decisions could be made. Presentation of an input word leads to a pattern of activation across semantic units. Without some independent lexical representation that specifies exactly what pattern of semantic unit activation is to be expected for each word, there is no way to determine whether a given activation pattern actually corresponds to a word or not.

R3.4. Merge and ART. The following line of reasoning is pursued by **Montant:** ART uses feedback, ART is good, therefore this is evidence in favour of feedback. Remember that our central claim is that "Feedback is never necessary." We also pointed out that the best a recognition system can do is to select the stored representation that best matches its input. This holds for ART as much as anything else. In ART, the feedback is part of the mechanism for assessing the degree of match between bottom-up input and stored representations. The same result could be achieved without feedback. Indeed, although most versions of ART use feedback, the feedback is not necessary and is not needed in ART2-A (Carpenter et al. 1991). Grossberg et al. (1997a) demonstrate the similarity between ART and the FLMP equations which do not require feedback to be implemented. Feedback is also absent from the related learning mechanisms proposed by Page (2000). So the fact that ART has such an impressive track record, and normally uses feedback, in no way counters our thesis about feedback not being necessary.

A potentially more interesting criticism based on ART comes from **Luce et al.** They also think that feedback is needed for ART and is therefore a good idea. But they argue that in ART phonemes and words are just lists of different lengths, so the whole issue of feedback between words and phonemes simply does not arise. Although it is true that phonemes are represented at the list level in ART, they are also represented at a lower level as the elements from which lists are composed. We can see this clear distinction between levels in action in Grossberg's (e.g., Cohen & Grossberg 1986) account of the word superiority effect in reading, which relies on feedback from the list (i.e., letter and word) level to the previous letter level. We presume that the account of lexical effects on phoneme identification would have a similar explanation in ART, in that phonemes could be identified directly from a phoneme level. The alternative that Luce et al. suggest is that attention can be shifted between words and phonemes by attending to different sizes of list. Normally longer list units like words would mask (inhibit) shorter units like phonemes. Such masking would be stronger for phonemes in words than in nonwords. So, while attention to lists of length 1 might overcome the problems faced by phonemes in words, there is no reason why it should lead to facilitation of responses to words. If the crucial representations for phonemes and words are both at the list level, then the model cannot explain the effects of lexical context on phoneme identification.

Overall, our view of ART is that it may well be able to provide the basic building blocks for implementing a model of speech perception. It has addressed many issues which have been ignored by other models. However, the basic principles of ART place few constraints on how the components might be put together to form a complete model, and it is not clear that feedback would necessarily be central to such a model. ART is now being used to simulate real data on speech perception, and we look forward to an ART-based testable psychological model of speech perception.

Grossberg himself argues that feedforward models fail to explain phenomena such as phoneme restoration and backward effects in time (Repp 1980; Repp et al. 1978). First, the argument concerning phoneme restoration is flawed because it depends on the assumption that the source of the restored phonemic percept is in the input representation rather than being derived from the lexical representation. Second, the existence of backward effects in time has nothing to do with the feedforward/feedback distinction. Shortlist, a strictly feedforward model, simulates a range of backward-in-time phenomena (e.g., Norris et al. 1995; 1997).

R4. Data

Thirty years ago Foss (1969) introduced the phoneme-monitoring task to Psycholinguistics (presumably precipitating **Warren** s musings over phoneme detection during a colloquium that year). We, like many users of the task, would not want to claim that it taps directly into necessary stages of speech processing. Indeed, this is one of the motivating factors for our development of Merge. However, we do believe that in the past three decades spoken-word perception has been an enormously active research field in which real theoretical progress has been made, and that this

is in part due to Foss and the other pioneers who equipped Psycholinguistics with the necessary empirical tasks. That these tasks often involved metalinguistic decisions is a consequence of our inability to measure perception directly; **Doeleman et al., Marslen-Wilson, Meyer & Levelt,** and **Murray** all remark on the undesirability of this situation. Meyer & Levelt further claim that the study of speech production is less bedevilled by the indirect measurement problem than the study of perception, because in their work on production they are able to measure (and model) onset of articulation. We suspect that this claim should be taken with a pinch of salt; articulation can be seen as the bottleneck of an otherwise far more rapid speech production process (Levinson 2000), and *this* allows for the possibility that production processes such as lexical access are not directly reflected in articulation speed at all. For perception, however, both Murray and Meyer & Levelt point to the usefulness of recently developed eye movement paradigms (Tanenhaus et al. 1995). So far these tasks can only be used with a restricted set of specified response options (a display of which the subject knows all the members in advance), which means that many issues are as yet outside their range of usefulness; we certainly hope that this type of approach will be incorporated in further tasks of greater refinement in the near future. Even better, of course, would be appropriate refinement of brain imaging techniques; these are still laughably far from being able to provide insight into the sort of questions dealt with in the experiments we have discussed (such as the processing difference involved in hearing two versions *of job* in which the jo- portion comes respectively *from jod* or from jog).

At the moment, however, the data on phonemic decision making provide the only insight into such questions, and none of the commentaries lead us to revise our conclusion that the Merge model currently provides the best available account of these data. In this section, we discuss the comments addressed to specific questions about the decision data. The presentation follows the order we adopted for our review of the data in the target article (sect. 4), but ends with a new subsection on acoustic-phonetic processing.

R4.1. **Variability of lexical effects.** The variability of lexical effects on phonetic categorization and phoneme monitoring tasks is a challenge to models with feedback. No commentator contests this claim. **Pitt,** however, draws attention to a specific kind of variability of lexical involvement in phonetic categorization which we did not discuss in the target article. This is that lexical effects in categorization change overtime (Fox 1984; McQueen 1991; Pitt & Samuel 1993). Pitt questions whether Merge could deal with this variability. It can. Lexical involvement in the model tends to build up, and then decay over time (although the situation that is being modelled is somewhat different, the lexical effects in Merge's subcategorical mismatch simulations [see Fig. 3b] increase as lexical activation builds up, and then decrease as phoneme node activation reaches asymptote). It is possible that experiments on word-initial categorization have tended to tap into the incrementing phase of lexical involvement (the standard finding is that there are larger lexical effects in slower responses), while those on word-final categorization (here there are smaller lexical effects in slower responses) have tended to tap into the decrementing phase. We have already begun to address this issue experimentally (McQueen et al. 1999b).

It is important to note that the pattern of lexical involvement in word-final categorization, though not problematic for Merge, is in fact problematic for models with feedback, like TRACE (as McQueen 1991 argued). TRACE simulations in McClelland (1987, Fig. 1.2, p. 12) show that, as processing continues, lexical feedback acts to increase the difference in activation between the phoneme nodes for the lexically-consistent and lexically-inconsistent phonemes (/t/ and /d/ given *dor?* in McClelland's example). TRACE therefore wrongly predicts that lexical involvement in word-final categorization should build up over time. We thank **Pitt** for reminding us about another of TRACE'S frailties.

R4.2. Facilitation versus inhibition in phoneme monitoring.
The results presented by **Connine & LoCasto** show that listeners were faster to detect the target /m/ in the nonword *chorum* (which is close to the word *chorus) than* in the control nonword *golum* (which is not close to any real word). This finding replicates Wurm and Samuel (1997) and supports Wurm and Samuel's argument that more word-like nonwords are easier to process than less word-like nonwords. This is one reason why Frauenfelder et al. (1990) may have failed to find inhibitory lexical effects in nonwords like *vocabulaire*. Despite **Samuel's** protestations, however, it remains the case that there was no *direct* evidence for fhis kind of inhibitory effect when the target article was written. Why Connine and LoCasto's commentary is important is that it now provides us with direct evidence of lexical inhibition. Listeners were slower to detect the ///, for example, in the nonword *chorush* than in the control nonword *golush*. It would appear that when die target phoneme is close enough to the sound it replaces in the base word (/// is only one feature different from the /s/ in *chorus) there* is sufficient support for the lexically-consistent sound (/s/) to overcome the benefit due to *chorush* being more word-like than *golush,* resulting in a small net inhibitory effect.

Connine & LoCasto claim that their results are inconsistent with the Merge model. Specifically, they suggest that the bottom-up priority rule in Merge might have to be abandoned. They are right to suspect that we would be loath to remove this rule; it serves the important function of preventing hallucinations. These new results are, however, consistent with Merge and in fact provide support for the rule. The inhibitory effect appears to occur only when the target phoneme is phonetically close to the phoneme it replaces, that is, when the target itself provides some bottom-up support for the replaced sound. Since /f/ activates the /s/ decision node, the word node for *chorus,* following the bottom-up priority rule, can also activate the /s/ decision node. Due to the resulting competition between the */s/* and */f/* nodes, /f/ decisions will be delayed. When the target provides no support for the sound it replaces *{/ml* in *chorum* differs in place, manner and voicing from /s/) the bottom-up priority rule will prevent lexical activation from supporting the lexically-consistent phoneme, and no inhibition will be observed.

We agree with **Connine & LoCasto** that attentional processes have an important role to play in language processing. No model of phonemic decision making has a satisfactory attentional component. Merge, like any other model, would be strengthened if it included a fuller account of attentional factors.

R4.3. Compensation for coarticulation.
The results of Pitt and McQueen (1998) are particularly important in the feedback debate. They found a dissociation between a lexical effect on the labeling of word-final fricatives and no lexical effect on the labeling of following word-initial stops (e.g., categorization of the ambiguous fricative in "jui? ?apes" as *Is/,* but no increased tendency to label the following stop as /k/, consistent with compensation for coarticulation following *Is/).* This dissociation is very problematic for models with feedback, like TRACE. If feedback modified the activation of the *Is/* node at the phoneme level in TRACE, the compensation for coarticulation mechanism at that level of processing ought to have been triggered. The results are however consistent with the Merge model, in which the lexicon can influence fricative decision nodes, but cannot influence the prelexical compensation mechanism.

Some commentators question this argument. **Pitt** points out that the compensation process may not be purely prelexical, while, as we have already discussed, **Massaro** and **Oden** wish to maintain their view that the process operates at a high-level integration stage in FLMP. The evidence, however, suggests strongly fhat compensation for coarticulation has a prelexical locus. Pitt and McQueen's (1998) data in fact suggest this: It would be hard to explain the dissociation in lexical involvement between the fricative and stop decisions if compensation did take place at the decision stage. Mann and Repps (1981) original demonstration of fricative-stop compensation was based on nonsense strings (like /ska/ and /ujta/), suggesting that the process does not depend on lexical access (contrary to Oden s suggestion). Most audiors therefore agree that fricative-stop compensation is prelexical (Elman & McClelland 1988; Pitt & McQueen 1998). **Brancazio & Fowler** argue that liquid-stop compensation is also owing to a prelexical process.

A particularly striking demonstration that liquid-stop compensation does not operate at the phoneme decision stage is provided by Mann (1986b). Japanese listeners who could not identify English /l/ and /r/ correcdy showed appropriate compensation in their labeling of stops following /]/ and /r/ (i.e., more /go/ responses after */all than* after */ar/).* These subjects showed fhe same amount of compensation as both native English speakers and Japanese listeners who were able to identify l\l and /r/. The process responsible for compensation for coarticulation between liquids and stops (and, by extension, probably the mechanism for fricative-stop coarticulation) therefore appears *to* operate at fhe prelexical stage, that is, at a level of processing below fhat at which explicit phoneme decisions are made.

Pitt and McQueen (1998) also showed fhat compensation for coarticulation following ambiguous fricatives could be triggered by Transitional Probability (TP) biases in die nonword contexts in which the ambiguous fricatives were placed. Previous demonstrations of lexical involvement in compensation for coarticulation in which the word contexts had TP biases (Elman & McClelland 1988) could thus be due to a prelexical process sensitive to TPs (and not to lexical feedback). The remarks of **Brancazio & Fowler, Doeleman et al.,** and **Massaro** suggest that they may have misunderstood Pitt and McQueen's results. It is therefore important to emphasize that Pitt and McQueen did not show fhat the compensatory effect was owing to a TP bias rather than to a bias based on sensitivity to coarticulation.

That is, they did not show that there was a TP mechanism instead of a compensation for coarticulation mechanism. Rather, they showed that the process which compensates for coarticulation could be triggered by a TP bias in the context. In other words, the compensation process can be activated either by unambiguous fricatives or by an ambiguous fricative in a context where, for example, TPs favor *Is/* over

Vroomen & de Gelder question the distinction between TPs and lexical information. They ask why, in the Merge account, statistical regularities can play a role in prelexical processing while lexical knowledge cannot. Our answer is that the data, in particular the dissociation between the effects of lexical biases and TP biases in Pitt and McQueen (1998), but also other similar dissociations (Vitevitch & Luce 1998; 1999), suggest that the effects result from two distinct sources of information, and that TP information is stored prelexically. Vroomen and de Gelder also question whether TP information can assist prelexical processing which is already very robust. As we discuss in section R4.7, we agree that prelexical processing is very efficient. Pitt and McQueen (1998) therefore suggest that TPs will be of most value when the speech signal is ambiguous (even if that is a relatively rare phenomenon outside the psycholinguistic laboratory). Vroomen & de Gelders suggestion that TPs could only be learned using some form of feedback is incorrect. TPs are statistical regularities in the speech signal, and thus can be learned from the signal by a feedforward system.

R4.4. Phonemic restoration and selective adaptation. In
his commentary, **Samuel** expresses concern that we are unwilling to accept the data in Samuel (1997) as evidence of feedback. He attributes two misconstruals to us. First, he tries to undermine our point that the adaptation produced by restored phoneme looks different from that obtained with real phonemes. There is no need to tally up just how many adaptation effects reported in the literature are limited to the category boundary and how many are spread over the entire continuum; the point remains that the effects with real and restored phonemes in Samuel (1997; see Figs. 1 and 2, pp. 102 and 104) do not look the same. This worried us, and still does.

Second, **Samuel** suggests that we have distorted the results of Samuel (1997) and of Samuel and Kat (1996), claiming that we suggested that the adaptation occurs at the lexical level. We did not. We agree with Samuel and Kat (and others) that adaptation may well operate at several different levels of processing, but we did not propose that the lexical level is one of them. The crucial issue is the locus of the adaptation effect with restored (noise-replaced) phonemes. We suggested that Merge could account for Samuels (1997) data if the locus of the adaptation effect with restored phonemes is found to be at the decision stage (which can indeed be equated with Samuel and Kat s "categorical" level; in both accounts, these levels are responsible for categorical decisions). We also argued for a type of bottom-up priority in selective adaptation, that is, that adaptation effects are driven primarily by the information in the speech signal, rather than by phonemic precepts. The failure to find lexical effects with intact adaptors (Samuel 1997, Experiment 3) is thus consistent with the proposed account in the Merge model. Lexical context may bias processing at the decision level with noise-replaced adaptors but not with

intact adaptors for two reasons: because diere is no ambiguity at the decision level which lexical context can act upon when the adaptors are intact; and because intact adaptors will produce adaptation primarily at lower levels, which (at least in the Merge model) can not be modulated by the lexicon.

In short, there is nothing in **Samuel's** commentary to change our view of the Samuel (1997) data. We agree that these are potentially crucial data in the feedback debate. However, the locus of the adaptation effect with noise-replaced adaptors remains to be established. Given the importance of these findings, we have attempted to replicate them, but have been unable to do so (McQueen et al. 1999c).

R4.5. Lexical effects on phonemic decisions in nonwords.
It is argued by **Newman** and **Brancazio & Fowler** that a prelexical mechanism sensitive to simple TPs cannot be responsible for the effects on phonetic categorization in nonwords reported by Newman et al. (1997). We agree that since simple (diphone) probabilities were controlled by Newman et al. they cannot be the source of the effect. Higher-order (longer range) probabilities may have played a role, however. Newman in fact suggests that the probabilities between the initial and final consonants in her materials may have had an effect (though Brancazio & Fowler argue that these probabilities were also controlled). But what about the probabilities ofthe complete strings? Brancazio and Fowler assume that these were all zero. The CVCs only have zero probability on a strictly syllabic account, however. Though all of Newman et al.'s items were nonwords, and none appear as syllables in English words, some of the sequences do nevertheless occur in English words (e.g., *beysh* appears *mprobation, kice* in *skyscraper*). In a count ofthe CELEX database (Baayen et al. 1993), we found that in one of Newman et al.'s sets *(beysh-peysh/beyth-peyth)* the triphone probability biases made the same (correct) predictions as die lexical neighborhood biases, while in another set *(gice-kice/gipe-kipe)* the triphone probabilities made the opposite (i.e., incorrect) predictions to the lexical neighborhoods. In die other four sets in Newman et al. (two which showed neighborhood effects, and two which showed no effects), we found no matching triphones (except for *toish* in *toyshop*).

Note that a dictionary-based count is a blunt instrument that can only approximate the frequencies of triphones (or diphones, or whatever) in continuous speech. The CELEX count misses strings across word boundaries in running speech, like *gice* in "big icecream," which at least theoretically might modulate prelexical TP sensitivities. The CELEX analyses nevertheless suggest that although some ofthe effects reported in Newman et al. are almost certainly due to the effects of lexical neighborhoods, some may be due to a prelexical mechanism sensitive to higher-order sequential probabilities. As we suggested in the target article, more work needs to be done to tie down the locus or loci of these effects. We also need to know more about the nature ofthe TP mechanism. As Pitt and McQueen (1998) pointed out, we do not yet know what the limits of listeners' TP sensitivity are (whether diphone, triphone, or even longer sequences are involved; whether syllable structure constrains sensitivity or not; and so on).

Whether these effects prove to be entirely prelexical, entirely lexical, or a mixture of the two, they do not challenge

the Merge model. Though **Newman** agrees that Merge could explain effects driven by a prelexical mechanism, she questions whether the model could explain effects at the lexical level, arising from the joint influence of gangs of words. She is right that there are strict limits on the number of words activated at the lexical level in Merge (as in the Shortlist model, whose name in fact reflects this property). In Shortlist, the default maximum number of candidate words considered to begin at any particular segmental position is 30. As we pointed out in the target article, the number of words can be reduced considerably witliout impairing Shortlists performance; that is, correct recognition is still achieved even when only two words are allowed in the shortlist (Norris 1994b). We did not mean to imply however that the maximum (in Merge or Shortlist) should be as small as two. Indeed, other effects of competitor neighborhood size on word recognition (Norris et al. 1995; Vroomen & de Gelder 1995) have suggested that the shortlist maximum should be larger than two. Shortlist is able to simulate such data successfully with the maximum remaining at the default of 30 (Norris et al. 1995; 1997). Although Newman et al.'s data have not yet been simulated, we think that it is reasonable to assume that Merge, operating with a similar shortlist size, would capture effects due to gangs of lexical neighbours (the largest gang in Newman et al. 1997 had 14 members).

R4.6. Subcategorical mismatch. The possibility is raised by **Whalen** that the mismatches in the cross-spliced items in McQueen et al. (1999a) and in Marslen-Wilson and Warren (1994) had overt ambiguities, and thus tJiat listeners used nonphonetic means to resolve these ambiguities. The data speak against this possibility. Although it is true that trained phoneticians could, with careful listening, possibly detect the mismatches in the materials, we do not believe that the naive subjects used in McQueen et al.'s experiments were able to detect the mismatches, at least when they were presented with the full items. Once the listeners had heard each final stop, they were able to identify it rapidly and accurately. If the materials had been overtly ambiguous, one would not expect mean phonetic decision latencies and error rates on the order of 650 msec and 5% and mean lexical decision latencies and error rates of about 470 msec and 8% (McQueen et al. 1999a, Experiments 1 and 3, cross-spliced conditions). The gating experiment in McQueen et al. shows that listeners were sensitive to the information in the pre-splice portions of the words (as does the forced-choice vowel identification task). But only in the earlier gates did listeners tend to respond with words consistent with the pre-splice portions of the cross-spliced items (e.g., *shot* responses to the W2W1 word *sloop,* made from the [slo] from *shot* and the [p] from *shop).* Once listeners had heard the final stop, over 85% of their responses reflected the identity of the release burst (e.g., *sloop* responses). We therefore believe that the effects in these experiments reflect the operation of bottom-up speech processing, as modeled in Merge, rather than conscious ambiguity-resolution processes.

R4,7. Speech processing. Speech recognition is a difficult and complex process. Several of the commentators seem to have based assumptions of top-down feedback solely on intuitions that a complex process must necessarily be error-prone, and hence incapable of succeeding on its own with-

out reference to other processing levels. Thus we read that speech "is characterized by noise and variability" **(Tanenhaus et al.)** and that ambiguity in speech signals "makes it very unlikely that a pure bottom-up analysis can be efficient" **(Montant)** so that "feedback would be helpful" **(Tanenhaus et al.);** the system should not be designed to be error-free and optimal because it is not actually error-free and optimal **(Connine & LoCasto).** These commentators are psychologists, and their intuitions do not appear to be shared by those commentators who are phonetic scientists. The assumption of error-prone front-end processing can be contrasted with the explicit detail of speech processing laid out in die commentary by **Kingston,** in which we see a picture of massive redundancy producing a bottom-up information flow of such richness tJiat there is room for portions of it to be lost without damage to the end result. Other commentators who are phonetic scientists **(Benki, Nearey, Stevens, Whalen)** likewise display no such intuition-based assumptions about defective front-end processing warranting feedback: for Whalen, the claim that lexical feedback cannot in principle improve speech processing is "sound"; for Benki our arguments against feedback are "convincing" and lexical effects should better be viewed in terms of bias; Nearey points out that the human system even in high noise with arbitrary and unpredictable input does a remarkably good job, far better than any existing ASR system; Stevens accepts bottom-up processing alone for the same situation of words in isolation.

A comparable contrast between views can be seen in the remarks of **Warren, Nearey,** and **Slowiaczek** on the issue of phonemic representations in speech processing. As we pointed out in the target article (sect. 7), the framework we have proposed is compatible with a range of possible front-end implementations. Certainly the experimental evidence (including of course that of our own work on subcategorical mismatch) indicates that listeners process speech input continuously and not as a sequence of independent phonemes. The evidence from his own laboratory which Warren so amply cites is fully consistent with the consensus position. Warren interprets such evidence as indicating that phonemes have no role to play in human speech processing. Nearey, however, on the basis of his own work, argues for "phoneme-like units," and Slowiaczek makes a strong case for phonemic representations on the basis of evidence from phonological priming. Our own position is closer to that of the latter two commentators, but the crucial point here is that nothing in the Merge/Shortlist framework depends on whether or not phonemic representations intervene in speech recognition. Phonemic decisions are based on output from the decision nodes, which are separate from the direct input-to-lexicon processing paffi.

Stevens, taking up the issue of the input-to-lexicon path, describes a casual-speech multi-word utterance the recognition of which, he maintains, involves the kind of top-down processes which the target article argues against. However, the processes he describes do not involve feedback. He proposes acoustic processing that produces a pattern of features; these features in turn generate a cohort of potential word sequences. This is exactly the process of multiple activation of candidate word sequences embodied in Shortlist and indeed most current spoken-word recognition models. Stevens then proposes application of rule-based transformations of the activated word forms. Rules are, of course, by definition not lexically stored information. Application

of the rules will then "lead to a pattern that matches the pattern derived from the acoustic signal." This is exactly the bottom-up priority embodied in Merge and Shortlist. Feedback, in contrast, would allow the reverse - transformation of the pattern derived from the acoustic signal to match the lexical form. That is, where Stevens's rules allow the system to accept, for instance, a nasal as a possible instantiation of a voiced fricative, top-down feedback would result in the system altering its analysis of the input, and deciding that what had been heard was a voiced fricative and not a nasal at all. Stevens does not think that this happens, and nor do we: there is no feedback in speech recognition.

Finally, the speech scientist commentators point to some levels of complexity which we had not considered explicidy in the target article: **Whalen** describes non-sequential context effects requiring integration of acoustic information across time, **Nearey** discusses the need for temporal sensitivity in the front-end processor, and **Stevens** (as also the commentary by **Gow**) highlights the fact that phonological processes can transform featural representations of phonetic information. There are many further aspects still to the complexity of speech processing. But complexity is not ipso facto a warrant for feedback.

R5. The wider context of language processing

Several commentators relate our arguments to aspects of human language processing beyond die circumscribed domain of the evidence we reviewed. We used research on phonemic decision-making in speech recognition as a clear case study in which to examine the need for feedback in modeling the research evidence. But speech recognition is just one function of the human language processing system. This system not only recognises speech but also produces it; the relationship between our model and models of speech production has been raised by **Meyer & Levelt.** The system processes auditory information for speech recognition; but it is also capable of drawing on visual information to the same end, as noted by **Brancazio & Fowler.** The system recognises words; but it also recognises sentence structure, raised in the commentaries by **Isel** and **Shillcock.** Furthermore, the adult listener's recognition system has developed from an initial state via a process of language acquisition in the child, as **Jusczyk & Johnson** discuss; and it is implemented, as a number of commentators stress, in the neural architecture of the human brain. All these comments provide welcome views of the place of Merge in the wider context of language processing.

R5.1. Production and perception. It is proposed by **Meyer & Levelt** that certain representational levels in the language processing system are shared between production and perception, and that feedback must therefore necessarily occur between those levels. This speculation prompts two obvious remarks. One is that sharing of resources at these levels is as yet unsupported by empirical evidence. Experiments summarised by Levelt et al. (1999) support tripartite lexical processing in production (lexical concepts, syntactic words, phonological forms), but to our knowledge such a division is not indicated by empirical evidence for perception (although note that **Caskell** proposes a division between lexical content and form, implemented without

feedback in the DCM). The second remark is that bidirectional connectivity is the prerequisite for feedback, but is not itself feedback; feedback occurs when die connections are used in both directions during the same processing operation. If connections between two levels are used in only one direction during language production, and only in the other direction during language recognition, there is no feedback. Certainly there is room for further investigation of such issues.

R5.2. Syntactic processing. As Shilleoek points out, the recognition of function words is dependent upon syntactic context and hence might be more likely to involve feedback. Studies in Dutch (Haveman 1997) have in fact shown comparable priming effects for function and content words, and no evidence for the prediction of function words from syntactic context. **Isel,** responding to our remarks in section 3.5.2 about the relationship between syntactic and semantic processing during comprehension, describes ERP studies which indicate early independence of the processing of lexical gender and of sentence semantics. Modulation of the effects of gender by semantic factors occurs only at a later processing stage. Gender is a property of words which does not alter its type (masculine, feminine, and neuter, in the case of the study cited by Isel) as a function of syntactic structure, but can alter its expression; for example, to mark case relations. Gender type can thus hardly serve as the prototypical measure of syntactic processing; indeed, disambiguation via gender type has been shown not to block concurrent availability of alternate parses of an ambiguous syntactic structure (Brown et al., in press; Van Berkum et al. 1999). There is however separate electrophysiological evidence that syntactic analysis of verb agreement is independent of semantic processing (Hagoort & Brown, in press). Similar studies have shown separate processing effects of content and function words, at least in visual processing (Brown et al. 1999; Ter Keurs et al. 1999). As we pointed out in section 3.5.2, current models of syntactic/semantic processing differ significandy with respect to feedback; we welcome the growing attention paid to sorting out these differences via neurophysiological investigations.

R5.3. Audio-visual processing. We note that **Brancazio & Fowler** observe that Merge, like other models of speech processing, fails to incorporate any obvious mechanism for exploiting visual information. Visual information is, of course, not *necessary* for speech perception. Indeed the McGurk effect is evidence that speech perception can be adversely affected by visual information - it is only when looking at the face producing /ga/ tiiat we decide we are hearing /da/; close the eyes and the speaker's production of /ba/ is veridically available to the listeners consciousness. Although it is tempting to relegate this effect to domains external to the speech perception model, the phenomenon is nonetheless robust and poses an intriguing set of questions (which, it should be remarked, **Massaro** and his colleagues have not shied from addressing in FLMP simulations). Moreover, as Brancazio & Fowler point out, die range of data currently available suggest a prelexical locus for the McGurk effect, which could make it a useful experimental tool. We are therefore very interested to hear of Brancazio & Fowler's planned exploitation of audio-visual effects to test predictions from autonomous models such as Merge

versus feedback models, and we look forward to the results of their study. (Although space constraints prevented them from describing their planned materials in detail, we hope that as well as controlling transition probability of the consonant-consonant sequences they also, for the reasons discussed above in sect. R4.3, will be able to control the probability of the vowel-to-consonant transitions.)

R5.4. The development of phonemic processing. As **Jusczyk & Johnson** point out, any speech recognition system in place in an adult listener has its beginnings in a system developed by an infant. And an infant begins by knowing no words, so the system must be capable of developing without the use of information flowing from word representations to prelexical representations. This is of course not in itself an argument that the adult system must also make no use of top-down information flow. As Jusczyk & Johnson observe, a reorganisation of the system to allow feedback in the stable state is conceivable. They also observe that the decision nodes of Merge may imply reorganisation or elaboration of the system beyond what is available in the initial state, for phonemic decision is not, as we argued in section 7, a necessary operation in infant development. Neuro-imaging evidence certainly exists, which suggests that such a reorganisation distinguishes phonological processing by literate versus illiterate language users (Castro-Caldas et al. 1998), and evidence from aphasic listeners also suggests a dissociation of phonemic decision-making and speech comprehension (Basso et al. 1977; Riedel & Studdert-Kennedy 1985).

Note that **Jusczyk & Johnson's** assumption that phonemic decision plays no role in language development stands in marked contrast to **Doeleman et al.'s** claim that phonemic decision-making *is* part of infant perception. Here Doeleman et al. confuse infants' ability to discriminate with adults' ability to identify. Years of speech perception research have been based on the difference between identification tasks and discrimination tasks; discriminating a difference between two inputs is not at all the same thing as identifying the nature of the difference. In fact the studies with very young infants to which Doeleman et al. refer have inter alia shown infants to be capable of discriminations that adults cannot make; thus infants in an English-language environment discriminate a change from dental to retroflex stops, both of which English-speaking adults unhesitatingly categorise as a /t/ (neither dental nor retroflex, but alveolar place of articulation in their dialect; Werker & Tees 1984). That the discrimination performance is actually not based on phonemic identification was shown by Moon et al. (1992): in their study, infants could tell the difference between *pat* and *tap* but not between *pst* and *tsp*. The phonemic changes in Moon et al.'s two pairs were identical; in the first pair, however, the medial vowel resulted in a possible syllable, while in the second pair the medial fricative resulted in non-syllabic input which the infants clearly could not decompose as adult listeners would have done.

Not all infant perception research involves simple discrimination; researchers can now also establish whether infants prefer one of two types of input which they can discriminate. **Jusczyk & Johnson** list an impressive array of evidence gleaned from such preference tasks concerning the speech perception capacities of very young infants, and the list could be much longer. But Jusczyk & Johnson hold that these discrimination and preference capacities do not consti-

tute phonemic decision, and we agree. Phonemic decision is knowing, for instance, that cup and *cat* begin in the same way, and it is not observed, even in societies which encourage such awareness, till age three or four (Bradley & Bryant 1983; Liberman 1973). Phonemic decision-making is, as we argue in section 7 of the target article, separate from the normal route from input to lexicon, which by that age is fully in place.

R5.5. Neural implementation. A number of commentators (**Doeleman et al., Grossberg, Luce et al., Montant, Tanenhaus et al.**) raise the question of whether the existence of widespread neural backprojections in the brain might undermine our case against feedback. The answer here is that it depends on what dlose backprojections actually do. For example, backprojections might be involved in non-specific attentional control over the entire prelexical system. The presence of such backprojections would be entirely consistent with our case against feedback (see definitions). More generally, we have very little understanding of how information processing algorithms are implemented in neural hardware. Backprojections might well be part of the neural mechanisms, such as gain control, required to implement an informationally feedforward system with neural hardware. That is, the existence of backprojections may not be manifest at all at the psychological or information processing level. Alternatively, backprojections might be involved in learning but play no role in processing learned material (see Norris 1993).

The relation between processing models and their neural implementation is surely one of the most exciting areas for future research. But we should remember that the gulf between psychological models and their possible neural implementation is currently enormous.

R6. Conclusion

The feedback from the commentaries leaves us convinced that feedback in spoken word recognition is never necessary. There is still no good theoretical case for assuming that there should be feedback from lexical to prelexical processing in speech recognition. The data are consistent with a feedforward model like Merge, but inconsistent with a feedback model like TRACE. A model based on ART might possibly be able to explain some of the data, but it is far from clear that dle feedback in ART is necessary. Advances in neurobiology might well illuminate this debate but, as we have cautioned, the mapping between neurobiological data and psychological theory is not straightforward. In the meantime progress will come from the development of Merge and other models to give a better computational account of human speech recognition, one that can be subjected to rigorous empirical test.

References

Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively

Aderman, D. & Smith, E. E. (1971) Expectancy' as a determinant of functional units in perceptual recognition. *Cognitive Psychology* 2:117-29. [aDN]
Allen, J. (1994) How do humans process and recognize speech? *IEEE Transactions on Speech and Audio Processing* 2(4):567-77. [TMN]

Alonso, J. M,, Cudeiro, J., Perez, R,, Gonzalez, F. & Acuna, C. (1993) Influence of layer 5 of area 18 of the cat visual cortex on responses of cells in layer 5 of area 17 to stimuli of high velocity. *Experimental Brain Research* 93:363-66.　[MM]

Altmann. G. T. M. & Steedman, M. (1988) Interaction with context in human sentence processing. *Cognition* 30:191-238.　[aDN]

Appelbaum, I. (1998) Fodor, modularity and speech perception. *Philosophical Psychology* 11:317-30.　[rDN'l

Azuma, T. & van Orden, G. C. (1997) Why safe is better than fast: The relatedness of a words meaning affects lexical decision times. *Journal ofMemory and Language* 36:484-504.　[MGG]

Baaven, R. H., Piepenbrock, R. & van Rijn, H. (1993) *The CELEX lexical database [CD-ROMJ.* Linguistic Data Consortium, University of Pennsylvania.　[rDN]

Bard, E. & Shillock, R. (19931 Competitor effects during lexical access: Chasing Zipfs tail. In: *Cognitive models of speech processing: The second Spedonga meeting,* ed. G. T. M. Altmann & R. Shillcock. Erlbaum.　[RSN]

Basso. A., Casati, G. & Vignolo, L. A. (1977) Phonemic identification defect in aphasia. *Cortex* 13:85-95,　[rDN]

Becker. C. A. (1980) Semantic context effects in visual word recognition: An analvsis of semantic strategies. *Memory and Cognition* 8:493-5)2.　[aDN]

Bertelson, P. (1986) The onset of literacy-: Liminal remarks. *Cognition* 24:1-30.
!JM

Boland, j. E. (1997) The relationship Ijetvveen syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes* 12:423-84. [aDM

Boland, J. E. *&c* Cutler, A. (1996) Interaction with autonomy; Multiple output models and the inadequacy of the Great Divide. *Cognition* 58:309-20. [aDN]

Boothrovd, A. & Nirtrouer, S. (198S) Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America* S4(D:101-14.　[JRB,TMN]

Bourlard, H. & Morgan, N. (1994) *Connectionist speech recognition: A hybrid approach.* Kluwer Academic.　[TMN]

Bradley D. C. & Forster. *K.* I. (1987) A reader's view of listening. *Cognition* 25:103-34.　[JCZ]

Bradlev, L. & Bryant. P. E. (1983) Categorising sounds and learning to re a d - a causal connection. *Nature* 301:419-21.　[arDN]

Bradlow, A. R. & Kingston, J. (1990) Cognitive processing in the perception of speech. *Journal of the Acoustical Societv/ of America* 88:S56. (Abstract 1.　[JK]

Brancazio, L. (1998^ Contributions of the lexicon to audiovisual speech perception. Unpublished doctoral dissertation. University of Connecticut.　[LB]
　　(1999) lexical influences on the McGurk effect. *Proceedings of the International Conference on Auditory-Visual Speech Processing '99,* 67-73.　[LB]

Brookshire, R. H. (1972) Visual and auditory sequencing by aphasic subjects. *journal of Communication Disorders* 5:259-69.　[RMW]

Brown. C. M.. van Berkum, J, J. A. & Hagoort, P. (in press) Discourse before gender: An event-related potential study on the interplay of semantic and syntactic information during spoken language understanding. *Journal of Psycholingu is tic Resea rch. [* rDN]

Brown. C. M.. Hagoort, P. & Ter Keurs, M. (1999) Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience* 11:261-81.　[rDN]

Bullier, J.. Munk, M. H. J. & Novvak, L. G. (1993) Corticocortical connections sustain interarea synchronization. *Concepts in Neuroscience* 4:101-17. [MM]

Burton, M. YV. (1992) Syllable priming in auditor) word recognition. Paper presented at the 33rd annual meeting of the Fsvehonomic Society, St. Louis, MO.　[LMS1

Burton. M.\Y,, Baum. S. R. & Blumstein, S. E. (1989) Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Perfonuance* 15:567-75. [aDN!

Cairns, P., Shillock. R., Chater, N. & Levy, J. (1995) Bottom-up connectionist modelling of speech. In: *Connectionist models of memortf and language,* ed. J. P. Lew. D. Bairaktaris, J. A. Buliinaria & P. Cairns. University College London Press.　[JRB. aDX]
　11997) Bootstrapping word boundaries: A bottom-up corpus based approach to speech segmentation. *Cognitive Psychology* 33:111-53.　[MGG]

Carello, C, Turvey M. T. & Lukatela, G. (1992) Can theories of word recognition remain stubbornly nonphonological? In: *Orthography, phonology, morphology, and meaning,* ed. R. Frost & L. Katz. North-Holland.　[jCZ]

Cannon, A. & Nachshon, I. (1971) Effect of unilateral brain damage on perception of temporal order. *Cortex* 7:410-18.　[RMW]

Carpenter, C. A. & Grossberg, S. (1987) A massively parallel architecture *for a* self-organizing neural pattern recognition *machine. Computer Vision, Graphics and Image Processing* 7:54-115.　[aDX]

Carpenter. C. A., Grossberg, S. & Rosen, D. (1991) ART2-A: An adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks* 4:493-504.　[rDN]

Carpenter, R. H. S. (1999) Visual selection: Neurons that make up their minds. *Current Biology* 9:R595-98.　[rDN]

Case, P., Tuller, B., Ding, M. Z. & Kelso, J. A. S. (1995) Evaluation of a dynamical model of speech perception. *Perception ami Psychophysics* 57:977-88.　[JCZ]

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S. & Ingvar, M. (1998) The illiterate brain; Learning *to* read and write during childhood influences the functional organization of the brain. *Brain* 121:1053-63.　[rDN]

Cheesman, M, F. & Greenwood, K. G. (1995) Selective adaptation by context conditioned fricatives. *Journal of the Acoustical Society of America* 97:531-38. [aDN]

Chomskv, N. & Halle, M. (1968) *The sound pattern of English.* Harper and Row. [KNS]

Churchland, P. S., Ramachandran, V S. & Sejnowski, T. J. (1994) A critique of pure vision. In: *Large-scale neuronal theories of the brain,* ed. C. Koch & J. L. Davis. MIT Press.　[MKT]

Guff, M. S. *6c* Luce, P. A. (1990) Similarity neighborhoods of spoken two-syllable words; Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception and Performance* 16:551-63.　[aDN]

Cohen, M. & Grossberg, S. (1986) Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short-term memory. *Human Neurobiology* 5:1-22.　[SG, rDN]

Colombo, L. (1986) Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance* 12:226-34.　[LMS]

Connine. C. M. (1994) Vertical and horizontal similarity in spoken word recognition. In: *Perspectives on sentence processing,* ed. C. Clifton, Jr., L. Frazier & K. Rayner. Erlbaum.　[CMC]

Connine, C. M. & Clifton, C. (1987) Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 13:291-99.　[CMC, aDN]

Connine, C M. & Titone, D. (1996) Phoneme monitoring. *Language and Cognitive Processes* 11:635-64.　[aDN]

Connine, C. M., Titone, D., Deelman, T. & Blasko, D. (1997) Similarity mapping in spoken word recognition. *Journal of Memory and Language* 37:463-80. [PAL, RSN, AGS]

Coulson, S., King, J. & Kutas, M. (1998) Expect the unexpected: Event-related brain potentials to morphosyntactic violations. *Language and Cognitive Processes* 13:21-58.　[FI] '

Crain, S. & Steedman, M. (1985) On not being led up the garden path: The use of context by the psychological syntax processor. In: *Natural language processing,* ed. D. R. Dowry, L. Kartunen & A. M. Zwicky. Cambridge University Press.　[aDN]

Cutler, A. & Chen, H.-C. (1995) Phonological similarity effects in Cantonese word recognition. *Proceedings of the International Congress on Phonetic Sciences, Stockholm, Sweden, vol. I,* 106-109.　[LMS]

Cutler, A., Mehler, J., Norris, D. & Segui, J. (1987) Phoneme identification and the lexicon. *Cognitive Psychology* 19:141-77.　[JRB, arDN]

Cutler, A. <k Norris, D. (1979) Monitoring sentence comprehension. In: *Sentence processing: Psycholingimtic studies presented to Merrill Garrett,* ed. VV. E. Cooper & E. C. T. Walker. Erlbaum.　[arDN]
　(1999) Sharpening Ockham's razor. *Behavioral and Brain Sciences* 22:40-41. [ASM]

Cutler, A. & Otake, T. (1994) Mora or phoneme? Further evidence for language-specific listening. *Jou rnal of Memory and Language* 33:824-44.　[aDN]

David, A. S., Woodruff, P. W. R., Howard, R., Mellers, J. D. C., Brammer, M., Bullmore, E,, Wright, I., Andrew, C. & Williams, S. C. R. (1996) Auditory hallucinations inhibit exogenous activation of auditor)' association cortex. *NeuroReport* 7:932-36.　[RS]

Davis, M., Marslen-Wilson, W. D. & Gaskell, M. G. (1997a) Ambiguity and competition in lexical segmentation. In: *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society, Stanford, CA.*　[WDM-W]

Davis, M., Gaskell, M. G. & Marslen-Wilson, W D. (1997b) Recognising words in connected speech: Context and competition. In: *Proceedings of the Fourth Neural Computation and Psychology Workshop: Connectionist representations,* ed. J. A. Buliinaria. Springer-Verlag.　[WDM-W]

Demonet. J. F., Price, C, Wise, R. & Frackowiak, R. (1994) A PET study of cognitive strategies in normal subjects during language tasks. Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain* 117:671-82.　[rDN]

Deschenes, M. or Hu, B. (1990) Electrophysiology and pharmacology of the corticothalamic input to lateral thalamic nuclei; An intracellular study in the cat. *European Journal of Neuroscience* 2:140-52.　[MM]

Donchin, E. & Coles, M.G.H. (1988) Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences* 11:357-74.　[rDN]

Duma\; N. & Radeau, M. (1997) Rime and syllabic effects in phonological priming between French spoken words. *Proceedings of Eurospeech V7, Rhodes, Greece, vol. 4,* 2191-94.　[LMS]

Eimas, P. D., Marcovitz Hornstein, S. B. & Payton. P. (1990) Attention and the role

of dual codes in phoneme monitoring. *Journal of Memory and Language* 29:160-80. faDN]

Eimas, P. D. & Nygaard, L. C. (1992) Contextual coherence and attention in phoneme monitoring. *Journal of Memory and Language* 31:375-95. [aDN]

Einstein, A. & Infeld, L. (1966) *The evolution of physics*. Simon and Schuster. (Originally published in 1938). [JCZ]

Elman, J. L. & McClelland, J. L. (1988) Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language* 27:143-65. [DWM, arDN, AGS]

Emmorey, K. D. (1989) Auditory morphological priming in the lexicon. *Language and Cognitive Processes* 4:73-92. [LMS]

Farah, M. (1989) Semantic and perceptual priming: How similar are the underlying *mechanisms? Journal of Experirnental Psychology: Human Perception and Performance* 15:188-94. [aDN]

Farmer, J. D. (1990) A Rosetta Stone for connectionism. *Physica D* 42:153-87. [JCZ]

Felleman, D. J. & van Essen, C. D. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1:1-47. [SG]

Finkel, L. H. & Edehnan, G. M. (1989) The integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *The Journal ofNeuroscience* 9:3188-208. [MM]

Fodor, J. A. (1983) *Tlie modularity of mind*. MIT Press. [WSM, arDN]
(1985) Precis of *The modularity of mind. Behavioral and Brain Sciences* 8:1-42. [aDN]

Forster, K. I. (1979) Levels of processing and the structure of the language processor. In: *Sentence processing: P-sycholinguistic studies presented to Merrill Garrett,* ed. W. E. Cooper & E. C. T Walker. Erlbaum. [WSM, arDN]

Foss, D. J. & Blank, M. A. (1980) Identifying the speech codes. *Cognitive Psychology* 12:1-31. [aDN]

Foss, D. J. (1969) Decision processes during *sentence* comprehension, effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior* 8:457-62. [rDN]

Foss, D. J. & Gernsbaeher, M. A. (1983) Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior* 22:609-32. [aDN]

Fowler, C. A., Brown, J. M. & Mann, V. A. (1999) Compensation for coarticulation in audiovisual speech perception. *Proceedings of the XlVth International Congress of Phonetic Sciences, San Francisco,* 639-42. [JK]
(in press) Contrast effects do not underlie effects of preceding liquids on stop consonant identification by humans. *Journal of Experimental Psychology: Human Perception and Performance*. [LB]

Fox, R. A. (1984) Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Percejrtion and Performance* 10:526-40. [rDN]

Frauenfelder, U. H. & Peeters, G. (1998) Simulating the time-course of spoken word recognition: An analysis of lexical competition in TRACE. In: *Localise connectionist approaches to human cognition,* ed. J. Grainger & A. M. Jacobs. Erlbaum. [aDN]

Frauenfelder, U. H. & Segui, J. (1989) Phoneme monitoring and lexical processing: Evidence for associative context effects. *Memory and Cognition* 17:134-40. [aDN]

Frauenfelder, U. H., Segui, J.& Dijkstra, T. (1990) Lexical effects in phonemic processing: Facilitatory or inhibitor)'? *Journal of Experimental Psychology: Human Perception and Performance* 16:77-91. [CMC, PAL, arDN, AGS]

Frazier, L. (1979) *On comprehending sentences: Syntactic parsing strategies*. Indiana Linguistics Club. [aDN]
(1987) Sentence processing: A tutorial review. In: *Attention and performance XII: The psychology of reading,* ed. M. Coltheart. Erlbaum. [aDN]

Freeman, W. (1991) The physiology of perception. *Scientific American* 264:78-85. [PAL, MM]

Friederici, A. D. (1995) The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language* 50:259-81. [FI]

Friederici, A. D., Pfeifer, E. & Hahne, A. (1993) Event-related brain potentials during natural speech processing: Effects *in* semantic, morphological, and syntactic violations. *Cognitive Brain Research* 1:183-92. [FI]

Frith, C. & Dolan, R. J. (1997) Brain mechanism associated with top-down processes in perception. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 352:1221-30. [MM]

Frost, S. J., Fowler, C. A. & Rueckl, J. G. (1998) Bidirectional consistency: Effects of a phonology common to speech and reading, (submitted). [JCZ]

Ganong, W. F. (1980) Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6:110-25. [LB, arDN, MP]

Gaskell, M. G. & Marslen-Wilson, W. D. (1995) Modeling die perception of spoken words. In: *Proceedings of the 17th Annual Conference of the Cognitive Science Society,* ed. J. D. Moore & J. F. Lehman. Erlbaum. [WDM-W, aDN]
(1997) Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes* 12:613-56. [MGG, DWG, JG, WDM-W, aDN]
(1998) Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 24:380-96. [aDN]
(1999) Ambiguity; competition and blending in speech perception. *Cognitive Science,* (in press). [WDM-W]

Gibbs, P. & Van Orden, G. C. (1998) Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 24:1162-87. [JCZ]

Goldinger, S. D. (1998) Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105:251-79. [PAL]
(1999) Only the shadower knows: Comment on Hamburger and Slowiaczek (1996). *Psychonomic Bulletin and Review* 6:347-51. [LMS]

Goldinger, S. D., Luce, P. A., Pisoni, D. B. & Marcario, J. K. (1992) Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:1211-38. [aDN]

Gotdob, L. R., Goldinger, S. D., Stone, G. O. & Van Orden, G. C. (1999) Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance* 25:561-74. [JCZ]

Gow, D. W. (submitted a) Assimilation and anticipation: Phonological modification and word recognition in connected speech. [DWG]
(submitted b) Does assimilation produce lexical ambiguity? [DWG]

Grainger, J. & Jacobs, A. M. (1994) A dual read-out model of word context effects in letter perception: Further investigations of the word-superioritv effect. *Journal of Experimental Psychology: Human Perception and Performance* 20:1158-76. [JG, arDN, JCZ]
(1996) Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review* 103:674-91. [JG, aDN]
(1998) On Iocalist connectionism and psychological science. In: *Local 1st connectionist approaches to human cognition,* ed. J. Grainger & A. M. Jacobs. Erlbaum. [JG]

Green, K. P. & Miller, J. L. (1985) On the role of visual rate information in phonetic perception. *Perception and Psychophysics* 38:269-76. [LBl

Grossberg, S. (1978) A theory of human memorv: Self-organization and performance of sensory-motor codes, maps, and plans. In: *Progress in theoretical biology,* ed. R. Rosen & F. SnelJ. Academic Press. [SG]
(1980) How does a brain build a cognitive code? *Psychological Review-* 87:1-51. [SG, MM]
(1986) The adaptive self-organization of serial order in behavior: Speech. language, and motor control. In: *Pattern recognition by humans and machines, vol 1: Speech perception,* ed. E. C. Schwab & H. C. Nusbaum. Academic Press. [SG, PAL]
(1995) The attentive brain. *American Scientist* 83:438-49. [SG]
(1999a) How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. *Spatial Vision* 12:163-85. [SG]
(1999b) The link between brain learning, attention, and consciousness. *Consciousness and Cognition* 8:1-44. [SG. MM]

Grossberg, S., Boardman, I. & Cohen, M. (1997a) Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance* 23:481-503. [SG, PAL, MM. rDN]

Grossberg, S., Mingolla, E. & Ross, W. (1997b) Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neiirosdrnces* 20:106-11. [SG]

Grossberg, S. & Mvers, C. (1999) The resonant dynamics of conscious speech: Interword integration and duration-dependent backward eitects. *Psychological Review,* (in press). [SG]

Grossberg, S. & Stone, G. O. (1986) Neural dvnamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review* 93:46-74. [SG, PAL]

Gunter, T C, Friederici. A. D. & Schriefers, H. (1998) Gender violations, semantic expectancy and ERPs. Poster presented at the XII International Conference on Event-Related Potentials, Cambridge, USA. [FI]

Gunter, T C., Stovve, L. A. & Mulder, G. (1997) When syntax meets semantics. *Psychophysiology* 36:126-37, [FI]

Hagoort, P. & Brown, C. M. (in press) Semantic and syntactic effects of listening to speech compared to reading. *Neuropsychologia*. [rDN]

Hahne, A. & Friederici, A. D. (1999) Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience* ll(2):194-205. [FI]

Hamburger, M. B. & Slowiaczek, L. M. (1996) Phonological priming reflects lexical competition in auditory word recognition. *Psychonomic Bulletin and Review* 3:520-25. [LMS]

(1999) On the role of bias in dissociated phonological priming effects: A reply to Goldinger (1999). *Psychonomic Bulletin and Revietv* 6:352-55.   [LMS]

Hanes, D. P. & Schall, J. D. (1996) Neural control of voluntary movement initiation. *Science* 274:427-30.   [rDN]

Haveman, A. P. (1997) *The open-/closed-class distinction in spoken-word recognition.* Ph. D. thesis, University of Nijmegen.   [rDN]

Hines, T. (1999) A demonstration of auditory top-down processing. *Behavior, Research, Methods, Instruments and Computers* 31:55-56.   [MM]

Hintzman, D. (1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Revietv* 93:411-28.   [PAL]

Hornik, K., Stinchcombe, M. & White, H. (1989) Multilayer feedforward networks are universal approximators. *Nettral Networks* 2:359-66.   [DWM, rDN]

Hupe, J. M., James, A. C, Paynes, B. R., Lomber, S. G., Girad, P. & Bullier, J. (1998) Cortical feedback improves discrimination between figure and background bv VI, V2 and V3 neurons. *Nature* 394:784-87.   [MM]

herson. P.. Bernstein, L. E. & Auer, E. T. (1998) Modeling the interaction *of* phonemic intelligibility and lexical structure in audknisual word recognition. *Speech Communication* 26:45-63.   [LB]

Jacobs, A. M.. Rev, A., Ziegler, J. C. & Grainger, J. (1998) MROM-P: An interactive activation, multiple read-out model of orthographic and phonological pnx-esses in visual word recognition. *In: iMcalist connecttonist approaches to human cognition,* ed. J. Grainger & A. M. Jacobs. Erlbaum.   [JCZ]

Jusczvk, P. W. (1997) *The discovery of spoken language.* MIT Press. [aDN]

Jusczvk, P. \V. &: Aslin, R. N. (1995) Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology* 29:1-23.   [PWJ]

Jusczvk, P. W., Hohne, E. A. & Bauman, A. (1999) Infants' sensitivity to allophonic cues for word segmentation. *Perception and Psychophysics* 61:1465-76. [PWJ]

Jusczvk, P. W, Houston, D. & Newsome, M. (1999) The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39:   [PWJ]

Kawamoto, A. H. & Zemblidge, J. (1992) Pronunciation of homographs. *Journal of Memory and Uinguage* 31:349-74.   [JCZ]

Kelso, J. A. S. {1995) *Dynamic patterns: The self-organization of brain and behavior.* MIT Press.   [JCZ]

Kingston, J. (1991) Integrating articulations in the perception of vowel height, *Phonetica* 48:149-79.   [JK]

Kingston. J, & Diehl, R. L. (1994) Phonetic knowledge. *Language* 70:419-54.   [JK]
(1995) Intermediate properties in the perception of distinctive feature values. In: *Papers in laboratory phonology* /V. ed. B. Connell & A. Arvaniti. Cambridge University Press.   [JK]

Kingston, J. & Maernillan, N. A. (1995) Integrality of nasalization and F_1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Gamer paradigm. *Journal of the Acoustical Society of America* 97:1261-85.   [JK]

Kingston, J., Maernillan, N. A., Walsh Dickey, L., Thorbum, R. & Barrels, C. (1997) Integrality in the perception of tongue root position and voice quality in vowels. *Journal of the Acoustical Society of America* 101:1696-709.   [JK]

Klatt, D. H. (1980) Speech perception: A model of acoustic-phonetic analysis and lexical access. In: *Perception and production of fluent speech,* ed. R. A. Cole. Erlbaum.   [DHW]

Koster, C. J. (1987) *Word recognition in foreign and native language.* Foris. [aDN]

Kugler. P. N. 6c Turvev, M. T. (1987) *Information, natural law, and the self-assembly of rhythmic movement.* Erlbaum.   [JCZ]

Kuhis, M. & Hillvard, S. A. (1983) Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition* 11:539-50.   [FI]

Kutas, M. & Van Petten, C. (1988) Event-related potential studies of language. In: *Advances in psychophysiology, vol. 3,* ed. P. K. Ackles, J. R. Jennings & M. G. H. Coles. JAI Press.   [FI]

Lakatos, I, (1970) Falsification and the methodology of scientific research programmes. In: *Criticism and the growth of knowledge,* ed. I. lakatos & A. Musgrave. Cambridge University Press,   [JCZ]

Levelt, W. J. M.. Roelofs, A. & Meyer. A. S, (1999) A theury of lexical access in speech production. *Behavioral and Brain Sciences* 22(1): 1—75.   [ASM, arDN]

LeWnson, S. C 12000) *Presumptive meanings: The theory of generalized conversational implicature.* MIT Press.   [rDN]

Liberman, I. Y. (1973) Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society* 23:65-77.   [arDN]

Lotto, A. J. & Kluender. K. R. (1998) General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception and Psychophysics* 60:602-19.   [JK]

Luce, P. A. & Pisoni, D. B. (1998) Recognizing spoken words; The neighborhood activation model. *Ear and Hearing* 19:1-36.   [MM]

Luce. R, D. (1959) *Individual choice behavior. Wiley.*   [aDN]
(19631 Detection and recognition. In: *Handbook of mathematical psychology, vol I,* ed. R. D. Luce, R. T. Bush & E. Galanter. Wiley.   [aDN]

Lurk. S, J., Chelazzi, L.. Hillvard, S. A. & Desimone, R. (1997) Neural mechanisms of spatial selective attention in areas VI, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology* 77:24-42. [MM]

Lukatela, G. & Turvey, M. T. (1998) Reading in two alphabets. *American Psychologist* 53:1057-72.   [JCZ]

Lupker, S- J. *&c* Colombo, L. (1994) Inhibitory effects in form priming: Evaluating a phonological competition explanation. *Journal of Experimental Psychology: Human Perception and Performance* 20:437-51.   [LMS]

Luria, A. (1973) *The working brain.* Penguin Books.   [PAL]

MacDonald, M. C, Pearlmutter, N. J. & Seidenberg, M. S. (1994) Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101:676-703.   [aDN]

MacKay, W. A. (1997) SjTichronized neuronal oscillations and their role in motor processes. *Trends in Cognitive Sciences* 1:176-83.   [MM]

Maernillan, N. A., Kingston, J., Thorbum, R., Walsh Dickey, L. *&c* Bartels, C. (1999) Integrality of nasalization and F_r II. Basic sensitivity and phonetic labeling. *Journal of the Acoustical Society of America* 106:2913-32.   [JK]

Mann, V. A. (1980) Influence oi preceding liquid on stop-consonant perception. *Perception and Psychophysics* 28:407-12.   [JK]
(1986a) Phonological awareness: The role of reading experience. *Cognition* 24:65-92.   [aDN]
(1986b) Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of "1" and "r". *Cognition* 24:169-96.   [rDN]

Mann, V. A. & Repp, B. H. (1981) Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America* 69:546-58.   [JK, arDN]

Manuel, S. Y. (1992) Vocal reduction and perceptual recovery in casual speech. *Journal of the Acoustical Society of America* 91:4(2):2388.   [DWG]
(1995) Speakers nasalize /Q/ after /n/, but listeners still hear /Q/. *Journal of Phonetics* 23:453-76.   [KNS]

Marrocco, R. T, McClurkin, J. W. & Young, R. A. (1982) Modulation of lateral geniculate nucleus cell responsiveness by visual activation of the corticogeniculate pathway. *The Journal of Neuroscience* 2:256-63.   [MM]

Marslen-Wilson, W. D. (1993) Issues of process and representation in lexical access. In: *Cognitive models of speech processing: The second Sperlonga meeting,* ed. G. T. M. Altmann & R. Shillock. Erlbaum.   [aDN]

Marslen-Wilson, W. D., Moss, H. E. & van Halen, S. (1996) Perceptual distance and competition in lexical access. *Journal of Experimental Psycfwhgy: Human Percef)tion and Performance* 22:1376-92.   [aDN]

Marslen-Wilson, W D. & Tyler, L. K. (1980) The temporal structure of spoken language understanding. *Cognition* 8:1-71.   [WDM-Wj

Marslen-Wilson, W D. & Warren, P. (1994) Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review* 101:653-75.   [LB, MGG, PAL, WDM-W, arDN, AGS, LMS, MKT, DHW]

Marslen-Wilson, W. D. & Welsh, A. (1978) Processing interactions during word recognition in continuous speech. *Cognition* 10:29-63.   [DWG]

Massaro, D. W. (1973) Perception of letters, words, and nonwords. *Journal of Experimental Psychology* 100:49-53.   [DWM]
(1975) *Understanding language: An infomuition-processing analysis of speech perception, reading, and Psycholinguistics.* Academic Press.   [DWM]
(1978) A stage model of reading and listening. *Visible Language* 12:3-26. [aDN]
(1979) Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception and Performance* 5:595-609. [DWM]
(1987) *Speech perception by ear and eye: A paradigm for psychological inquiry.* Erlbaum.   [LB, aDN]
(1988) Some criticisms of connectionist models of human performance. *Journal of Memory and Language* 27:213-34.   [DWM]
(1989a) Testing between the TRACE model and the Fuzzy Logical Model of Speech Perception. *Cognitive Psychology* 21:398-421.   [aDN]
(1989b) Multiple book review *of Speech perception by ear and eye: A paradigm for psychological inquiry. Behavioral and Brain Scietwes* 12:741-94.   [aDN]
(1996) Integration of multiple sources of information in language processing. In: *Attention and performance X\'l: Information integration in perception and communication,* ed. T. Inui & J. L. McClelland. MIT Press,   [DWM, aDN]
(1998) *Perceiving talking faces: Frotn speech perception to a behavioral principle.* MIT Press.   [DWM, aDN]

Massaro, D. W. & Cohen, M. M. (1983a) Categorical or continuous speech perception: A new test. *Speech Communication* 2:15-35.   [GCO]
(1983b) Phonological context in speech perception. *Perception and Psychophysics* 34:338-48.   [DWM, aDN]
(1991) Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology* 23:558-614.   [DWM, arDN, GCO]

Massaro, D, W, Cohen, M. M., Campbell, C. S. & Rodriguez, T. (submitted) Bayesian method of model selection validates FLMP. *Psychonomic Bulletin andBevieiv.*   [DWM]

Massaro, D. W. & Cowan, N. (1993) Information processing models: Microscopes of the mind. *Annual Review of Psychology* 44:383-425. [aDN]

Massaro, D. W. & Oden, G. C. (1980a) Evaluation and integration of acoustic features in speech perception. *Journal of the Acoustical Society of America* 67:996-1013. [GCO]

(1980b) Speech perception: A framework for research and theory. *In: Speech and language: Advances in basic research and practice, vol 3*, ed. N. J. Lass. Academic Press. [GCO]

(1995) Independence of lexical context and phonological information in speech perception. *Journal of Experimental Psychology: Learning Memory, and Cognition* 21:1053-64. [aDN, GCO]

Mattys, S. L. & Jusczyk, P. W. (submitted) Phonotactic cues for segmentation of fluent speech by infants. [PWJ]

Mattys, S. L., Jusczyk, P. W., Luce, P. A. & Morgan, J. L. (1999) Word segmentation in infants: How phonotactics and prosody combine. *Cognitive Psychology* 38:465-94. [PWJ]

McClelland, J. L. (1986) The programmable blackboard model of reading. In: *Parallel distributed processing vol. 2: Psychological and biological models,* ed. J. L. McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [rDN]

(1987) The case for interactionism in language processing. In: *Attention and performance XIV. Tfie psychology of reading,* ed. M, Coltheart. Erlbaum. [rDN]

(1991) Stochastic interactive processes and the *effect* of context on perception. *Cognitive Psychology* 23:1-44. [arDN]

McClelland, J. L. & Elman, J. L. (1986) The TRACE model of speech perception. *Cognitive Psychology* 18:1-86. [TMN, aDN, JV]

McClelland, J. L. fir Johnston, J. (1977) The role of familiar units in perception of words and nonwords. *Perception and Psychophysics* 22:249-61. [aDN]

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

McClelland, J. L., St. John, M. & Taraban, R. (1989) Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes* 4:287-336. [aDN]

McGurk, H. & MacDonald, J. (1976) Hearing Ups and seeing voices. *Nature* 264:746-48. [LB, aDN]

McNeill, D. & lindig, K. (1973) The perceptual reality of phonemes, syllables, words, and sentences. *Journal of Verbal Learning and Verbal Behavior* 12:419-30. [PAL]

McQueen, J. M. (1991) The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance* 17:433-43. [arDN, MP]

(1996) Phonetic categorisation. *Language and Cognitive Processes* 11:655-64. [aDN]

McQueen, J. M., Cutler, A., Briscoe, T. & Norris, D. (1995) Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes* 10:309-31. [aDN]

McQueen, J. M., Norris, D. & Cutler, A. (1994) Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning Memory, and Cognition* 20:621-38. [aDN]

(1999a) Lexical influence in phonetic decision-making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance* 25:1363-89. [MGG, PAL, WDM-W, arDN, AGS, DHW]

(1999b) The time course of lexical involvement in phonetic categorisation. Presented at the 137th meeting of the Acoustical Society of America, Berlin, March 1999. (Abstract in *Journal of the Acoustical Society of America* 105:1398). [rDN]

(1999c) Lexical activation produces impotent phonemic percepts. Presented at the 138th meeting of trie Acoustical Society of America, Columbus, OH, November 1999. (Abstract in *Journal of the Acoustical Society of America* 106:2296). [rDN]

Mecklinger, A., Schriefers, H., Steinhauer, K. fit Friederici, A. D. (1995) Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition* 23:477-94. [FI]

Mehler, J. & Segui, J. (1987) English and French speech processing. In; *The psycfiophysics of speech perception,* ed. M. E. H. Schouten. Martinus Nijhoff. [aDN]

Mergner, T, Huber, W. & Becker, W. (1997) Vestibular-neck interaction and transformation of sensory coordinates. *Journal of Vestibular Research* 7:347-67. [MM]

Miller, G. A., Heise, G. A. & Lichten, W. (1951) The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psyciology* 41:329-35. [JRB]

Miller, J. (1994) On the internal structure of phonetic categories: A progress report. Cognition 50:271-85. [DWG]

Miller, J. L. fit Liberman, A. M. (1979) Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception and Psychophysics* 25:457-65. [DHW]

Monseil, S. fit Hirsh, K. W. (1998) Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24:1495-520. [LMS]

Moon, C, Bever, T. G. fit Fifer, W. P (1992) Canonical and non-canonical syllable discrimination by two-day-old infants. *Journal of Child Language* 19:1-17. [rDN]

Morais, J. (1985) Literacy and awareness of the units of speech: Implications for research on *the* units of perception. Linguistics 23:707-21. [aDN]

Morais, J., Bartelson, P., Cary, L. & Alegria, J. (1986) Literacy training and speech segmentation. *Cognition* 24:45-64. [aDN]

Morais, J., Cary, L., Alegria, J. fie Bartelson, P. (1979) Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7:323-31. [aDN]

Morton, J. (1969) Interaction of information in word recognition. *Psyclwlogical Review* 76:165-78. [aDN]

Morton, J. & Long, J. (1976) Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior* 15:43-51. [aDN]

Miinte, T. F., Heinze, H. fir Mangun, G. R. (1993) Dissociation of brain activity related to syntactic and semantic aspects of language. *Jottrnal of Cognitive Neuroscience* 5:335-44. [FI]

Murray, W. S. fit Rowan, M. (1998) Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research (CUNY Special Issue)* 27:1-22. [WSM]

Myung, I. J. (in press) The importance of complexity in model selection. *Journal of Mathematical Psyciology.* [MP]

Myung, I. J & Pitt, M. A. (1997) Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin and Review* 4:79-95. [DWM]

(1998) Issues in selecting mathematical models of cognition. In: *Localist connectionist approaches to human cognition,* ed. J. Grainger fit A. Jacobs. Erlbaum. [MP]

Nearey, T. (1997) Speech perception as pattern recognition. *Journal of the Acoustical Society of America* 101(6):3241-54. [TMN]

(forthcoming) The factorability of phonological units *in* speech perception: Simulating results on speech reception in noise. In: *Festschrift for Bntce L. Deriving* ed. R. Smyth. [JRB, TMN]

Neville, H. J., Nicol, J., Barss, A., Forster, K. & Garrett, M. (1991) Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3:155-70. [FI]

Newell, A. (1973) You can't play 20 questions with nature and win: Projective *comments on* the papers of this symposium. In: *Visual information processing,* ed. W. G. Chase. Academic Press. [aDN]

Newman, R. S., Sawusch, J. R. fit Luce, P. A. (1997) Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance* 23(3):873-89. [LB, RSN, arDN]

Nobre, A. C, Allison, T. 6c McCarthy, G. (1998) Modulation of human extrastriate visual processing by selective attention to colours and words. *Brain* 121:1357-68. [MM]

Norris, D. G. (1980) *Serial and parallel models of comprehension.* Ph. D. thesis, University of Sussex. [rDN]

(1986) Word recognition: Context effects without priming. Cognition 22:93-136. [JRB, aDN]

(1987) Syntax, semantics and garden paths. In: *Progress in the psychology of language, vol. 3,* ed. A. W. Ellis. Erlbaum. [aDN]

(1992) Connectionism: A new breed of bottom-up model? In: *Connectionist approaches to language processing,* ed. N. Sharkey fit R. Reiley. Erlbaum. [rDN]

(1993) Bottom-up connectionist models of "interaction." In: *Cognitive models of speech processing: The second Sperlonga meeting,* ed. G. T. M. Altmann & R. Shillcock. Erlbaum. [JRB, WDM-W, aDN]

(1994a) A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psyciology: Human Perception and Performance* 20:1212-32. [arDN]

(1994b) Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52:189-234. [MGG, JG, TMN, arDN, JV]

(1995) Signal detection theory and modularity: On being sensitive to the power of bias models of semantic priming. *Journal of Experimental Psychology: Human Perception and Performance* 21:935-39. [aDN]

(submitted) Feedback consistency in visual word recognition: Don't try going the wrong way up a one-way street. [rDN]

Norris, D. G., McQueen, J. M. & Cutler, A. (1995) Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:1209-28. [arDN]

Norris, D. G., McQueen, J. M., Cutler, A. & Butterfield, S. (1997) The possible-word constraint in the segmentation of continuous speech. *Cognitive Psyciology* 34:193-243. [aDN]

Norris, D. G. & Wise. R. (1999) The study *of* prelexk-al and lexical processes in

comprehension: Psycholinguistics and functional neuroimaging. In: *The cognitive neurosciences,* ed. M. Gazzaniga. MIT Press. [rDN]

Oden, G. C. (1978) Semantic constraints and judged preference for interpretations of ambiguous sentences. *Memory and Cognition* 6:26-37. [GCO]

(1983) On the use of semantic constraints in guiding syntactic analysis. *International Journal of Man-Machine Studies* 19:335-57. [GCO]

Oden, G. C. & Massaro, D. W. (1978) Integration of featural infonnation in speech perception. *Psychological Review* 85:172-91. [aDN, GCO]

Oden, G. C. & McDowell, B. D. (in preparation) The gradedness of perceptual experience in identifying handwritten words. [GCO]

Oden, G. C., Rueckl, J. G. & Sanockt, T. (1991) Making sentences make sense, or words to that effect. In: *Understanding word and sentence,* ed. G. B. Simpson. North-Holland. [GCO]

Ohala, J. J. & Feder, D. (1994) Listeners' normalization of vowel quality is influenced bv "restored" consonantal context. *Phonetica* 51:111-18. [JK]

Osterhout, L. & Holcomb, P. J. (1992) Event-related brain potentials elicited bv syntactic anomaly. *Journal of Memory and Language* 31:785-804. [FI]

(1993) Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes* 8:413-37. [FI]

Osterhout, L. & Mobley, L. A. (1995) Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 34:739-73. [FI]

Otake. T., Yoneyamu, K., Cutler, A. & van der Lugt, A. (1996) The representation of Japanese morale nasals. *Journal of the Acoustical Society of America* 100:3831-42. [aDX]

Paap. K., Newsome, S. L., McDonald, J. E. & Schvaneveldt, B. W. (1982) An activation- verification model for letter and word recognition: The word superiority effect. *Psychological Revictv* 89:573-94. [aDN]

Page, M. (2000) Connectionist modelling in psychology-: A localist manifesto. *Behavioral and Brain Sciences* 23(3). [JG, arDN]

Peereman, B., Content, A. & Bonin, P. (1998) Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Menwry and language* 39:151 -74. [JCZ]

Peeters, G., Franenfelder, U. & Wittenburg, P. (1989) Psychological constraints upon connectionist models of word recognition: Exploring TBACE and alternatives. In: *Connectionism in perspective,* ed. R. Pfeifer, Z. Schreter, F. Fogelman-Soulie & L. Steels. Elsevier. [aDN]

Peitgen, H.-O., Jiirgens, H. & Saupe. D. (1992) *Chaos and fractals: New frontiers of science.* Springe r-Yerlag. [JCZ]

Pitt. M. A. (1995) The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:1037-52. [JRB, arDN]

Pitt, M. A. & McQueen, J. M. (1998) Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39:347-70. [LB, TED, JK, DWM, RSN, arDN, MP, DHW]

Pitt, M. A. & Samuel, A. G. (1993) An empirical and meta-analytic evaluation of the phoneme identification *task. Journal of Experimental Psychology: Human Perception and Performance* 19:699-725. [JRB, arDN]

Press. \V. H., Planners; B. P., Teukolsky, S. A. & Vetterling, W. T. (1986) *Numerical recipes: The art of scientific computing.* Cambridge University Press. [aDN]

Raczaszek, J., Tuller, B., Shapiro, L. P., Case, P. & Kelso, S. (1999) Categorization of ambiguous sentences as a function of a changing prosodic parameter: A dynamical approach. *Journal of Psycholinguistic Research* 28:367-93. [JCZ]

Radeau, M. (1995) Facilitation and inhibition in phonological priming. Paper presented at the 36th Annual Meeting of the Psychonomics Society, Los Angeles. [LMS]

Radeau, M. & Colin, C. (1996) Task effects in phonological priming between spoken words. Paper presented at the 37th Annual Meeting of the Psychonomics Society, Chicago. [LMS]

Radeau, VI, Morais, J. & Dewier, A. (1989) Phonological priming in spoken word recognition: Task effects. *Memory and Cognition* 17:525-35. [LMS]

Radeau, M., Morais, J. & Segui, J. (1995) Phonological priming between monosyllabic spoken words. *Jcnirnal of Experimental Psychology: Human Perception and Performance* 21:1297-311. [LMS]

Radeau, *M.,* Segui, J. & Morais, J. (1994) The effect of overlap position in phonological priming between spoken words. *Proceedings oflCSLP "94, Yokohama, Japan, vol.* 3, 1419-22. [LMS]

Ratcliff, R., Van Zandt, T. & McKoon. G. (1999) Connectionist and diffusion models of reaction time. *Psychological Review,* (in press). [rDN]

Read, C. A., Zhang, Y., Nie, H. & Ding, B. (19861 The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition* 24.31-44. [arDN]

Reicher, G. M. (1969) Perceptual recognition as a function of meaningfulness of stimulus *material. Journal of Experimental Psychology* 81:274-80. [aDN]

Repp, B, H, (1980) A range-frequency effect on perception of silence in speech. *Haskim iMhoratories Status Report on Speech Research* 61.151-65. [SG, rDN]

(1982) Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin* 92:81-110. EJK]

Repp, B. H,, Liberman, A. M., Eccardt, T. & Pesetsky, D. (1978) Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance* 4:621-37. [SG, rDN]

Rhodes, G., Parkin, A. J. & Tremewan, T. (1993) Semantic priming and sensitivity in lexical decision. *Journal of Experimental Psychology: Human Perception and Performance* 19:154-65. [aDN]

Riedel, K. & Studdert-Kennedy, M. (1985) Extending formant transitions may not improve aphasics' perception of stop consonant place of articulation. *Brain and language* 24:223-32. [rDN]

Roberts, M. & Summerfield, Q. (1981) Audiovisual presentation demonstrates that selective adaptation in speech perception is purely auditory. *Perception and Psychophysics* 30:309-14. [aDN]

Rodd, J., Gaskell, M. G. & Marslen-Wilson, W. D. (1999) Semantic competition and the ambiguity disadvantage. In: *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society, Vancouver, Canada, August 1999.* Erlbaum. [MGG]

Roelfsema, P R., Engel, A. K., Konig, P. & Singer, W. (1996) The role of neuronal synchronization in response selection: A biologically plausible theory of structured representations in the visual cortex. *Journal of Cognitive Neuroscience* 8:603-25. [MM]

Rosier, ¥., Friederici, A. D., Putz, P. & Hahne, A. (1993) Event-related brain potentials while encountering semantic and syntactic constraint violation. *Journal of Cognitive Neuroscience* 5:345-62. [FI]

Rubin, P., Turvey, M. T. & van Gelder, P. (1976) Initial phonemes are detected faster in spoken words than in nonwords. *Perception and Psychophysics* 19:394-98. [aDN]

Rueckl, J. G. & Oden, G. C. (1986) The integration of contextual and featural information during word identification. *Journal of Memory and Language* 25:445-60. [GCO]

Rumelhart, D. & McClelland, J. (1982) An interactive activation model of context effects in letter perception-. Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review* 89:60-94. [aDN]

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) Statistical learning by 8-month-old infants. *Science* 274:1926-28. [PWJ]

Saldana, H. M. & Rosenblum, L. D. (1994) Selective adaptation in speech perception using a compelling audiovisual adaptor. *Journal of the Acoustical Society of Anwrica* 95:3658-61. [aDN]

Satin, P.-A. & Bullier, J. (1995) Cortico-cortical connections in the visual system: Structure and function. *Physiological Reviews* 75:107-54. [MM]

Samuel, A. G. (1981) Phonemic restoration: Insights from a new methodology. *Jimmal of Experimental Psychology: General* 110:474-94. [SG, aDN, AGS]

(1981b) The rule of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance* 7:1124-31. [SG]

(1986) Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive Psychology* 18:452-99. [aDN]

(1987) Lexical uniqueness effects on phonemic restoration. *Journal of Memory and language* 26:36-56. [aDN]

(1996a) Does lexical information influence the perceptual restoration of pho-nemesp/onrna/ *of Experimental Psychology: General* 125:28-51. [aDN]

(1996b) Phoneme restoration. *Language and Cognitive Processes* 11:647-54. [aDN]

(1997) Lexical activation produces potent phonemic percepts. *Cognitive Psychology* 32:97-127. [arDN, AGS]

Samuel, A. G. & Kat, D. (1996) Early levels of analysis of speech. *Journal of Experimental Psychology: Human Perception and Performance* 22:676-94. [arDN, AGS]

Sawuscii, J. R. & Jusczyk, P. (1981) Adaptation and contrast in the perception of voicing. *Journal of Experimental Psychology: Human Perception and Performance* 7:408-21. [aDN]

Segui, J. & Frauenfekfer, U. (1986) The effect of lexical constraints upon speech perception. In: *Human memory and cognitive capabilities: Mechanisms and performances,* ed. F. Klix & H. Hagendorf. North-Holland. [aDN]

Seidenberg, M. S. & McClelland, J. L. (1989) A distributed, developmental model of visual word-recognition and naming. *Psychological Review* 96:523-68, [WDM-W, rDN]

Shepard, R. (1984) Ecological constraints in internal representation; Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review* 91:417-47. [PAL]

Shillock, R. C. & Bard, E. G. (1993) Modularity and the processing of closed class words. In: *Cognitive modeh of speech processing: The second Sperionga meeting,* ed. ed. G. T M. Altmann & R. C. Shillock. Erlbaum. [RS]

Sillito, A. M., Cudeiro, J. & Murphy, P. C. (1993) Orientation sensitive elements in

the corticofugal influence on centre-surround interactions in the dorsal lateral geniculate nucleus. *Experimental Brain Research* 93:6-16. [MM]

Silverstein, S. M., Matteson, S. & Knight, R. A. (1996) Reduced top-down influence in auditory perceptual organization in schizophrenia. *Journal of Abnormal Psychology* 105:663-67. [MM]

Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* 270:758-64. [MM]

Slaney, M. (1998) A critique of pure audition. In: *Computational auditory scene analysis,* ed. D. Rosenthal & H. G. Okuno. Erlbaum. [MM]

Slowiaczek, L. M. & Hamburger, M. R. (1992) Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:1239-50. [aDN, LMS]

Slowiaczek, L. M., McQueen, J. M., Soltano, E. G. & Lynch, M. (2000) Phonological representations in prelexical speech processing: Evidence from form-based priming. *Journal of Memory and Language,* (forthcoming). [LMS]

Slowiaczek, L. M., Soltano, E. G. & McQueen, J. M. (1997) Facilitation of spoken word processing: Only a rime can prime. Paper presented at the Psychonomic Society Meeting, Philadelphia. [LMS]

Steedman, M. J. & Altmann, G. T. M. (1989) Ambiguity in context: A reply. *Language and Cognitive Processes* 4:SI105-22. [aDN]

Stevens, K. N. (1995) Applying phonetic knowledge to lexical access. In: *Proceedings of the 4th European Conference on Speech Communication and Technology, Madrid,* 1:3-11. [KNS]

Stevens, K. N. & Halle, M. (1967) Remarks on analysis by synthesis and distinctive features. In: *Models for the perception of speech and visual form,* ed. W. Wathen-Dutin. MIT Press. [KNS]

Stone, G. O., Vanhoy, M. D. & Van Orden, G. C. (1997) Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language* 36:337-59. [JCZ]

Stone, G. O. & Van Orden, G. C. (1993) Strategic processes in printed word recognition. *Journal of Experimental Psychology: Hitman Perception and Performance* 19:744-74. [JCZ]

Streeter, L. A. & Nigro, G. N. (1979) The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America* 65:1533-41. [aDN]

Taft, M. (1982) An alternative to grapheme-phoneme conversion rules? *Memory and Cognition* 10:465-74. [JCZ]

Tallal, P. & Piercy, M. (1973) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241:468-69. [RMW]

Tallon-Baudry, C, Bertrand, O., Delpuech, C. & Pernier, J. (1997) Oscillatory 7-band (30-70 Hz) activity induced by a visual search task in humans. *Tfie Journal of Neuroscience* 17:722-34. [MM]

Tanenhaus, M. K. & Donnenwerth-Nolan, S. (1984) Syntactic context and lexical access. *The Quarterly Journal of Experimental Psychology* 36A:649-61. [ERS]

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. & Sedivy, J. (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* 268:1632-34. [ASM, rDN]

Taraban, R. & McClelland, J. L. (1988) Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language* 27:1-36. [aDN]

Ter Keurs, M., Brown, C. M., Hagoort, P. & Stegemann, D. F. (1999) Electrophysiological manifestations of open- and closed-class words in patients with Broca's aphasia with agrammatic comprehension. *Brain* 122:839-54. [rDN]

Tincoff, R. & Jusczyk, P. W. (1996) Are word-final sounds perceptually salient for infants? Poster presented at LabPhon V, Northwestern University, Evanston, IL. [PWJ]

Tononi, G., Sporns, O. & Edelman, G. M. (1992) Reentry and the problem of integrating multiple cortical areas: Simulation of dynamics integration in the visual system. *Cerebral Cortex* 2:310-35. [MM]

Trueswell, J. C. & Tanenhaus, M. K. (1994) Towards a lexicahst framework of constraint-based syntactic ambiguity resolution. In: *Perspectives on sentence processing,* ed. C. Clifton, L. Frazier & K. Rayner. Erlbaum. [aDN]

Tuller, B., Case, P., Ding, M. & Kelso, J. A. S. (1994) The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception and Performance* 20:3-16. [JCZ]

Turner, C. W. & Doherty, K. A. (1997) Temporal masking and the "active process" in normal and hearing-impaired listeners. In: *Modeling sensorimotor hearing loss,* ed. W. Jesteadt. Erlbaum. [MM]

Ullman, S. (1995) Sequence seeking and counter streams: A computational model for bidirectional information flow in the visual cortex. *Cerebral Cortex* 1:1-11. [MM]

Usher, M. & McClelland, J. (1995) On the time course of perceptual choice: A model based on principles of neural computation. Technical Report PDP.CNS.95.5, Department of Psychology, Carnegie-Mellon University. [rDN]

Van Berkum, J. J. A., Brown, C. M. & Hagoort, P. (1999) When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research* 28:555-71. [rDN]

Van Orden, G. & Goldinger, S. (1994) Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance* 20:1269-91. [PAL, JCZ]

Van Orden, G. C, Holden, J. G., Podgomik, M. N. & Aitchison, C. S. (1999) What swimming says about reading: Coordination, context, and homophone errors. *Ecological Psychology* 11:45-79. [JCZ]

Van Petten, C. & Kutas, M. (1991) Electrophysiological evidence for the flexibility of lexical processing. In: *Understanding word and sentence,* ed. G. B. Simpson. Elsevier. [FI]

Vitevitch, M. S. & Luce, P. A. (1998) When words compete: Levels of processing in spoken word recognition. *Psychological Science* 9:325-29. [JRB. RS\, arDN]

(1999) Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40:374-408. [PAL, rDN]

Vroomen, J. & de Gelder, B. (1995) Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of 'Experimental Psychology: Human Perception and Performance* 21:98-108. [arDN]

Wandell, B. A. (1995) *Foundations of vision.* Sinauer. [MKT]

Warren, R. M. (1970) Perceptual restoration of missing speech sounds. *Science* 167:392-93. [RMW[7]]

(1971) Identification times for phonemic components of graded complexity and for spelling of speech. *Perception and Psychophysics* 9:345-49. [RMW]

(1984) Perceptual restoration of obliterated sounds. *Psychological Bulletin* 96:371-83. [SG, MM]

(1999) *Auditory perception: A new analysis and synthesis.* Cambridge University Press. [RMW]

Warren, R. M. & Ackroff, J. M. (1976) Two types of auditory sequence perception. *Perception and Psychophysics* 20:387-94. [RMW]

Warren, R. M. & Bashford, J. A., Jr. (1993) When acoustic sequences are not perceptual sequences: The global perception of auditory patterns. *Perception and Psychophysics* 54:121-26. [RMW]

Warren, R. M,, Bashford, J. A., Jr. & Gardner, D. A. (1990) Tweaking the lexicon: Organization of vowel sequences into words. *Perception and Psychophysics* 47:423-32. [RMW]

Warren, R. M. & Gardner, D. A. (1995) Aphasics can distinguish permuted order of phonemes - but only if presented rapidly- *Journal of Speech ami Hearing Research* 38:473-76. [RMW]

W'arren, R. M., Gardner, D. A., Brubaker, B. S. & Bashford. J. A.. Jr. (1991) Melodic and nonmelodic sequence of tones: Effects of duration on perception. *Music Perception* 8:277-90. [RMW]

Warren, R. M,, Healy, E. W. & Chalikia, M. H. (1996) The vowel-sequence illusion: Intrasubject stability and intersubject agreement of syllabic forms. *Journal of the Acoustical Society of America* 100:2452-61. [RMW]

Warren, R. M. & Obusek, C. J. (1971) Speech perception and phonemic restorations. *Perception and Psychophysics* 9:358-62. [RMW]

Warren, R. M. & Sherman, G. L. (1974) Phonemic restorations based on subsequent context. *Perception and Psychophysics* 16:150-56. [SG]

Warren, R. M. & Warren, R. P. (1970) Auditory illusions and confusions. *Scientific American* 223:30-36. [RMW]

Watt, S. M. & Murray, W. S. (1996) Prosodic form and parsing commitments. *Journal of Psycholinguistic Research (CUNY Special Issue)* 25:91-318. [WSM]

Werker, J. & Tees, R. (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49-63. [rDN]

Whalen, D. H. (1984) Subcategorical phonetic mismatches slow phonetic judgments. *Perception and Psychophysics* 35:49-64. [PAL, aDN, DHW]

(1989) Vowel and consonant judgments are not independent when cued by the same information. *Perception and Psychophifsics* 46:284-92. [DHW]

(1991) Subcategorical phonetic mismatches and lexical access. *Perception and Psychophysics* 50:351-60. [aDN, DHW[7]]

Whalen, D. H. & Samuel, A. G. (1985) Phonetic information is integrated across intervening nonlinguistic sounds. *Perception and Psychophysics* 37:579-87. [DHW]

Wurm, L. H. & Samuel, A. G. (1997} Lexical inhibition and attenrional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory and Language* 36:165-87. [CMC, arDN. AGS]

Yost, W. A. & Nielsen, D. W (1977) *Fundamentals of hearing: An introduction.* Holt, Rinehart and Winston. [MKT]

Zatorre, R. J., Evans. A. C, Meyer, E. & Gjedde, A. (1992) PET studies of phonetic processing of speech: Review', replication and reanalvsis. *Cerebral Cortex* 6:21-30. [rDN]

Ziegler, J. C. & Ferrand, L. (1998) Orthography shapes the perception of speech:

The consistency effect in auditory word recognition. *Psychonomic Bulletin and Revietc* 5:683-89. [JCZ]

Ziegler, J. C. & Jacobs, A. M. (1995) Phonological information provides early sources of constraint in the processing of letter strings. *Journal of Memory and Language* 34:567-93. [JCZ]

Ziegler. j. C, Montant, M. & Jacobs, A. M. (1997a) The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language* 37:533-54. [JCZ]

Ziegler, J. C. & Van Orden, G. C, (1999) Feedback consistency and subjective familiarity, (submitted). [JCZ]

Ziegler, J. C, Van Orden, G. C. & Jacobs, A. M. (1997b) Phonology can help or hurt the perception of print. *Journal of Experimental Psychology: Human Perception and Performance* 23:845-60. [JCZ]

Zue, V. & Shattuck-Hufnagel, S. (1979) The palatalization of voiceless alveolar fricative *Isl* in American English. In: *Proceedings of the 9th International Congress of Phonetic Sciences,* ed E. Fischer-Jorgensen, 1:215. [KNS]