as an additional source of information. We have, in fact, predicted many studies in which higher-order constraints such as transition probability influence segment and word identification (Massaro & Cohen 1983b). Thus, our published mathematical fit (Massaro 1996) still holds but the additional source of information is now happily acknowledged as transition probability rather than coarticulation.

I have argued over the years that quantitative models are necessary to distinguish among theoretical alternatives in psychological inquiry. There has been a resurgence of interest in model testing and selection, with exciting new developments in evaluating the falsifiability and flexibility of models (Massaro et al., submitted; Myung & Pitt 1997). Merge is formulated in terms of a miniature neural network that predicts activation levels that are qualitatively compared to empirical measures of RT. The network requires something between 12 and 16 free parameters to predict the desired outcomes, which basically involve the qualitative differences among a few experimental conditions. In an unheeded paper, I demonstrated that neural networks with hidden units were probably not falsifiable (Massaro 1988), which was later substantiated in a more formal proof (Hornik et al. 1989). I'm worried that mini-models may have the same degree of flexibility, and mislead investigators down a path of limited understanding.

Finally, for once and for all, we would appreciate it if the field would stop claiming that somehow these mini-neural networks are modeling the "mechanisms leading to activation" (sect. 6.3, para. 8), whereas the FLMP is doing something less. The authors claim that "FLMP is not a model of perception in the same way that Merge and TRACE are" (ibid.). One might similarly criticize Sir Isaac Newton's Law of Universal Gravitation, which simply states that the gravitational force FG between any two bodies of mass *m* and M, separated by a distance *r*, is directly proportional to the product of the masses and inversely with the square of their distance. As any "dynamic mechanistic" model should, we have formalized, within the FLMP, the time course of perceptual processing and have made correct predictions about the nature and accuracy of performance across the growth of the percept (Massaro 1979; 1998, Ch. 9; Massaro & Cohen 1991).

Merging speech perception and production

Antje S. Meyer and Willem J. M. Levelt
Max Planck Institute for Psycholinguistics, NL 6500 AH Nijmegen, The Netherlands. **{asmeyer;pim}@mpi.nl**

Abstract: A comparison of Merge, a model of comprehension, and WEAVER, a model of production, raises five issues: (1) merging models of comprehension and production necessarily creates feedback; (2) neither model is a comprehensive account of word processing; (3) the models are incomplete in different ways; (4) the models differ in their handling of competition; (5) as opposed to WEAVER, Merge is a model of metalinguistic behavior.

In their commentary on our recent *BBS* target article on lexical access in speech production (Levelt et al. 1999), Cutler and Norris (1999) praised our rigorous application of Ockham's razor, that is, our effort to design the simplest possible model of lexical access that would be consistent with the available evidence. We proposed a model minimizing inter-level feedback. Our commentary on the target article by Norris, McQueen & Cutler is an obvious place to return the compliment. We are pleased to see them propose a model of spoken word recognition in which there is no feedback from higher to lower level processing units. As Norris et al. point out, the functions of language production and comprehension are intimately related, and the corresponding models should be compatible in their general architecture.

1. Our first comment is that it is precisely this intimate relation between perception and production that forces us to assume some

feedback in the system. In our target article we proposed that the representational levels of lemmas (syntactic words) and lexical concepts are shared between perception and production. Hence, there should be bi-directional activation spreading from concepts to lemmas (in production) and from lemmas to concepts (in perception), that is, full feedback. Uni-directionality of processing (i.e., non-feedback) can only be claimed for those parts of the system that are not shared between perception and production. These are the prelexical and word levels in Merge, and the morphophonological and phonetic levels in WEAVER. (We leave undiscussed here the issue of self-monitoring, which involves the perceptual system in still another way.)

So far there seems to exist perfect complementarity between WEAVER and Merge. Still the two models do not yet fit together like pieces of a jigsaw puzzle, forming a comprehensive and consistent picture of spoken word processing. Rather, each research team has extensively studied certain areas of language processing, leaving others largely uncharted. That is the topic of our next two comments.

2. Some core areas of language processing have not been systematically incorporated in either model. For instance, both models were designed as computational accounts of single word processing. But a comprehensive picture of word processing must include a computational account of the processing of words in their multiword syntactic and semantic contexts. It is not predictable how such a comprehensive account will ultimately affect our partial models of single word processing.

3. There are some areas that have received ample attention in modeling production, but not in modeling comprehension, or vice versa. For instance, the model proposed by Levelt et al. and its computational implementation (WEAVER) include specific assumptions about the mapping from lexical concepts to lemmas and from lemmas to phonological forms. The model proposed by Norris et al. concerns the mapping of the speech input onto lexical forms; the activation of syntactic properties and meanings of words are not part of the model. Precisely where the two systems may be shared (see [1]), no modeling of comprehension is available. On the other hand, Shortlist, which is part of Merge, provides a more detailed treatment of word processing in context than does WEAVER in its account of phonological word formation. What is worse, there are clear mismatches between the Merge and WEAVER:

4. WEAVER has opted for a Luce rule treatment of competition, but Norris et al. opt for a threshold treatment. One argument for the latter approach is that adding a Luce treatment would involve unnecessary reduplication, because Merge already has inhibitory connections among word nodes. There are no inhibitory connections in WEAVER; it fares very well without. We cannot judge whether Merge (or Shortlist for that matter) could be made to run successfully without inhibitory connections, only using Luce's rule (which is somewhat like asking a diesel owner to drive her automobile on gas). There is, however, a crucial point in the background: WEAVER is a model of reaction times (speech onset latencies), whereas Merge is a model of activation levels; it suffices for Merge to display the correct monotonic relation between activation levels and (lexical or phoneme decision) reaction times. This brings us to our final comment:

5. Merge is a model of metalinguistic judgment, whereas WEAVER models the primary word naming process. This reflects marked differences in the production and comprehension research traditions. The major empirical methods in comprehension research have been metalinguistic: phoneme decision, lexical decision, word spotting, and so on. There is nothing wrong with this as long as modeling these tasks involves as a core component the primary process of word recognition. That is the case for Merge, which essentially incorporates Shortlist. One should only start worrying when a different, ad hoc core component is designed for every metalinguistic task to be modeled. In production research the tradition has been to model the chronometry of the primary process of word production, or alternatively the distribution of speech errors. Metalinguistic tasks, for instance lexical decision, gender decision or production phoneme monitoring, are occasionally used in studying word production, but they lead a marginal existence. Granting the significance of these research traditions, we would still see it as an advantage if more primary tasks were used in word comprehension research. The eye scanning paradigm (Tanenhaus et al. 1995) is one possibility, picture/word verification may be another one.

Feedback: A general mechanism in the brain

Marie Montant[1]

Department of Psychology, Carnegie Mellon University, Pittsburgh PA 15213-3890.[1]**montant@lnf.cnrs-mrs.fr**

Abstract: Norris, McQueen & Cutler argue that there is no need for feedback in word recognition. Given the accumulating evidence in favor of feedback as a general mechanism in the brain, I will question the utility of a model that is at odds with such a general principle.

In the neuroscience literature, a large body of evidence suggests that feedback is used by the brain for various aspects of perception, action, language, and attention. In the visual system, for instance, feedback connections are not the exception but the rule. Many anatomical studies have shown that most connections between cortical areas are reciprocal, and, in some cases, like the ventral occipito-temporal pathway (the "what" pathway), feedback connections are predominant (for a review, see Salin & Bullier 1995).

Feedback connections seem to have various roles. They can filter the visual input and improve its quality by changing the sensitivity of the afferent pathways to some aspects of the stimulation. This is done by modifying the balance of excitation and inhibition in lower order cortical areas and in the subcortical nuclei of the afferent pathways (Alonso et al. 1993; Deschenes & Hu 1990; Marrocco et al. 1982; Sillito et al. 1993; Ullman 1995).

Feedback connections are suspected to play an important role in figure-ground separation because they convey top-down expectations about the visual environment that make it possible to segment the visual scene (Hupé et al. 1998). By conveying top-down expectations, feedback connections are also involved in attention-driven modulation of visual processing (e.g., Luck et al. 1997) and visual word recognition (e.g., Nobre et al. 1998).

Feedback is also crucial in the synchronization of adjacent populations of neurons in the cortex, a phenomenon that is considered as the neural mechanism of "feature binding." That is, the binding of the different features of an object for the construction of a unified percept (Bullier et al. 1993; Finkel & Edelman 1989; Freeman 1991; Roelfsema et al. 1996; Singer 1995; Tononi et al. 1992; Ullman 1995). Synchronous neuronal activity would reflect recurrent bottom-up and top-down activation of neural assemblies that code different aspects of a same object during the process of recognition (Tallon-Baudry et al. 1997).

The above described feedback mechanisms are not specific to visual perception. Similar mechanisms have been described in vestibular perception (e.g., Mergner et al. 1997), in auditory perception (e.g., Hines 1999; Slaney 1998; Turner & Doherty 1997), and also in the domain of sensorimotor integration and action (see MacKay 1997, for a review).

If feedback mechanisms are used in such crucial aspects of perception, in most sensory modalities, why should it be that they are not used in speech perception? The authors argue that this is the case because (1) the system can perform an optimal bottom-up analysis of the auditory input on the first pass, (2) feedback cannot improve the quality of the input but, on the contrary, can make the system hallucinate, and (3) feedback solutions are less parsimonious than pure bottom-up solutions. However, these three assumptions are highly questionable.