Modelling vocabulary acquisition in spoken word recognition

Gareth Gaskell

Most current models of spoken word recognition relate to the system in its "steady state" (e.g., McClelland and Elman, 1986; Norris, 1993). That is, they apply to the situation where the listener has already acquired a fully fledged list of spoken words. Even in cases where these models incorporate a learning component (e.g., Gaskell & Marslen-Wilson, 1997), the learning phase does not accurately capture the microstructure of learning - it is merely a way of setting the weights in the steady-state model based on a representation of the language environment.

Recent research (e.g., Gaskell & Dumay, 2003) has emphasised the timecourse of the lexicalisation of novel words, suggesting a system of dual representation, with both fast and slow learning of different aspects of words. When we encounter a new spoken word (e.g., "cathedruke"), the phonological form is easily stored and made use of immediately, but engagement in lexical competition, as measured by its effect on the recognition of other words (e.g., "cathedral"), emerges over a longer timecourse, and possibly relies on consolidation processes related to sleep. This project will employ computational, behavioural and brain imaging methods to add an extra dimension to models of spoken word recognition - the ability to capture short-term changes in the state of the mental lexicon related to the changes in the language environment over the course of a lifetime.

In terms of computational research, integration with brain imaging data will require a shift to more biologically plausible network model, with in particular a realistic implement of learning mechanisms. The behavioural data so far suggest that initial exposure to a novel word generates a relatively robust episodic representation of the form of the word, coupled with other relevant aspects. This "one-shot" learning contrasts with the protracted and delayed consolidation of the novel word representation in the lexical competition network for speech. One computational account of this delay relates to avoidance of "catatstrophic interference" (cf., McClelland, McNaughton and O'Reilly, 1995), but there are other viable mechanisms that could be explored. Other key properties include the incorporation of a continuous time dimension, and the development of realistic distributed semantic representations of words. A network implementation of the model seems crucial for incorporating the st! rong learning aspects of the behavioural data, but this implementation may well be usefully supplemented by statistical/mathematical simulations (cf. Gaskell & Marslen-Wilson, 1999), which have the advantages of greater precision and lower computational costs.

On the neuroimaging side, work is just beginning on the neural consequences of lexicalisation, in collaboration with Matt Davis. Speculation about the nature of the short- and long-term representations formed after exposure to novel items needs full investigation using fMRI, with the intention of identifying brain areas found to be related to immediate and delayed learning of novel words. The use of MEG holds an even greater promise, given its millisecond time resolution and the suggestion that aspects of lexical compensation can be captured in terms of both location and timecourse (e.g., Pulvermuller, Shtyrov & Ilmoniemi, 2003). Here, behavioural and imaging data can be collected in highly comparable ways, offering greater certainly about the location and nature of competitive processes in spoken word recognition.

The goal of the project is to have a computationally and neurally specific model of the acquisition and storage of spoken words. The model will be evaluated with respect to two bodies of data - existing data on the steady-state process of spoken word recognition, plus existing and planned behavioural data on the acquisition and lexicalisation of novel words. Although not part of the planned research, the expectation is that the model should be further applicable to developmental data on language acquisition.

References

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, 89, 105-132.

Gaskell, M. G. & Marslen-Wilson, W.D. (1997). Integrating form and meaning: a distributed model of speech perception. *Language and Cognitive Processes,* 12, 613-656.

Gaskell, M. G. & Marslen-Wilson, W.D. (1999). Ambiguity, competition and blending in spoken word recognition. *Cognitive Science*, 23, 439-462.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition,* 52(3), 189-234.

McClelland, J. L. and Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology, 18*, 1-86.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419-457.

Pulvermüller F., Shtyrov Y., & Ilmoniemi R.J. (2003). Spatiotemporal Dynamics Of Neural Language Processing: An MEG study using Minimum-Norm Current Estimates. *NeuroImage*, 20, 1020-5.