Theories that develop

Journal Article

216 Bilingualism: Language and Cognition 5 (3), 2002, 216–217 © 2002 Cambridge University Press DOI: 10.1017/S1366728902273010

Theories that develop

MICHAEL S. C. THOMAS
School of Psychology, Birkbeck College, University of London, Malet St., London WC1E 7HX, UK.
E-mail: *m.thomas@psychology.bbk.ac.uk, m.thomas@ich.ucl.ac.uk*

The target article represents a significant advance in the level of sophistication applied to models of bilingual word recognition, and Dijkstra and van Heuven are to be congratulated on this endeavour. Bearing in mind the success of the (computational) BIA model in capturing detailed patterns of experimental data, I look forward to future simulation results from the BIA+ when the proposals of this new framework are implemented. It is an essential step to draw a distinction between recognition systems and the decision mechanisms that drive responses, and the authors have provided a novel way of apportioning empirical evidence of context effects in bilingual word recognition across this divide. Given the explanatory weight now being placed on decision mechanisms rather than the word recognition system itself, perhaps indeed it is now time to make some simplifying assumptions about the recognition system and start building detailed computational models of the decision component of the system. Implementation will provide the clarity of theorisation and evaluation of theory viability that have been the hallmark of the BIA model thus far.

However, in this commentary, I want to focus on one particular avenue of theory development which is as yet under-explored in the BIA+ framework. This is how the cognitive structures of the bilingual's word recognition system are acquired and maintained. Dijkstra and van Heuven comment "there are many other aspects of bilingual word recognition that should be considered in the future (for instance, how the model would develop over time and during learning)" (p. 181). It is quite right, of course, to take one step at a time. It is enough work trying to explain the mass of bilingual word recognition data via a static model of the adult system, and the authors have outlined a proposal with great potential on that front. However, many of the interesting issues that characterise the bilingual language system are *dynamic* ones (Thomas & van Heuven, in press).

For example, how are two languages acquired, and the appropriate integrated-yet-differentiated representations established? To what extent are there critical periods or age of acquisition effects in the acquisition of an L2? To what extent are there transfer effects between first and second languages? How is an L2 best acquired – by initial association to an existing L1 or by a strategy that encourages direct contact with semantics (such as picture naming)? How is each language maintained, in terms of on-going patterns of relative dominance and/or proficiency, and to what extent are effects modality specific (i.e., differential across spoken and written language, comprehension and production)? How is each language lost, in terms of aphasia after brain damage in bilinguals, or in terms of the natural attrition of a disused language?

All of these questions demand a dynamic aspect of models of the bilingual system. But the point is not just about the features of bilingualism that remain to be explained. It also concerns the way one goes about constructing a model. For example, at times as I have watched the development of the BIA and now BIA+ framework, I have had a sense that the model is chasing the empirical data rather than predicting it. The authors have rarely put forward strong opinions about which features should be in the model and which not: whether there should be top-down connections or not; whether these connections should be inhibitory or not; whether interlingual homographs should have one or two representations; whether orthographic neighbourhoods should be facilitatory or inhibitory; whether phonological similarity should be inhibitory or facilitatory; and so on. There is little in the way of theory-driven expectation. Rather, data are presented and the model's assumptions altered accordingly. Indeed, sometimes the model's assumptions are altered between experiments, as in the case of the cross-language top-down inhibition in the BIA invoked to simulate one set of results but not another.

Now, this isn't necessarily a bad way of model building. Dijkstra, van Heuven and colleagues have made a great deal of progress building on existing static models from the monolingual literature. However, a complementary approach might seek to generate expectations based on what *function the relevant cognitive system has developed for*. What tasks is the system required to achieve in its normal usage? What representations is it therefore likely to learn to optimally succeed in these tasks (based on reasonable assumptions of the learning mechanisms available)?

These are not idle questions. A developmental perspective suggests theoretical possibilities for the structure of the *adult system* as well. Two examples will suffice, both theoretical issues that Dijkstra and van Heuven spend some time considering.

What is the role of language membership? The BIA/BIA+ framework proposes that word nodes are present in an integrated network where each node has a link to a language node, which marks its membership. Part of the debate concerns the subsequent role of the activity that arrives at each language node. Inspection of the BIA implementation reveals that the connections between word nodes and language nodes are of identical value. Language membership is postulated to be a uniform construct. But is this really a sensible outcome of a plausible learning system? Let us say that the strength of the connection between a word node and its language node is established by Hebbian learning, where the language context establishes the activity on the language node at the same time as

the word node is activated, and their correlated activity strengthens the connection. Will every word node be associated with an equal degree of language context/language node activation? Will every word node be activated equally frequently? It seems unlikely.

In contrast, if we take a learning perspective, the expectation is that the system will evolve the processing structures sufficient to achieve the task at hand. In this case, the aim of the system is to provide the correct output (meaning) given the input (word form). Perhaps language context information will only be used in situations where it is necessary, for example, to disambiguate inputs that are similar across languages but require different outputs. Under this view, language membership may not be an all or nothing tag, but a continuum that depends on the demands of the task. Words with phonological or orthographic features unique to a language may not need to be tagged, whereas words with common features may lean more heavily on language membership information to disambiguate them.

(One might argue that all of a bilingual's words must have a language tag in order that the individual can answer the question "To which language does this word belong?" But such information could simply constitute encyclopaedic knowledge, part of the word's meaning. The extent to which language membership needs to assume an active *processing status* in the bilingual lexicon is a separate question.)

The second example follows on from this idea of the *constraints of learning*. Dijkstra and van Heuven debate whether interlingual homographs might have one or two representations in the bilingual lexicon, and whether cognates might have a special status. In constructing a static model, such distinctions must indeed take the form of *a priori* decisions. From a learning perspective, however, the system must find its own answer to this question, depending on the task that is being asked of it. If word form is to be identified, the representations need be no different for homographs/homophones, and increased frequency of exposure to the form across languages will likely improve response times and accuracy. But if the system is required to map from word form to semantics, interlingual homographs must have quite different representations, and likely must exploit language context information to drive this. Moreover, the degree to which each input form is associated with each meaning in the two languages is likely to have a material effect on the efficiency with which the system can generate each meaning; that is, the low frequency meaning of an interlingual homograph will be accessed more slowly (see Thomas, 1997 for a computational learning system exhibiting these characteristics).

In the same way that Dijkstra and van Heuven have built on existing static models of monolingual word recognition to account for phenomena within the bilingual domain, there now exists a host of monolingual *developmental* models addressing diverse phenomena. In the future, it will be possible to speak to many of the salient issues in bilingualism identified above, including issues of language acquisition, language dominance, language control and language loss (Thomas and van Heuven, in press).

The target article represents impressive progress in the development of theories of bilingual word recognition. In the future, one hopes that the models themselves will be permitted to develop too.

References

Thomas, M. S. C. (1997). Distributed representations and the bilingual lexicon: One store or two? In J. Bullinaria, D. Glasspool & G. Houghton (eds.), *Proceedings of the 4th Annual Neural Computation and Psychology Workshop*, pp. 240–253. Berlin: Springer.

Thomas, M. S. C. & van Heuven, W. J. B. (in press). Computational models of bilingual comprehension. In A. M. B. De Groot & J. F. Kroll (eds.), *The handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press.