 Bilingualism: Language and Cognition 5 (3), 2002, 202–205 © 2002 Cambridge University Press DOI: 10.1017/S1366728902223019

The BIA++: Extending the BIA+ to a dynamical distributed connectionist framework

MAUD JACQUET
ROBERT M. FRENCH
Quantitative Psychology and Cognitive Science (B32), Department of Psychology, University of Liège, 4000 Liège, Belgium.
E-mail: *mjacquet@ulb.ac.be, rfrench@ulg.ac.be*

Dijkstra and van Heuven have made an admirable attempt to develop a new model of bilingual memory, the BIA+. Their article presents a clear and well-reasoned theoretical justification of their model, followed by a description of their model. The BIA+ is, as the name implies, an extension of the Bilingual Interactive Activation (BIA) model (Dijkstra and van Heuven, 1998; Van Heuven, Dijkstra and Grainger, 1998; etc), which was itself an adaptation to bilingual memory of McClelland and Rumelhart's (1981) Interactive Activation model of monolingual memory.

The authors provide a wealth of background on bilingual memory cross-lingual interference and priming effects in what amounts to a veritable review of the literature in this area. The model that they propose is designed to account for many of these empirically observed effects. In what follows we will center our discussion around three points related to the design of their model. These issues are:

- the use of modular vs. distributed representations;
- learning;
- emergence and self-organization of lexical items.

We will discuss each of these points in turn.

Overview

From the earliest days of the development of computer models of human cognitive capacities, one of the most significant problems that hung over the entire endeavor was the problem of hand-coded representations, or the Problem of Representation, as it is sometimes called (see, for example, Chalmers, French and Hofstadter, 1992; Elman, 1995). Over and over again, exaggerated claims were made for programs that supposedly discovered new mathematical theorems, solved complex problems, made scientific discoveries, took creative leaps or discovered analogies. And, each time, upon closer inspection, one discovered that the real reason these programs were able to do anything at all was because they had been given input data that had been carefully tailored so that the desired solution was, if not necessarily inevitable, at least not too difficult to produce. This does not, of course, mean that one can *never* use hand-coded representations. Obviously, at some point, modelers have to make decisions about the form of the input data that will inevitably influence the output of their programs. So the real issue is this: To what extent does the fit of a program's output to empirical data depend on the way in which the raw data was "pre-processed" (i.e., by a human) before being presented to the program. This will become a crucial issue in the discussion that follows.

Modular vs. distributed representations

We are by no means eliminativists (e.g., Churchland, 1995) who decry any use of any modular structures in modeling cognition. The need for modularity in programs simulating complex cognitive abilities is obvious. On the other hand, the addition of arbitrary processing modules every time new data conflicts with a previous "boxological" structure is not acceptable either. We believe that the necessary modules should be able to be explained *as an emergent by-product of the architecture*. This allows modelers to make use of modularity when necessary but only insofar as it can be explained as an emergent product of lower level mechanisms.

This leads us to our first point of contention with the BIA+. This model is a far cry from the old Information Processing models of the pre-connectionist era. In keeping with the philosophy of McClelland and Rumelhart's 1981 IA model, the letter level in the BIA+ emerges from the primitive-feature level and the word level from the letter level. But then the problems begin. The letter level gives rise to the word level but, unlike the monolingual IA case, at the word level *the lexical items are separated into their respective languages*. We are never told how this might work. Of course, if the program already knows that a word belongs to a particular language, then the job is much easier. But that's cheating. Little children learning two languages are rarely told explicitly that "chien" is a word in French, while "bug" is a word in English. They simply hear the two words and they gradually *learn* where they belong. So, the question for Dijkstra and van Heuven is: How could these two "language modules" (Dutch and English) have come about in the first place? What would "gate" a particular word to incorporate it in the Dutch module and how would an English word be incorporated into the English module? And, most seriously of all, what do the "language nodes" associated with each module correspond to? In a standard spreading-activation account with distributed internal representations, these language nodes do not strike us as necessary. People know they are writing in English (as opposed to, say, French) because a coherent, highly-interconnected set of representations of "English" items are currently active in their brains, period. There is no need to have an additional, explicit "language" node continually reminding the writer that he or she is writing in English.

Further, in our reading of their description of the BIA+, explicit language nodes would seem to pose problems that would not arise in the distributed network account we are suggesting. Consider, for example, bilingual orthographic neighborhood effects (for example, see French and Ohnesorge, 1996, among many other papers on the subject). It has been shown that the recognition of a word, *W*, belonging to L1, the active language, can be significantly affected by a large orthographic neighborhood in L2, the non-active language. Now, according to the BIA+, how would this work? If the language node corresponding to L2 is not active, where did the inhibition of *W* come from? On the other hand, the authors might argue that the L2 node is partially active, due to the partial activation of the W's orthographic neighbors in L2, but that its activation is below conscious threshold. Nonetheless, if they claim that this sub-conscious activation level is sufficient to have an inhibitory effect on the word in L1, then what is the use of having a separate language node whose partial activation correlates perfectly with the activation of the orthographic neighbors? Why not eliminate the language node altogether, since the real work is being done, not by a partially activated language node but, rather, by the activated and partially activated elements *at the word level*?

Another point concerns the addition of a "semantics" module. It is certainly true that a word's semantic features should be taken into account in a word identification model. For this reason, the BIA+ is, indeed, more complete than the original BIA. However, this addition comes with a price. For example, exactly *how* is semantics to be included in the model? In particular, given that theirs is a localist account, does this mean that there is one meaning per node? In this case, how are the relations between and overlap among concepts implemented? In a distributed model with learning, one can at least suggest ways in which this could be accomplished, but how do the authors propose dealing with this problem in the non-learning, localist framework in which they are working?

The final point concerns BIA+'s "task module". While it is unquestionably appropriate to consider the role of task demands, is a new module really necessary to accomplish this? Instead, different sets of elementary processing operations could be activated depending on the task requirements. In this way, task processing becomes "distributed" task processing. Task demands would then influence the operation of the system, as they should, but there would be no need for a separate "task module". A good example of a memory system with integrated task specificity is Minerva II (Hintzman, 1984), in which there is no "task module" per se, instead a different combination of processes is triggered depending on the intended goal.

Learning

Our most significant problem with the BIA+ model is its lack of learning mechanisms. This means, inevitably, that all the model's inhibitory and excitatory connections between items at the same level and between levels must be set by hand, with all the potential problems that this can entail. One of the problems the authors run into with this approach is the necessity of encoding inter-lexical homographs twice, once in the L1 module and once in the L2 module (see section 2.5 of their article). Is this reasonable from an interactive activation perspective? At the letter and feature level, a Dutch-English homograph, such as "ROOM" (meaning "cream" in English) has rigorously identical characteristics. Then, suddenly, at the word level, there are *two* distinct lexical items, one for Dutch and one for English. The only justification for this particular hand-coding is that the *authors themselves* know that "ROOM" is an interlexical homograph and, in order for their model to produce the interference effects observed in Dutch-English bilinguals, this particular lexical item (as opposed to items like "COW") must be represented separately in the two separate language modules within their model. It strikes us that there should be a single lexical item, "ROOM", that would become active (along with the appropriate language-specific semantics) depending on the active language context, a situation that would occur if the model learned the items of the languages and the relationships among them on its own. One could imagine, for example, the statistical learning model of Christiansen, Allen and Seidenberg (1998) being extended to the case of bilingual language learning. In short, our view is that it would be far better for this type of model *to develop its own representations through learning*, as has often been proposed for the acquisition of a single language. Not only would this avoid potential problems with hand-coding but the connection strengths would be a function of directly experiencing the languages (in written form), rather than the work of programmers attempting to make the program work correctly.

As we discussed in the previous section, without learning, how do the authors propose to incorporate the semantic level? Will they hand-code all necessary semantic representations and their relationships to each other, each with varying strengths? This leads to the problem of scaling up to real languages. While it might be potentially possible for programmers to hand-code the relationships among a very small number of features, letters and lexical items of two different languages, it is much more difficult – almost certainly impossible – to do so as the number of items increases towards those required for a real language. If the program were able to learn to set its own synaptic weights, the problem of scaling-up associated with hand-coding could quite possibly be overcome.

Finally, we propose looking at the question of an integrated vs. separate lexical store. Indeed, the authors argue for integrated lexicons, but one has to wonder what, exactly, they mean by this. We will examine this question next, again in the context of learning.

Integrated vs. separate lexicons

An overwhelming body of recent experimental evidence in bilingual memory militates in favor of the idea of *unified bilingual lexicons* (Hermans, Bongaerts, de Bot and Schreuder, 1998; Costa and Caramazza, 1999; Costa, Caramazza and Sebastián-Gallès, 2000; Colomé, 2001), even though, functionally, bilingual memory gives the appearance of lexical separation. In the BIA+ this is achieved by having common feature and letter levels for the lexical items in both languages. Thereafter, at the word level, functional language separation seems for all the

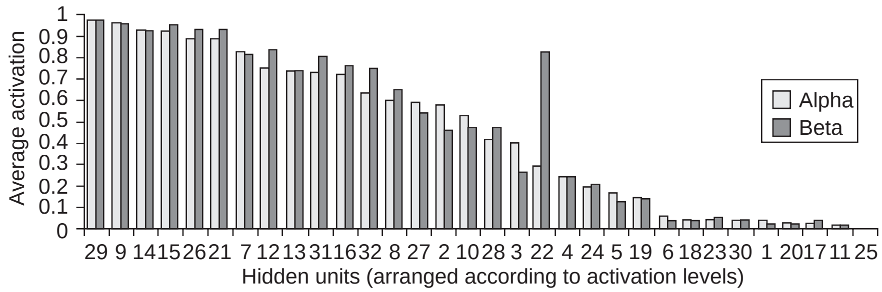

Figure 1. Overall overlap of the hidden-unit activations of all word repreentations for both languages.

world to have been implemented as *physical* lexical separation. In other words, at the word level, the BIA+ seems to posit separate lexicons. We suggest that this language-based word-level separation would be unnecessary in a distributed connectionist setting.

But how could a connectionist network's highly overlapping internal representations of the items in both languages *also* produce the high degree of functional language separation that we experience as a bilingual? French (1998) used a simple recurrent connectionist network (Elman, 1990) to show how this question might be approached from the standpoint of bottom-up bilingual language learning. This very simple model could arguably serve as a starting point for a statistical, emergent approach to bilingual language learning and storage. French produced two micro-languages, Alpha and Beta, and generated random series of sentences in each language, switching from one language to the other from time to time, as bilinguals do. There were no explicit markings of sentences or language switches. Below is a sample of text presented to the network:

BOY LIFTS TOY MAN SEES PEN MAN TOUCHES BOOK GIRL PUSHES BALL WOMAN TOUCHES TOY BOY PUSHES BOOK FEMME SOULEVE STYLO FILLE PREND STYLO GAR ON TOUCHE LIVRE FEMME POUSSE BALLON FILLE SOULEVE JOUET WOMAN PUSHES PEN BOY LIFTS BALL WOMAN TAKES BOOK

After exposing the network to some 60,000 items as above, he then examined the average hidden unit representations (i.e., activation patterns) over all words in each language and compared them. The amount of overall overlap between the representations of the words in the two languages was enormous, giving the impression of very highly overlapping lexicons (see Figure 1 above).

In contrast, however, when a cluster analysis of the internal representations of the words in each language was performed, French discovered that not only were all of the parts of speech in each language clustered correctly, but also the two languages themselves had developed the exactly appropriate clusters, as can be seen in Figure 2.

The simulation was later run with two larger micro-languages containing 768 words apiece and a distributed (i.e., non-localist) encoding for the words in each language. The same high-dimensional language separation emerged spontaneously. French was also able to show inter-lingual homograph priming effects with this extremely simple model. The point is that it is at least *possible* to develop bilingual memory models along these lines.

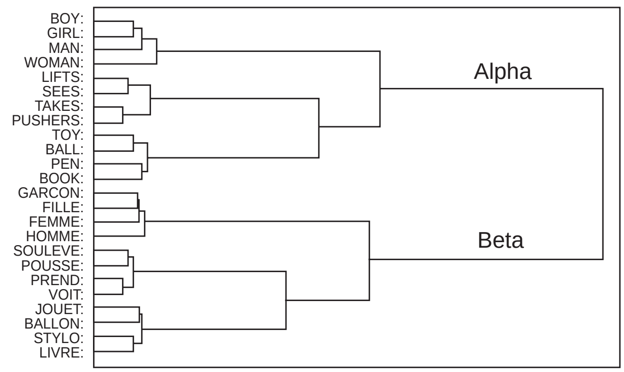

Figure 2. Cluster analysis of the hidden-unit representations for the words in both languages.

Conclusion

Dijkstra and van Heuven's BIA+ model is a well-thought out attempt to incorporate low-level mechanisms into a high-level model of bilingual memory that accounts for the extensive experimental data showing inter-lingual priming and interference effects. Just as the original Interactive Activation model paved the way for the connectionist revolution in cognitive psychology, we hope that the ground-breaking work of these authors will naturally evolve towards broader-based distributed connectionist network models and related dynamical models of bilingual memory, capable of learning and being able to incorporate both the bottom-up and the top-down processing that we know to be an integral part of bilingual language processing.

Acknowledgments

This work was funded by a grant from the European Commission HPRN-CT-2000-00065 to R. M. French.

References

Chalmers, D. J., French, R. M. & Hofstadter, D. R. (1992). High-level perception, representation, and analogy: A critique of Artificial Intelligence methodology. *Journal of Experimental and Theoretical and Artificial Intelligence*, 4, 185–211.

Christiansen, M. H., Allen, J. & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 221–268.

Churchland, P. (1995). *The engine of reason, the seat of the soul*. Cambridge, MA: MIT Press.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *Journal of Memory and Language*, 45, 721–736.

Costa, A. & Caramazza, A. (1999). Is lexical selection in bilinguals language specific? Further evidence from Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 2, 231–244.

Costa, A., Caramazza, A. & Sebastián-Gallès, N. (2000). The cognate facilitation effect: implications for model of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296.

Dijkstra, A. (Ton) & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (eds.), *Localist connectionist approaches to human cognition*, pp. 189–225. Mahwah, NJ: Erlbaum.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.

Elman, J. (1995). Language as a dynamical system. In R. Port & T. van Gelder (eds.), *Mind as motion: Explorations in the dynamics of cognition*, pp. 195–223. Cambridge, MA: MIT Press.

French, R. M. (1998). A Simple Recurrent Network model of bilingual memory. In *Proceedings of the Twentieth Annual Cognitive Science Society Conference*, pp. 368–373. Hillsdale, NJ: Lawrence Earlbaum.

French, R. M. & Ohnesorge, C. (1996). Using interlexical nonwords to support an interactive-activation model of bilingual memory. *Proceedings of the Eighteenth Annual Cognitive Science Society Conference*, pp. 318–323. Hillsdale, NJ: Lawrence Erlbaum.

Hermans, D., Bongaerts, T., de Bot, K. & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1, 213–230.

Hintzman, D. L. (1984). Minerva II: A simulation model of human memory. *Behavior Research Methods, Instrument & Computers*, 16, 96–101.

McClelland, J. & Rumelhart, D. (1981). An Interactive-Activation model of context effects in letter perception, part 1: An account of basic findings. *Psychological Review*, 88, 375–405.

Van Heuven, W. J. B., Dijkstra, A. (Ton) & Grainger, J. (1998). Orthographic neighbourhood effects in bilingual recognition. *Journal of Memory and Language*, 39, 458–483.