Bilingualism and the single route/dual route debate

Conference Presentation

Thomas, M.S.C. (1998) Bilingualism and the single route/dual route debate
The 20th Annual Conference of the Cognitive Science Society, 1-4 August 1998, Madison, U.S. - pp.1061-1066

Bilingualism and the Single route / Dual route debate.

Michael S. C. Thomas,
Department of Psychology,
King Alfred's College,
Sparkford Road,
Winchester, SO22 4NR, UK.
thomas@dragon.psych.ox.ac.uk

Abstract

The debate between single and dual route accounts of cognitive processes has been generated predominantly by the application of connectionist modeling techniques to two areas of psycholinguistics. This paper draws an analogy between this debate and bilingual language processing. A prominent question within bilingual word recognition is whether the bilingual has functionally separate lexicons for each language, or a single system able to recognize the words in both languages. Empirical evidence has been taken to support a model which includes *two* separate lexicons working in parallel (Smith, 1991; Gerard and Scarborough, 1989). However, a range of interference effects has been found between the bilingual's two sets of lexical knowledge (Thomas, 1997a). Connectionist models have been put forward which suggest that a *single* representational resource may deal with these data, so long as words are coded according to language membership (Thomas, 1997a, 1997b, Dijkstra and van Heuven, 1998). This paper discusses the criteria which might be used to differentiate single route and dual route models. An empirical study is introduced to address one of these criteria, *parallel access*, with regard to bilingual word recognition. The study fails to find support for the dual route model.

Introduction

Two well known connectionist models have questioned the necessity of dual route processing accounts within the psychology of language. In inflectional morphology, connectionist models of the formation of the past tense have suggested that a single network is sufficient to learn both rule-based past tense formations as well as idiosyncratic formations and family clusters (e.g. Plunkett and Marchman, 1991). The naming and word recognition models of Seidenberg and McClelland (1989) and latterly Plaut, McClelland, Seidenberg, and Patterson (1996) have suggested that a single network is sufficient to learn both a rule-based method of constructing pronunciations of novel word strings, as well as the various irregular and idiosyncratic pronunciations of English monosyllabic words.

The single route / dual route debate has often revolved around whether it is necessary to postulate *a priori* rule-based processing mechanisms to underlie various aspects of cognition; because rules have exceptions, a rule-based processor must be accompanied by an exception mechanism. However, the debate can also be seen simply as a question of whether we need to postulate one or two functional processing mechanisms to underlie a given cognitive ability.

In this paper, I will propose that the one or two lexicons debate within bilingual word recognition is analogous to this question. After outlining this debate, I will suggest four criteria which may be used to distinguish between single and dual route models of a given cognitive ability. Lastly, I will introduce empirical evidence from a bilingual priming study which addresses one of these criteria. This study produces results which appear inconsistent with a dual route account of bilingual word recognition.

Bilingual word recognition

A primary issue when investigating the bilingual language processing system is its relation to that of the monolingual system. At one extreme, one could envisage that the bilingual's two languages serve merely as subdivisions within a single system. At the other extreme, one might envisage an entirely separate system for each of the bilingual's two languages.

In terms of bilingual word recognition, the traditional account takes an intermediate position. Assuming that knowledge about words can be separated into knowledge about word meanings and knowledge about word forms, the traditional account postulates a common semantic system for both languages, but separate systems to recognize the word forms of each language (Smith, 1991). Thus in a task like visual lexical decision, priming paradigms show that short term *semantic* priming occurs between the bilingual's languages as well as within them (e.g. Chen and Ng, 1989); but long term *lexical* priming between the first and second presentations of a word is only found for repetitions within a language, not between translation equivalents in different languages (e.g. Kirsner, Smith, Lockhart, King, and Jain, 1984).

For many pairs of languages, there are word forms that exist in both languages. These are known as homographs. *Non-cognate homographs* on the other hand are words which have the same form but a different meaning in each language (e.g. the French words MAIN and FIN mean 'hand' and 'end' in English). Since these words have a different meaning in each language, they often have a different frequency of occurrence. When bilinguals are required to recognize the visual forms of non-cognate homographs, they do so according to their frequency of occurrence within each language (French and Ohnesorge, 1995; Gerard and

Scarborough, 1989); that is, the *same* word form is recognized quickly in the language context where it has a high frequency of occurrence, and slowly in the language context where it has a low frequency of occurrence. The behavior of these words appears to be unaffected by the presence of the same word form in the other language. This is taken as additional evidence of independent lexicons for each language. And indeed, as if to confirm that a different representation underlies each version of the visual word form, presentation of a non-cognate homograph in one language context does not facilitate later recognition of the word form in the other language context (Thomas, 1997a).

In broad metaphorical terms, then, the bilingual word recognition system comprises separate mental dictionaries to look up the word forms in each language, but dictionaries pointing to a common set of meanings. In order to retrieve the meaning for a given input, there are *dual* lexical routes.

Two criticisms of the dual route model of bilingual word recognition

There are two kinds of criticisms that have been made of this dual route model.

The first is based on the assumption that morphology is the principle which guides organization of the bilingual's lexical representations. On the basis that words of different languages often sound or look different, morphology alone could be sufficient to produce the apparent effects of independence in a single undifferentiated system. Kirsner, Lalor, and Hird (1993, p. 228) comment: "morphology may be the critical feature for lexical organization, providing the pegs around which clusters of words are organized, regardless of language". In this kind of system, one would not expect long term lexical priming effects between translation equivalents (such as DOG and CHIEN in English and French) when they are orthographically different, in the same way that orthographically different synonyms within a single language do not prime each other over long intervals (e.g. DOG and HOUND). However, the morphological approach cannot account for why non-cognate homographs, with the same orthographic form in each language, nevertheless show different recognition latencies depending on the language context in which they are presented.

The second criticism of the dual route model of bilingual word recognition arises because, despite evidence for the independence of lexical representations, there is nevertheless a wide range of cross-language interference effects found when stimuli exhibit some degree of similarity between the two languages. That is, recognition of a word in one language may be facilitated or inhibited if it resembles a word in the other language. Thus Klein and Doctor (1992) found that non-cognate homographs (MAIN, FIN) were recognized more slowly than matched cognate homographs (words which have the same form and meaning each language, such as TRAIN in English and French); Cristoffanini, Kirsner, and Milech (1986) and Gerard and Scarborough (1989) found that cognate homographs in a bilingual's weaker language were recognized more quickly than words of comparable frequency in that language; Beauvillain (1992) found that when operating in a single language context, bilingual subjects recognized words more quickly if the words possessed orthographic patterns specific to that language, rather than common between the two languages; and Grainger and Dijkstra (1992) found that the number of orthographic neighbors in the non-active language could affect lexical decision times for words in the active language. If a word had more orthographic neighbors in the other language than its own, it was recognized more slowly.

Evidence of cross-language interference effects has led researchers to suggest that the bilingual has a single set of lexical representations, but representations that are structured by a combination of language membership and morphology (Grainger and Dijkstra, 1992; Thomas, 1997a, 1997b; Dijkstra and van Heuven, 1998). This approach has been supported by computational modeling within the connectionist tradition.

Two single route connectionist models

Single route connectionist models have been put forward within both the interactive activation and distributed frameworks.

Grainger and Dijkstra (1992) proposed that a bilingual version of the interactive activation model could account for the cross-language neighborhood effects found in their lexical decision study. Dijkstra and van Heuven (1998) implemented a version of this model and simulated a number of results from experiments with Dutch-English bilinguals demonstrating inhibitory effects of cross-language neighborhoods. The bilingual interactive activation model (BIAM) contains an integrated lexicon at the word unit level, with units for the words in both languages. All word units compete with each other regardless of the language to which they belong. In addition, an extra representational layer is added which contains two "language nodes", one representing each language. The units for the words in each language have excitatory connections to their respective language node. The language nodes in turn have inhibitory connections to the word units of the *other* language. During recognition, excitation of the word units of one language will inhibit the activity of the word units in the other language, via their language node. In this model, there is a single lexicon differentiated only by the addition of information concerning language membership.

Thomas (1997a, 1997b) has put forward a distributed model of bilingual word recognition, based on an extension to Seidenberg and McClelland's (1989) framework for monolingual word recognition and naming. The Bilingual Single Network (BSN) model postulates a single representational resource to perform each of the mappings between orthography, phonology, and semantics for both of the languages. However, words are tagged for language membership as part of their input and output representations.

The BSN model can account both for empirical data suggesting the independence of lexical representations (French and Ohnesorge, 1995; Kirsner et al, 1984), and also for between language interference effects (Cristoffanini, Kirsner, and Milech, 1986; Gerard and Scarborough, 1989; Klein and Doctor, 1992). For example, in generating the semantic output for a non-cognate homograph like FIN in

one of the languages, the model shows a frequency response determined by how often the model is trained on the orthographic code in association with the appropriate language code. Thus FIN would show a within language frequency response – a behavior taken to suggest independence. However, when attempting to learn non-cognate homographs, the network is presented with two mappings that are very similar at input but dissimilar at output (i.e. word form + language code *a* => meaning *a*; same word form + language code *b* => meaning *b*). Such inconsistent mappings are less well learnt than consistent mappings where similar inputs generate similar outputs. Thus FIN would be recognized more slowly and less accurately than matched cognate homographs – evidence of interference between the two languages at the lexical level.

The dual route model strikes back

We have seen, then, two models which account for the cross-language interference effects apparently incompatible with the traditional dual route account. However, an alternative avenue is open to the dual route theorist which may allow them to deal with these problematic effects. The dual route theorist may propose that interference effects at the lexical level arise merely from the competition of two independent routes. Thus Sharwood Smith (1991, p. 16-17) comments:

> *…evidence from crosslinguistic influence in language loss and language acquisition… would suggest that L1 and L2 knowledge is not rigidly separated and that knowledge in one system can indeed affect the shape of knowledge in the other. What this precisely means with respect to the idea of autonomy or encapsulation is, of course, theory-dependent. Since knowledge is only "visible" in performance, any crosslinguistic influence between L1 and L2 could still be the on-line accessing of what are still separate systems, as in code-switching. In other words, crosslinguistic influence could be a purely real-time control phenomenon.*

In this version of the dual route model, the lexical representations are still functionally independent, but now any evidence of interference between the two routes is taken to arise because the routes work in parallel and both contribute (for good or bad) to the overall response. For the dual route theorist, evidence of between language interference is evidence of a failure to properly control independent processing mechanisms, not evidence that there is a single representational resource.

In light of this response, it is no longer clear how we may distinguish between single route and dual route accounts.

Four criteria to distinguish single route and dual route models

The dual route theorist's response to evidence of between route interference effects is found not only within bilingual word recognition, but also within the past tense debate and the monolingual word recognition debate.

In the past tense debate, the claim made for the single route model is that a single processing mechanism can produce the U-shaped learning characteristic of children (e.g. Plunkett and Marchman, 1991). A single connectionist network demonstrates over-generalizations, whereby irregular past tense formations (go => went) are temporarily treated as regular verbs (go => goed), before additional training draws the model towards overall competence. However, the traditional dual route account does not take such interference as evidence of a single mechanism. It suggests that the interference effect arises because two independent mechanisms (one for regular past tense formations, one for irregular past tense formations) are generating conflicting responses, and, at a certain stage in development, there is a failure to control the mechanisms. For a brief period, the irregular route fails to 'block' or override the regular route (e.g. Pinker, 1994).

With regard to monolingual naming, the traditional account suggests that two processing mechanisms are necessary to generate the pronunciation of a written word. The word may be recognized and its stored pronunciation retrieved (lexical route). Or the pronunciation may be constructed on the basis of the letters that make up the word, using a set of rules which convert letters into sounds (GPC route)[1]. The single route model (Plaut et al, 1996) suggests that a single processing mechanism may achieve both of these functions, that is, store the pronunciations for written words, and construct the pronunciations of letter strings (at least as well as humans can). In addition, the single route model can account for *consistency* effects. In the dual route model, pronunciations for letter strings such as ZAID and ZAKE are constructed in the GPC route. Over in the word store, it turns out that when -AID arises within known words, it does not generate consistent pronunciations: MAID, PAID, BRAID etc. but also SAID. On the other hand, when -AKE appears in words, this segment is always pronounced the same: BAKE, TAKE, WAKE, SHAKE etc. The interesting thing is that human subjects reliably show longer latencies when naming strings like ZAID than they do when naming strings like ZAKE (Glushko, 1979). Presumably, this is because the construction of the pronunciation is influenced by the fact that, over in the word store, there is a difference of opinion about -AID words but not about -AKE words.

For the single route account, this interference effect is to be expected, since both functions are realized over a single representational resource. In the dual route account, construction of pronunciations and retrieval of stored pronunciations occur in independent mechanisms. So where does the interference effect come from? By now we can guess the response: interference is the result of the parallel function of two independent mechanisms where both compete to generate a response. Coltheart, Curtis, Atkins, and Haller, (1993) continue "…because competition increases naming latency, this is a way in which a pure dual-route model can explain consistency effects on nonword (and word) naming latency" (p. 605).

Here then are three examples of single route / dual route debates in which both sides can claim to provide equally valid explanations of the data. How can we distinguish

[1] There is a third route via semantics which both 'dual route' and 'single route' accounts agree upon.

between these accounts? In the next section, I will discuss four possible criteria:

1. Are there between route interference effects in the empirical data?
2. Can the two routes be dissociated?
3. Can the two routes be acquired independently?
4. Is there evidence of parallel access?

1. Are there between route interference effects in the empirical data?

Such interference effects might be taken as support for the single route account. It seems a reasonable starting point to require that, if a model wants to postulate two independent mechanisms, then during the normal functioning of one mechanism, there should be no interference from the other. Consistency effects might be seen as a prototypical example of a between route interference effect supporting a single route model. Empirical evidence of consistency effects has been found not only in naming (Glushko, 1979), but also in past tense formation (Seidenberg and Bruck, 1990), and in bilingual word recognition (Grainger and Dijkstra, 1992). However, in the light of the preceding discussion, this criterion may well prove inconclusive: dual route models have a way to account for between route effects. One might merely note that given that single route models have provided computational accounts of consistency data, the onus is on dual route models to provide a computational account of how the activity of the separate routes may be glued together to produce the requisite interference patterns. Thus far, such dual route computational accounts have been thin on the ground.

2. Can the two routes be dissociated?

A second possible criterion to distinguish single route and dual route models would be the presence of dissociations between the two routes. Evidence of dissociations might support a dual route account. Pinker (1991) and Coltheart et al (1993) have pointed to cognitive neuropsychological evidence of double dissociations between regular / irregular routes in past tense, and lexical / GPC routes in naming. In the bilingual case, there is some evidence of differential impairment of languages, although this generally appears to be associated with levels of proficiency and histories of acquisition (see Paradis, 1996; Vaid and Hall, 1991).

For this criterion, the onus is on single route models to show how damage to a single mechanism may lead to separate impairment to one or other function (be it regular / irregular, lexical / GPC, or L1 / L2). Some work has been carried out in the lesioning of distributed networks, which has shown both the potential difficulties in this enterprise (e.g. Bullinaria and Chater, 1995) but also the possibility of generating double dissociations in such networks (Plaut, 1995).

3. Can the two routes be acquired independently?

Evidence that routes can be acquired independently might support a dual route account. This criterion is not usefully applied to the naming or past tense debates, where both routes are acquired more or less simultaneously. In the bilingual case, however, both simultaneous and sequential acquisition of languages are possible. Moreover, sequential acquisition of two languages produces a functional word recognition system not markedly different from the simultaneously acquired system (Magiste, 1984; Potter, So, von Eckhardt, and Feldman, 1984).

This places constraints on a single route model, particularly one using distributed representations. Although the bilingual interactive activation model aims only to provide an account of final state performance, the distributed network model would hope to include an account of acquisition. However, it is well known that distributed networks which are trained sequentially on different tasks can experience interference effects between the first and second training set (see e.g. McCloskey and Cohen, 1989). In bilingual terms, this would predict damage to L1 during acquisition of L2. This does not appear to be a prominent characteristic of second language acquisition (although it is not clear that anyone has specifically looked for such an effect, for instance under conditions of intense L2 acquisition).

Given the apparent independence of acquisition, the onus here is on the single route model to show how a single mechanism can integrate sequentially acquired functions. A distributed model of bilingual word recognition employing a single representational resource for both languages would need to address the sequential learning problem. Some (limited) progress has been made on this issue (see e.g. McClelland, McNaughton, and O'Reilly, 1995), but much remains to be done. It is worth noting that this issue is not limited to bilingual models. Even a distributed monolingual model of word recognition needs to establish how vocabulary expansion can progress in a single mechanism without disruption of pre-stored distributed knowledge.

4. Is there evidence of parallel access?

To account for between route interference effects, dual route models employ the notion of parallel access. Both routes are working at once, and their combination causes interference. It is important to note that in a single route model, such parallel access is *not possible*: there is only a single representational system, which may take on different forms in different contexts (regular vs. irregular, lexical vs. GPC, L1 vs. L2). Evidence which demonstrates that parallel access *must have occurred* would be evidence for a dual route model.

We have to be reasonably careful, here, in what we accept as pertinent empirical evidence. For example, monolingual research on how the meanings of ambiguous words are retrieved has suggested that during word recognition, both meanings of an ambiguous word are initially accessed, but that only one meaning remains active after a delay (Seidenberg, Tanenhaus, Leiman, and Bienkowski, 1982, Experiment 1). We could take this to mean that there are two independent representations of word meaning accessed in parallel, one of which is subsequently suppressed. However, using a recurrent connectionist network, Kawamoto (1993) demonstrated that the empirical results taken to suggest such parallel access could be simulated by a distributed system

where both meanings were represented over the *same* representational resource.

Evidence for parallel access might work as follows. A subject is required to perform some task which should only invoke activity in one of the 'routes'. In the dual route account, this task must have also activated the other route (to potentially cause interference effects). We then attempt to search for some residue of the activity in the second route. One candidate residue would be a priming effect. As an illustration of this approach, I will briefly discuss an empirical study carried out to evaluate the dual route model of bilingual word recognition (see Thomas, 1997a, for further details).

Empirical study: a test of the parallel access criterion in bilingual word recognition

In the language exclusive bilingual lexical decision task, one of the subject's languages is defined as active. The subject is then required to separate words from plausible letter strings, but to respond positively only to stimuli that are words in the currently active language. If a letter string appears which is a word in the currently *inactive* language, this should be rejected as if it were a non-word.

In the standard monolingual lexical decision task, if stimuli are repeated after a gap of 20 or so trials, responses to words are accelerated, but plausible letter strings gain little from repetition (Scarborough, Cortese, and Scarborough, 1977). This priming effect is assumed to arise because accessing the word representation on the first occasion makes subsequent access easier. Plausible letter strings on the other hand, are assumed not to have a representation in the lexicon that can be so facilitated.

In the current experiment, stimuli were blocked into sets of 50. The active language alternated between each block. Stimuli were repeated either within a block (i.e. within a language) or across blocks (between the languages).

Let us take a given word from L1; we will focus on the situation where the second presentation of the L1 word occurs in a block where L1 is the active language, and thus where the stimulus should be accepted as a word. The first presentation of this word may either have occurred when the subject was also responding according to L1; or it may have occurred when the subject was responding according to L2, where (assuming the word form does not also exist in L2) it should have been rejected as a nonword. Now according to the dual route model, all word recognition should activate both routes in parallel. So long as accessing a word's representation causes later access to be facilitated, it shouldn't matter whether the word was first viewed in an L1 context or an L2 context. Parallel access predicts that the priming effect should be the same.

An experiment was run according to this design using 16 undergraduate English-French bilinguals, whose native language was English, but who had studied French at least until age 18. Half of the subjects were also studying French at undergraduate level[2]. In order to ensure that subjects did not develop a strategy based on the orthographic characteristics of stimuli, all stimuli formed orthographically legal strings in both languages.

Table 1 shows the results for repetitions of English and French words either within language or between languages. For English words, in the subjects' dominant language, within language repetition caused the expected priming effect (56 msec, $p=0.031$). On the other hand, preview of an English word from a French language context did not cause any facilitation at all ($p>0.5$). For French, the subjects' weaker language, within language repetition again caused the expected priming effect (95 msec, $p<0.001$). On this occasion, preview of a French word from an English language context did cause a significant facilitation effect (53 msec, $p=0.035$), but a smaller effect than the within language repetition. Overall, between language priming effects were significantly smaller than within language effects (simple factorial anova, $F(1,3)=4.39$, $p=0.036$).

Table 1: Within and between language word repetition.

English words.

Presentation context		Response Time (msecs)		Diff.	t-test
Prime	Target	Unprimed	Primed		p-value
English	*English*	*630*	*574*	*56*	*0.031*
French	*English*	*630*	*646*	*-16*	*0.589*

French words

Presentation context		Response Time (msecs)		Diff.	t-test
Prime	Target	Unprimed	Primed		p-value
French	*French*	*669*	*574*	*95*	*0.000*
English	*French*	*669*	*617*	*53*	*0.035*

In terms of the dual route model, the simplest interpretation of these results is that recognition in L1 is associated with some limited parallel access of the L2 lexicon, but that recognition in L2 *is not associated with any parallel access of L1*. However, this interpretation is not consistent with the explanation the dual route model gives for an interference effect we encountered earlier. Cristoffanini et al and Gerard and Scarborough found that low frequency L2 cognate homographs experienced a facilitation effect in the lexical decision task compared to matched words. For the dual route model, this would imply that, that since the word exists in both routes, during recognition in L2, parallel access of L1 is indeed required to produced the facilitation effect.

Conclusion

Connectionist models of bilingual word recognition are an exciting prospect. It is clear from the criteria reviewed here, that single route models in this domain still must overcome a number of hurdles to establish superiority over traditional dual route accounts. Nevertheless, with regard to one criterion, we saw empirical evidence against a dual route account of between route interference effects, effects which single route models can handle with ease.

[2] These subjects may more properly be referred to as second language learners, but see discussion of criterion 3.

References

Beauvillain, C. (1992). Orthographic and Lexical Constraints in Bilingual Word Recognition. In R. J. Harris (Ed.) *Cognitive Processing in Bilinguals.* Elsevier Science Publishers.

Bullinaria, J. A. and Chater, N. (1995). Connectionist modeling: Implications for Cognitive Neuropsychology. *Language and Cognitive Processes, 10 (3/4),* 227-264.

Chen, H.-C. and Ng, M. L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory and Cognition, 17,* 454-462.

Coltheart, M., Curtis, B. Atkins, P., and Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches*, Psychological Review, 100,* 589-608.

Cristoffanini, P., Kirsner, K., and Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *Quarterly Journal of Experimental Psychology, 38A,* 367-393.

Dijkstra, T. & van Heuven, W. J. B. (1998). The BIA-model and bilingual word recognition. In J. Grainger and A. Jacobs (Eds.) *Localist Connectionist Approaches to Human Cognition.* LEA.

French, R. M. and Ohnesorge, C. (1995). Using Non-cognate interlexical homographs to study bilingual memory organization. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society.* LEA. Pp. 31-36.

Gerard, L. D. and Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 305-315.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674-691.

Grainger, J. and Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.) *Cognitive Processing in Bilinguals.* Elsevier Science Publishers. Pp. 207-220.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32,* 474-516.

Kirsner, K., Lalor, E., and Hird, K. (1993). The bilingual lexicon: Exercise, meaning, and morphology. In R. Schreuder and B. Weltens (Eds.) *The Bilingual Lexicon.* Amsterdam/ Philadelphia: John Benjamins. Pp. 215-248.

Kirsner, K., Smith, M. C., Lockhart, R. L. S., King, M. L., and Jain M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior, 23,* 519-539.

Klein, D. and Doctor, E. A. (1992). Homography and polysemy as factors in bilingual word recognition. *South African Journal of Psychology, 22(1),* 10-16.

Magiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol. 10(2),* 304-315.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. G. (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.

McCloskey, M. and Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.) *The Psychology of Learning and Motivation,* Vol. 24, New York: Academic Press.

Paradis, M. (1996). Selective deficit in one language is not a demonstration of different anatomical representation. Comments on Gomez-Tortosa et al. (1995). *Brain and Language, 54,* 170-173.

Pinker, S. (1991). Rules of language. *Science*, *253,* 530-535.

Pinker, S. (1994). *The Language Instinct.* Penguin books.

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from Connectionist Neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17(2),* 291-321.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., and Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56-115.

Plunkett, K. and Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition, 38,* 1-60.

Potter, M. C., So, K-F., von Eckhardt, B., and Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23-38.

Scarborough, D. L., Cortese, C., and Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1-17.

Seidenberg, M. S. and Bruck, M. (1990). Consistency effects in the generation of past tense morphology. Paper presented at the thirty-first meeting of the Psychonomic Society (November: New Orleans).

Seidenberg M. S. and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., and Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14,* 489-537.

Sharwood Smith, M. (1991). Language modules and bilingual processing. In: E. Bialystok*, Language processing in bilingual children.* Cambridge University Press.

Smith, M. C. (1991). On the recruitment of semantic information for word fragment completion: Evidence from bilingual priming. *Journal of Verbal Learning and Verbal Behavior, 17,* 234-244.

Thomas, M. S. C. (1997a). Connectionist networks and knowledge representation: The case of bilingual lexical processing. D. Phil thesis, University of Oxford, UK.

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

Vaid, J. & Hall, D. G. (1991). Neuropsychological Perspectives in Bilingualism: Right, Left, and Center. In A. G. Reynolds and W. E. Lambert (Eds.) *Bilingualism, Multiculturalism, and second language learning.* Erlbaum.