Modeling the Bilingual Lexicon
of an Individual Subject

Risto Miikkulainen[1] and Swathi Kiran[2]

[1] The University of Texas at Austin, Austin, TX 78712, USA
[2] Boston University, Boston, MA 02215, USA

Abstract. Lexicon is a central component in any language processing system, whether human or artificial. Recent empirical evidence suggests that a multilingual lexicon consists of a single component representing word meanings, and separate component for the symbols in each language. These components can be modeled as self-organizing maps, with associative connections between them implementing comprehension and production. Computational experiments in this paper show that such a model can trained to match the proficiency and age of acquisition of particular bilingual individuals. In the future, it may be possible to use such models to predict the effect of rehabilitation of bilingual aphasia, resulting in more effective treatments.

Key words: Lexicon, semantics, speech, language acquisition

1   Introduction

The mental lexicon, i.e. the storage of word forms and their associated meanings, is a major component of language processing. It is also one that is perhaps the best understood in terms of computational modeling. Partly the reason is that abundant data exists about how the lexicon develops, how it is organized, how it functions, and how it breaks down in dyslexia and aphasia; partly the reason is that lexical processes are rather modular and therefore tend to be amenable to computational theories.

Although the physiological implementation and even the location of the lexicon in the brain is still open to some debate, there is evidence from MRI and electrophysiology that the lexicon may be laid out as a map, or multiple maps [16]. As a result, self-organizing map (SOM) models are a natural way to model the lexicon. SOM models have been developed to understand e.g. how ambiguity is processed by the lexicon, and how it breaks down in dyslexia and aphasia [12,13], and how the lexicon is acquired during development [11]. The SOM-based lexicon has also turned out useful as a component in larger systems of natural language processing [12].

Given that the majority of the world's population is bilingual (or multilingual) [17], an important extension of the SOM-based lexicon models, as well as lexical research in general, is to account for lexica with multiple languages. A theoretically based approach is developed in this paper, with the eventual aim of using the model to help rehabilitate bilingual patients with aphasia.

More specifically, a bilingual lexicon model is built for individuals with different proficiencies in the two languages, and different age at which the two languages were

acquired. Based on recent theoretical models, a single semantic map is used, with separate maps for the words in the two languages, and six separate sets of connections between the maps. Frequency of exposure, as opposed to age, is shown to primarily determine the proficiency, with age affecting the way the maps are organized. In the future, the model can be used to study how the multilingual lexicon breaks down with damage and how it can be rehabilitated. By matching the proficiency and age of acquisition with those of patients suffering from aphasia, it may be possible to use the model to derive most effective treatment regimes for them individually.

2 Bilingual Lexical Processing

Current theoretical models of the bilingual lexicon generally agree that bilingual individuals have a shared semantic (or conceptual) system and that there are separate lexical representations of the two languages. However, the models differ on how the lexica interact with the semantic system and with each other.

The concept-mediation model [15] (Fig. 1a), proposes that both the first (L1) and the second-language lexica directly access concepts. In contrast, the word-association model assumes that second-language words (L2) gain access to concepts only through first-language mediation (Fig. 1b). Empirical evidence [8] suggests that the word association model is appropriate for low-proficiency bilinguals and concept mediation model for high-proficiency bilinguals. As an explanation, De Groot [5] proposed the mixed model (Fig. 1c), where the lexica of a bilingual individual are directly connected to each other as well as indirectly (by way of a shared semantic representation). This model was further revised with asymmetry by Kroll & Stewart [9] (Fig. 1d). The associations from L2 to L1 are assumed to be stronger than those from L1 to L2, and the links between the semantic system and L1 are assumed to be stronger than those between the semantic system and L2.

A second important issue is whether activation of the semantic system spreads to both lexica or only within that of the language being used. The prevailing theory suggests that lexical access is target-language nonspecific [2], although, this view is controversial [3]. A third issue is the extent to which proficiency in the two languages and the age at which they are acquired (AoA) affect lexical access. There is evidence that language proficiency, and not AoA, primarily determines the nature of semantic processing [6]. For instance, Li and Farkas [10] showed that novice bilinguals have a fuzzier representation of semantics and phonology than proficient bilinguals.

Adopting the asymmetric mixed model of Fig. 1d, this paper will systematically examine the extent to which language proficiency and AoA influence activation of targets in the lexicon. The work will form a foundation for studying damage and rehabilitation of aphasia in bilingual patients in the future.

3 Computational Models of the Lexical System

Artificial neural networks have been used to model various aspects of the lexical system for over two decades. Most of the models aim to explain lexical processing with low-level mechanisms, focusing on the timing of the process as well as on certain types of
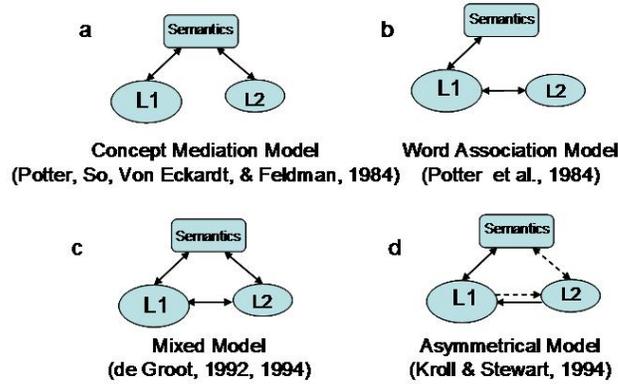
Fig. 1. Theoretical models of the bilingual lexicon. All four theories posit a common semantic system with language specific representations in L1 and L2. The most recent theory (d) includes connections between all maps, with connections of the most dominant language (L1 in this figure) stronger than the others (solid lines indicate strong connections and dashed lines weak connections). This theory is used as the starting point for the computational model.

performance errors and deficits. They are primarily process models, detached from the physical structures, and designed as controlled demonstrations of how disambiguation and production could be carried out in the lexical system [4,7,14].

One exception is the DISLEX model by Miikkulainen [12,13], which was further developed as DEVLEX by Li and colleagues [11]. Its organization is modeled after the cortical maps that underlie many perceptual processes and may also be the substrate for the lexical system in the brain [16]. DISLEX consists of two self-organizing maps, one for lexical symbols and the other for word meanings, as well as associative connections between them (Fig. 2). The lexical map is a layout of the orthographic or phonetic symbols in the language (orthography is used in the examples in this paper). It is a two-dimensional array of computational units, or neurons, trained to represent the symbols using the self-organizing map method. The symbols are vectors of gray-scale values [0..1], representing the orthographic features of the word. Each word is represented by a "blurry bitmap", i.e. a coarse visual image of the word as a series of letters.

During training, such vectors are presented to the map one at a time, and each unit computes the Euclidean distance d between its weight vector w and the symbol representation v:

$$d = \sqrt{\sum_k (w_k - v_k)^2}. \tag{1}$$

The unit with the smallest distance (unit (r, s)) is then found, and the weights of that unit and those in its neighborhood (units (i, j)) are adapted towards the input vector:

$$w'_{k,ij} = w_{k,ij} + \alpha(v_k - w_{k,ij})h_{rs,ij}, \tag{2}$$

where $h_{rs,ij}$ is a Gaussian function defining and α is the learning rate. This process has two effects: the weight vectors become representations of the symbol vectors, and the
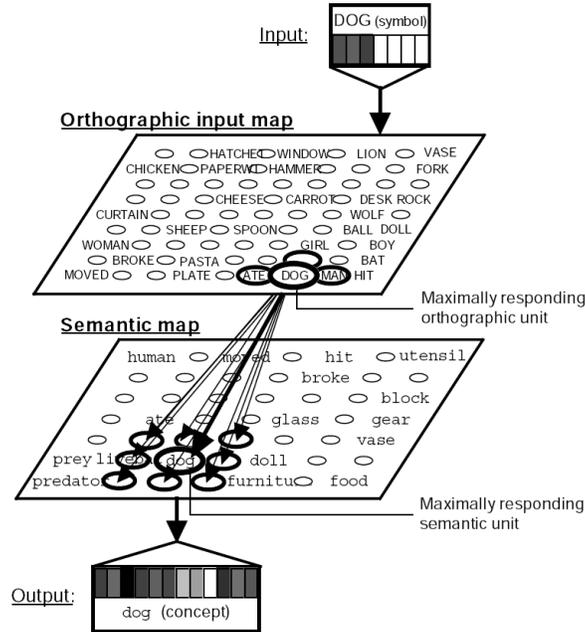
Fig. 2. A single-language DISLEX model and representations. The orthographic input symbol `DOG` is translated into the semantic concept `dog` in this example. The representations are vectors of gray-scale values between 0 and 1, stored in the weights of the map units. The size of the unit on the map indicates how strongly it responds. Only a few strongest associative connections of the orthographic input unit `DOG` (and only that unit) are shown.

neighboring weight vectors become similar. Over several presentations of each lexical symbol, the array of units then learns to represent the space of symbols in the language.

The semantic map is organized in a similar manner. The input vectors to this map represent semantic meanings of words. Each component of the vector represents a semantic microfeature (learned automatically based on word cooccurrence [12], and the vectors as a whole represent similarities between word meanings. The self-organizing map therefore learns the layout of the semantic space, i.e. the possible meanings of the words in the language.

Associations between the two maps are learned at the same time as the two maps are organized. A lexical symbol and its meaning are presented at the same time, resulting in activations on both maps. Associative connections between the maps are then adapted based on Hebbian learning, i.e. by strengthening those connections that link active units, and normalizing all connections of each unit:

$$a'_{ij,mn} = \frac{a_{ij,mn} + \alpha\eta_{ij}\eta_{mn}}{\sum_{uv}(a_{ij,uv} + \alpha\eta_{ij}\eta_{mn})}, \tag{3}$$

where $a_{ij,mn}$ is the weight on the associative connection from unit (i,j) in one map to unit (m,n) in the other map and η_{ij} is the activation of the unit. As a result of this

learning process, when a word is presented to the lexical map, its associated meaning is activated in the semantic map, and vice versa. DISLEX therefore models both comprehension and production in the lexicon.

The DISLEX model was evaluated in prior work in three ways: First, it was shown to function well as a practical lexicon component in a large language processing system called DISCERN [12], for processing script-based stories. Second, it was tested as a cognitive model of human lexical processing. The spatial organization of the lexicon is motivated by the maps in the brain, such as those suggested to underlie the lexical system [16]. Because of this structure, local lesions to the system result in category-specific impairments, similar to those documented in aphasic patients [18]. Noisy propagation between maps gives a possible computational explanation to many dyslexic phenomena [1], such as lexical errors ("ball" → "doll"), semantic errors ("lion" → "tiger") and combined errors ("sympathy" → "orchestra"). Third, the model was extended to model lexical development [11,10,19]. The extended model, called DEVLEX, is trained with gradually more words. It accounts for a range of phenomena in lexical acquisition, including effects of lexical categories such as representation of nouns/verbs, word frequency, word length and word density. AoA is considered to be an important factor in the developing bilingual lexicon as well. For instance, the timing of L2 acquisition impacts the structural representation of L1 and L2 maps. When L2 is acquired later than L1, it becomes dependent upon the L1 semantic map in a "parasitic" way and induces higher rates of errors [19].

DISLEX therefore forms a solid foundation for modeling bilingual lexical processing and its breakdown and recovery as well. The first step is to extend it to two languages with different proficiency and AoA, as described below.

4   The Bilingual DISLEX model

The original DISLEX model was extended to include lexica for two languages, L1 and L2. By varying the amount and timing of training in the two languages, the model can represent an individual with a given proficiency and AoA of L1 and L2.

The bilingual lexicon model was constructed to match typical subjects in the empirical studies: L1 is Spanish, L2 is English, and the AoA of English varies. Furthermore, because the patients are typically more proficient in English than Spanish, English is the dominant language (L2d) and Spanish is the weaker language (L1w). The L1 and L2 lexical symbols are each laid out in a different map in the model (Fig. 3). To construct such a model, 30 corresponding words, represented orthographically, were used in each language, together with 18 distinct semantic meanings; the mapping from words to meanings was thus many-to-many. As an example, the model was organized to simulate both early (Fig. 3a) and late acquisition of L2 (Fig. 3b) where L2 was the dominant language.

In the early L2 acquisition model, each Spanish symbol - English symbol - meaning triple was presented 45 times, and during the same time period, each English symbol - meaning pair was presented an additional 15 times, modifying the input and associative weights each time as described above. Such different amount of training in English and Spanish corresponds to the different exposure to the two languages during acquisition.
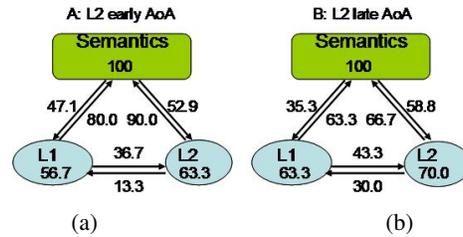
Fig. 3. Bilingual DISLEX models of two typical (hypothetical) individuals. Following the general architecture of Kroll & Stewart [9] (Fig. 1d), the bilingual DISLEX model consists of a map of lexical symbols for each of the two languages, a common map of word meanings, and associative connections between them. The solid arrows indicate strong and dashed weak associative connections; the numbers stand for percentage proficiency of each component. The model was trained in two different ways to match two different hypothetical patients, each with a different L2 AoA but high L2 proficiency. In this manner, the level of AOA, proficiency and impairment can each be altered to understand the individual contribution of each variable.

In the late L2 acquisition model, the Spanish symbol - meaning pairs were first presented 25 times; then, the English - Spanish - meaning triples were shown 20 times, interleaved with 40 presentations of the English symbol - meaning pairs. In both cases, the lexicon self-organized to represent the two languages and the mapping between them (Fig. 4). Each lexical map is organized according to the similarity of the word shapes (mostly word length and matching letters, which are the most prominent characteristics of the "blurry bitmap" representation used). The semantic map is organized according to the word meaning (i.e. animate words are clustered together, as are verbs and objects).

When a meaning representation is given an input to the semantic map, the corresponding units in the English or Spanish map are activated through associative connections. The main difference between early and late L2 maps is that while the early L2 map is organized relatively smoothly, the late L2 map is irregular and uneven. To quantify the behavior of these models, the 18 meaning representations were each presented to the semantic map in turn and propagated to the English map and to the Spanish map, modeling naming in the two languages; similarly, the 30 English words and 30 Spanish words were each presented as input to the appropriate map and propagated to the semantic map, modeling word comprehension. In each case, the unit with the maximally responding unit was found in the input map and in the associated map. If its weight vector was closest to the correct representation, the output for the word in the lexicon was correct.

A typical bilingual performance was observed in this process in both early and late L2 models (Fig. 3). With early English AoA, English (L2d) dominated Spanish (L1w) in production (53 vs. 47% accuracy) as well as in comprehension (90 vs. 80%). Importantly, the same was true of late English AoA both in production (58 vs. 35%) and in comprehension (67 vs. 63%). (Note that only the relative proficiency is important; the absolute proficiency can be adjusted by changing the amount and rate of learning in the model). Most interestingly, an important asymmetry emerged between the L1 and L2

Fig. 4. A DISLEX model of normal bilingual behavior. The L1 and L2 maps are organized according to the perceptual similarities (mostly word length), and the semantic map according to semantic meaning (e.g. animals, objects, verbs). A word presented to one of the maps (in this case semantic meaning `doll`) activates that map, and the activity propagates through associative connections to the other maps (in this case L1 and L2, activating the units representing `DOLL` and `MUNECA`, as well as a few neighboring units). The colors white → yellow → red → black indicate increasing response. The connections to L1 are less specific (activating a wider area) and therefore result in more errors than those to L2. Also, although both the early AoA L2 (bottom right) and late AoA L2 (top right) result in roughly equal performance, the early map is better organized than the late map (which consists of several small discontinuous clusters). In this manner, DISLEX models bilingual naming and comprehension with different proficiencies and different age of acquisition in the two languages.

maps: An L1 word activates the corresponding L2 word more strongly than the other way around (37 vs. 13% early, 43 vs. 30% late), showing more proficiency in English (L2d) than in Spanish (L1w), and demonstrated dominance of English over Spanish in transfer between the languages. This asymmetry is consistent with behavioral data (Section 2) showing that lexical activation in the non-dominant language results in activation of corresponding representations of the dominant language.

In a subsequent experiment, the role of the languages was reversed, resulting in a model with Spanish as the dominant (L1d) and English as the weaker (L2w) language. In a third experiment, the model was exposed equally frequently to the two languages (L1 = L2), and neither language dominated the other in the resulting model. These results were obtained both with early and with late L2 training, with consistent differences in the L2 map organization (Fig. 4). Together these experiments suggest that by training the model with different starting times and frequencies, the relative proficiency and

organization in the two languages can be tuned, making it possible to fit the model to a particular individual's performance and learning history.

5   Discussion and Future Work

The model is consistent with behavioral data in two important ways. First, relative language proficiency is modulated by the amount of exposure to each specific language. Second, various levels of L2 language performance can be achieved with both early and late L2 AoA. Further, the late L2 AoA results in less well organized L2 map even when L2 achieves eventual high proficiency.

In the future, it may be possible to use the model to fit the pre-stroke performance and learning history of an individual patient, and then used to derive an optimal treatment for that patient. Damage to the lexical system can be modeled in DISLEX in two ways: (a) Units or connections can be deleted from the model, and (b) noise can be added to the connections. By controlling the type and extent of damage, it will be possible to fit the model to the profile of an individual patient. In rehabilitation training, then, the model will be presented with selected word-meaning pairs in the two languages and it will continue self-organizing using the same mechanisms as during initial training. By varying the types of words (such as concrete vs. abstract words, rare vs. frequent words, short vs. long words, and word categories) and numbers of training pairs in the two languages systematically, it should be possible to determine a training recipe that leads to fastest and most complete recovery. Such an ability could greatly improve our ability to treat bilingual aphasia in the future.

6   Conclusion

A bilingual lexicon model consisting of a common map of word meanings and a separate maps for the words in the two languages is consistent with current theory of bilingual processing in the mental lexicon. By varying the frequencies with which the maps are trained, and the timing of the training in the two languages, models of individual proficiency and order in the two languages can be developed. The approach can be used to develop models of individuals, and in the future, may form a foundation for discovering individually optimized recipes for treatment of aphasia.

References

1. M. Coltheart, K. Patterson, and J. C. Marshall, eds., *Deep Dyslexia*, second edition, Routledge and Kegan Paul, London; New York (1988).
2. A. Costa, W. Heij, and E. Navarette, The dynamics of bilingual lexical access, *Bilingualism: Language and Cognition*, 9:137–151 (2006).

3. A. Costa, M. Miozzo, and A. Caramazza, Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection?, *Journal of Memory and Language*, 43:365–397 (1999).

4. G. W. Cottrell and S. L. Small, A connectionist scheme for modelling word sense disambiguation, *Cognition and Brain Theory*, 6:89–120 (1983).

5. A. de Groot, Determinants of word translation, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18:1001–1018 (1992).

6. L. Edmonds and S. Kiran, Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection., *Aphasiology*, 18:567–579 (2004).

7. G. E. Hinton and T. Shallice, Lesioning an attractor network: Investigations of acquired dyslexia, *Psychological Review*, 98:74–95 (1991).

8. J. Kroll and J. Curley, Lexical memory in novice bilinguals: The role of concepts in retrieving second language words, in: *Practical Aspects of Memory, Vol. 2*, Wiley, New York, NY (1988).

9. J. F. Kroll and E. Stewart, Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations, *Journal of Memory and Language*, 33:149–174 (1994).

10. P. Li and I. Farkas, A self-organizing connectionist model of bilingual processing, in: *Bilingual Sentence Processing*, R. Heredia and J. Altarriba, eds., Elsevier, Amsterdam (2002). In press.

11. P. Li, X. Zhao, and B. MacWhinney, Dynamic self-organization and early lexical development in children, *Cognitive Science*, 31:581–612 (2007).

12. R. Miikkulainen, *Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory*, MIT Press, Cambridge, MA (1993).

13. R. Miikkulainen, Dyslexic and category-specific impairments in a self-organizing feature map model of the lexicon, *Brain and Language*, 59:334–366 (1997).

14. D. C. Plaut, A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing., *Cognitive Science*, 23:543–568 (1999).

15. M. Potter, K. So, B. von Eckardt, and L. Feldman, Lexical and conceptual representation in beginning and proficient bilinguals, *Journal of Verbal Learning and Verbal Behavior*, 23:23–38 (1984).

16. M. Spitzer, U. Kischka, F. Gückel, M. E. Bellemann, T. Kammer, S. Seyyedi, M. Weisbrod, A. Schwartz, and G. Brix, Functional magnetic resonance imaging of category-specific cortical activation: Evidence for semantic maps, *Cognitive Brain Research*, 6:309–319 (1998).

17. G. R. Tucker, A global perspective on bilingualism and bilingual education, Technical Report EDO-FL-99-4, Center for Applied Linguistics (1999).

18. E. K. Warrington and R. A. McCarthy, Categories of knowledge: Further fractionations and an attempted integration, *Brain*, 110:1273–1296 (1987).

19. X. Zhao and P. Li, Bilingual lexical representation in a self-organizing neural network, in: *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, D. S. McNamara and J. G. Trafton, eds., 755–760, Cognitive Science Society, Austin, TX (2007).