In D. Reitter & F. E. Ritter (Eds.), Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016). University Park, PA: Penn State.

A Computational Model of Semantic Convergence in Bilinguals

Shin-Yi Fang (szf4@psu.edu)
Department of Psychology, Pennsylvania State University
University Park, PA 16802 USA
Benjamin D. Zinszer (bzinser@gmail.com)
Department of Brain and Cognitive Sciences, University of Rochester
Rochester, NY 14627 USA
Barbara C. Malt (barbara.malt@lehigh.edu)
Department of Psychology, Lehigh University
Bethlehem, PA 18015, USA
Ping Li (pul8@psu.edu)
Department of Psychology, Pennsylvania State University
University Park, PA 16802 USA

Abstract

Patterns of object naming were influenced by the language of the people used. In this study, we aimed to simulate the bilinguals' naming patterns and investigate the underlying mechanisms. Our results replicated the empirical findings that (1) bilingual speakers develop converged naming patterns in their two languages that are distinct from those of monolingual speakers of each language and (2) how the bilingual semantic convergence was manifested in their lexical representations. We also demonstrated that both orthography and connections between two languages are important to establish converged naming patterns in bilinguals. Furthermore, our modeling data suggested that the strength of name agreement in the two languages are involved in the cooperation and competition relationships on the object naming. Our model provides a foundation for examining the factors contributed to the patterns of object naming in bilinguals.

Keywords: object naming; lexical categories; modeling; self-organizing map; bilingual lexicon

Introduction

Across languages, objects are not always been classified into the same categories. For example, Malt, Sloman, Gennari, Shi, and Wang (1999) asked speakers of American English, Argentinean Spanish, and Mandarin Chinese to name 60 common household containers and found that the naming patterns differ substantially as a function of the language spoken. For speakers of two languages, how they dealt with the inconsistent mapping relationships between objects and names could enhance our understanding about language learning. Ameel, Storms, Malt, and Sloman (2005) investigated the naming patterns of adult Dutch-French simultaneous bilinguals, Dutch monolinguals, and French monolinguals. They found that object naming patterns by the bilingual speakers converge toward a pattern that is different from the naming patterns of monolinguals of each language, suggesting that simultaneous bilinguals do not behave like monolinguals in lexical categorization and the bilingual lexical representations are not simply the sum of two separate monolingual representations.

Ameel, Malt, Storms, and Van Assche (2009) further investigated the naming patterns in Ameel et al. (2005) and the typicality ratings between language groups. They examined if the bilinguals' semantic convergence was manifested in the centers and/or at the boundaries of lexical categories. Ameel et al. (2009) found that the translational equivalents have closer category centers in the naming patterns of the bilinguals than those of the monolinguals and the closeness in bilinguals were contributed by more than the boundary exemplars alone. Ameel et al. also found that bilinguals needed fewer dimensions to separate categories than monolinguals, suggesting that the category structures were less complex for bilinguals and that naming were more likely to be based on similarity to the prototypes.

In this study, we build a model based on self-organizing maps (SOM; Kohonen, 2001) to study bilingual object naming. By building and testing this computational model against existing data of Ameel et al. (2005; 2009), this work will provide the foundation to investigate underlying mechanism and further modeling study to manipulate learner characteristics such as age of exposure, frequency of input and proficiency in each language, as well as lexical input variables such as similarities between the lexical items.

Method

We constructed is a multi-layer SOM network, which includes three basic SOMs (i.e., semantic, phonological, or orthographic). The three SOMs are connected via associative links updated by bi-directional Hebbian learning. In addition to the basic SOM architecture, we added lateral between languages in the model to simulate between-language interactions. The lateral connections are implemented with the nodes that are fully connected with each other. The connection weights are updated via Hebbian learning rule. We used the monolingual naming data from Ameel et al. (2005) as the basis of input to the model. We trained the model on representations of pictures of 73 bottle-like objects that are typically named as bottle, jar, or container in American English or else to have one or more

salient properties in common with objects called by those names. We call this the "standard" model.

In order to identify the role of orthographic information and the effects of lateral connections, we constructed two comparison models to contrast with the standard model, respectively. The comparison models are identical to the standard model in every respect, except that in each comparison model one component is removed: Comparison Model A was constructed without orthographic information and Comparison Model B was constructed without lateral connections between languages.

Results

We found that, similar to Ameel et al. (2005), our computational model shows higher correlation between bilinguals' two language (0.97) than between two monolingual languages (0.63), indicating that our model simulated empirical naming patterns and captured bilinguals' lexical categorization. Both comparison Model A and B showed significantly worse performance than standard Model, suggesting that the both orthography and lateral connections were important in the model's object naming. Furthermore, the Model B performed worse than Model A suggesting that the role of lateral connection is more critical than the role of orthography. We also replicated that the main findings in Ameel et al. (2009), indicating that the translational equivalents have closer category centers in the naming patterns of the bilinguals than those of the monolinguals and bilinguals have less complex category structures than monolinguals.

We further examined the model to explore what properties in the model might have influenced the naming patterns. Our model shows that if an object elicited a strong level of activation for a word in the target language, the output name of the model for bilingual naming will be the same as the name for monolingual naming. However, if the activation level is weak in the target language and the cross-activation from the non-target language is strong, the output names of the model could be different between bilingual naming and monolingual naming. For example, if a bottle-like object elicited strong activation of the word *fles* in Dutch, both the monolinguals and bilinguals will produce *fles* in Dutch; whereas the activation of *fles* in Dutch is weak and the activation of *bus* in Dutch outperformed *fles*, due to combination of its original activation from Dutch and the strong lateral activation from French. In this example, the monolinguals will produce *fles*, but the bilinguals will produce *bus*.

Discussion

In this study, we successfully built a bilingual lexical categorization model based on connectionist SOM architecture that has been previously tested in other domains of language acquisition and processing. Our model simulated bilingual semantic convergence in the naming of common household objects as reported in the empirical literature (Ameel et al., 2005; 2009).

Our standard model performed significantly better than the two comparison models in which either lateral connection or orthography components not included. This is particularly important as our model is designed to simulate the dynamic interactions between two languages, and the orthography and the lateral connections play a critical role in bilingual lexical categorization. Our results demonstrate how, for simultaneous bilinguals, the processing of one language can be influenced by the other language (i.e., bi-directional influences between languages). Our simulation also showed that the strength of name agreement is an important factor to determine lexical naming patterns for bilinguals. If the object has high name agreement in one language, the influence from other language through lateral connection cannot easily change its category and vice versa.

The viability of our model paves the way to use modeling to study a wide range of learner and object name variables that may influence behavioral outcomes for simultaneous and sequential bilinguals (such as variables discussed before, including age of onset, proficiency, and frequency of input).

Conclusion

This study used a connectionist self-organizing model to simulate object naming patterns in bilinguals and to identify mechanisms of lexical semantic convergence. We successfully replicated the lexical convergence patterns reported in empirical data from Ameel et al. (2005), and we further investigated the mechanisms and important factors that modulate bilinguals' naming categorization. We demonstrated that the lateral connections play an important role in lexical convergence. Finally, we have identified the role of name agreement strength on bilinguals' object naming. This study provides a first computational model that examines the dynamic interaction between two lexicons in the process of naming objects using single versus multiple languages.

Acknowledgments

This research was supported by a grant from the National Science Foundation (BCS-1057885) to BCM and PL.

References

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009).Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, 60(2), 270–290.

Ameel, E., Storms, G., Malt, B. C., &Sloman, S. A. (2005).How bilinguals solve the naming problem. *Journal of Memory and Language*, 53, 60–80.

Kohonen, T. (2001). *The self-organizing maps*. (3rd ed.). Berlin: Springer.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, 40, 230–262.