CHALLENGES FACING A COMPLEMENTARY-SYSTEMS APPROACH TO ABSTRACT AND EPISODIC SPEECH PERCEPTION

McLennan, C. T.

Cleveland State University

c.mclennan@csuohio.edu

ABSTRACT

It has been nearly a decade since the publication of Goldinger's [4] *Psych. Review* paper in which he presented his episodic theory of lexical access. Moreover, Goldinger's (and others') empirical work [3, 14] providing evidence for episodic representations predates the formal presentation of his episodic theory. This is an appropriate time to note how the field has progressed in the past decade with respect to the debate over the nature of lexical representations. As evidenced by the two main papers, the emphasis is no longer on whether there are abstract and/or episodic representations. Instead, the focus is now on the ideal framework that can account for their coexistence. Goldinger's presentation of a complementary systems (hereafter CS) approach offers a glimpse into the direction that many future investigations of lexical representation may be headed. I discuss some of the challenges facing a CS approach in an effort to stimulate further discussion, and to help provide an impetus for future empirical, theoretical, and modeling studies.

Keywords: Word recognition, speech perception, language, abstract, episodic

1. INTRODUCTION

Despite differing theoretical perspectives, the need for both abstract and episodic representations is acknowledged in both of the main papers. Therefore, if our task as researchers working in this area were simply to determine whether representations were abstract and/or specific, our task would seemingly be complete. However, our aim is not simply determining which representations affect processing, but more importantly, determining the circumstances under which each type (abstract and episodic) is more or less likely to do so. Goldinger's discussion of a CS approach brings us closer to achieving this more demanding goal by appropriately redirecting our focus. I contend that even more will be gained by considering some of the challenges that confront models of lexical representation and processing, including the CS approach.

2. CHALLENGES

Although the CS approach offers a new framework for the coexistence of abstract and episodic representations, there are some important issues that merit further consideration. First, either abstract or episodic representations may be the dominant or default type of representation, with the other type only in evidence under precise, and perhaps relatively rare, circumstances. Second, abstract and episodic representations affect perception at different points during processing. Third, which type of representation dominates, and when during processing each type affects perception, are both likely to vary in certain populations (e.g., L2 learners).

2.1. Default Representation

I agree with Goldinger's current theoretical stance, and appreciate the need to account for both abstract and episodic representations [15]. Indeed, my colleagues and I have advocated such a position previously [9]. However, it is possible that one type of representation is dominant, and thus primarily responsible for the speed and accuracy with which listeners typically process spoken language. If there were such a default representation, then I suspect it would be relatively abstract in nature (although it is

unlikely that either type of representation will serve as the default across all situations and populations).

Much of the evidence for the representational status of episodic representations has been obtained using the long-term repetition-priming paradigm. Although numerous studies have provided evidence in support of detailed episodic representations, it has been my experience (and I suspect the experience of many other researchers as well) that while repetition-priming effects are quite robust, and observable under a wide variety of conditions, specificity effects, indicative of episodic representations, are typically only observed under certain conditions. Furthermore, many of the studies that provide evidence for episodic representations may not be tapping into on-line processing. Nevertheless, there are sufficient data to support the idea that episodic information is stored in memory and can have long-lasting effects, contrary to extreme abstractionist positions that posit no role for episodic representations. However, few studies offer support for the involvement of episodic representations during the immediate on-line perception of spoken language (e.g., by reporting reaction-times).

I agree with Cutler and Weber that the eye-tracking paradigm is a useful tool, and may be used to investigate a number of different issues in L1 and L2 speech perception. For example, eye-tracking has been used to demonstrate that bilingual lexical activation may be constrained by fine-grained acoustic-phonetic information [7]. Eye-tracking is also well suited for examining lexical processing early and as perception unfolds, providing new opportunities to examine whether episodic information can in fact affect the immediate on-line perception of spoken language.

2.2. Time-course Effects

Goldinger concludes his paper by stating that the CS account "unites long-term memory with real-time perception". However, it is not entirely clear whether this framework can account for fine-grained temporal effects. In particular, recent evidence suggests that rate- and talker-independent abstract representations affect processing early, and rate- and talker-specific episodic representations affect processing relatively late [13].

One of my colleagues (González) and I are currently building on this previous work by examining the time-course of talker-specific effects in the perception of foreign-accented speech. If listeners process foreign-accented speech relatively slowly, then the prediction based on previous findings with slowed processing is that talker-specific episodic representations should be in evidence. Although this work is still ongoing, our preliminary data are consistent with this prediction. When Spanish listeners participate in a lexical decision task in which the Spanish stimuli are spoken by native Spanish speakers, there is no evidence for talker-specific representations in an easy lexical decision task (i.e., when unword-like nonwords are included as the filler stimuli, and processing is relatively fast), consistent with previous results [13]. However, when the stimuli are spoken by native English speakers, and thus produced with a foreign accent, specificity effects emerge in both easy and hard lexical decision tasks.

The CS framework could potentially account for findings in which abstract representations come into play early and episodic representations come into play relatively late. In particular, the hippocampal system responsible for episodic effects may respond to the input more slowly than the more stable cortical system responsible for abstract patterns. However, relatively few time-course investigations have been conducted, and it is possible that under other circumstances episodic representations will affect perception earlier during processing than abstract representations. Future investigations should provide a more complete picture of the full range of possible time-course effects. Nevertheless,

accounting for, and perhaps predicting, time-course effects can be considered one of the challenges facing the CS approach.

2.3. Population Differences

Evidence for abstract and episodic representations has been obtained in a number of populations, including hearing impaired listeners [18] and cochlear implant patients [8]. However, the role that these representations play relative to normal hearing listeners is unclear. It would be interesting to learn whether the CS approach predicts that one type of representation is more likely to affect processing in certain populations than others (e.g., episodic more so in hearing impaired listeners).

Cutler and Weber argue that because L1 and L2 learning both occur in the same brain, abstract and episodic sources of knowledge probably play a role in both L1 and L2. Although episodic representations have been found to play a role in L2 vocabulary learning [19], and to have long-lasting effects on the learning of difficult contrasts [1], the circumstances in which abstract and episodic representations affect processing may be different in L1 and L2. Perhaps abstract representations are more involved in L1 perception, with episodic representations only coming into play in precise circumstances, and episodic representations play a more central role in L2 perception[i]. However, the representational and processing realities are likely to be more complex. That is, L2 proficiency and other factors regarding a bilingual's languages (e.g., the degree of similarity between L1 and L2) are likely to influence the roles that episodic and abstract representations play in L2 perception.

Evidence for both types of representations has also been found throughout the lifespan, including infants [6], preschool children [2, 17], and older adults [18]. As Cutler and Weber explain, L2 learners bring many sources of knowledge to bear on the task of language learning that are not available to infants acquiring an L1. Furthermore,

time-course effects may change throughout the lifespan [12]. Therefore, the roles that abstract and episodic representations play in young children and infants may differ from the roles that these types of representations play in adult L2 learners.

In addition to the general issues discussed thus far, there are also challenges more specific to the CS model. According to the CS model, extensive exposure to unique traces will slowly affect abstract information. What qualifies as "extensive" and "unique", and just how slowly abstract information will be affected, are all questions that need to be addressed. At what point do traces qualify as unique? How does the system(s) determine whether variable surface information associated with phonetic input is representative of a new accent, as opposed to inter- and/or intra-talker variability? Although an elegant account of the consolidation process has been proposed [11], additional work should clarify how this process works in the context of L1 and L2 language perception.

3. DISCUSSION

The field has entered into a new phase in which, rather than debating over abstract versus episodic representations, efforts are now focused on determining the ideal framework that can account for their coexistence. Indeed, even in the field of linguistics, where abstractionist approaches have dominated, there are arguments against traditional representational approaches [16].

The CS model discussed by Goldinger has some clear advantages, including its ability to account for hemispheric differences in font-specific [10] and talker-specific [5] priming, results that are more challenging for unitary accounts. Furthermore, the ability of the CS model to successfully simulate behavioral data, as Goldinger demonstrates, further illustrates the utility of this approach. Nevertheless, future empirical work aimed at addressing the issues discussed above may reveal shortcomings of this model. However, I suspect that such

shortcomings, while providing new challenges (and thus stimulating additional research), will be limited to the specifics of the current version of the model, rather than more general problems associated with a CS framework. For a variety of reasons (addressed by Goldinger, e.g., compatibility with neuroscientific literature, etc.), a CS account - an openly hybrid approach - is likely to be a central part of future studies on L1 and L2 speech perception and spoken word recognition.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1] Bradlow, A.R., Akahane-Yamada, R., Pisoni, D.B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Percep. & Psychophy., 61*, 977-985.

[2] Fisher, C., Hunt, C., Chambers, K., & Church, B. (2001). Abstraction and specificity in preschoolers' representations of novel spoken words. *J. of Mem. & Lang., 45,* 665-687.

[3] Goldinger, S.D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *J. Exp. Psych: Learn., Mem., & Cog., 22*, 1166-1183.

[4] Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psych. Rev., 105,* 251-279.

[5] González, J. & McLennan, C.T. (2007). Hemispheric differences in indexical specificity effects in spoken word recognition. *J. Exp. Psych: Human Perc. & Perf., 33,* 410-424.

[6] Houston, D.M., & Jusczyk, P.W. (2000). The role of talker-specific information in word segmentation by infants. *J. Exp. Psych: Human Perc. & Perf., 26*, 1570-1582.

[7] Ju, M., & Luce, P.A. (2004). Falling on Sensitive Ears: Constraints on Bilingual Activation. *Psych. Sci., 15*, 314-318.

[8] Kirk, K.I., Pisoni, D.B., & Miyamoto, R.C. (1997). Effects of stimulus variability on speech perception in listeners with hearing impairment. *J. Sp., Lang. Hear. Res., 40,* 1395-1405.

[9] Luce, P.A., McLennan, C.T., & Charles-Luce, J. (2003). Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition

priming. In Bowers, J. & Marsolek, C. (Eds.), *Rethinking implicit memory*, pp. 197-214. Oxford University Press.

[10] Marsolek, C.J., Kosslyn, S.M., & Squire, L.R. (1992). Form-specific visual priming in the right cerebral hemisphere. *J. Exp. Psych: Learn., Mem., & Cog., 18,* 492–508.

[11] McClelland, J., McNaughton, B., & O'Reilly, R. 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. *Psych. Rev., 102*, 419-437.

[12] McLennan, C.T. (2006). The time course of variability effects in the perception of spoken language: Changes across the lifespan. *Lang. & Speech, 49,* 113-125.

[13] McLennan, C.T. & Luce, P.A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *J. Exp. Psych: Learn., Mem., & Cog., 31,* 306-321.

[14] Mullennix, J.W., Pisoni, D.B., & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *J. Acoust. Soc. Am., 85,* 365–378.

[15] Pisoni, D.B. & Levi, S.V. (2005). Some observations on representations and representational specificity in speech perception and spoken word recognition. *Research on Spoken Lang. Processing, 27,* 3-26.

[16] Port, R.F. & Leary, A.P. (2005). Against formal phonology. *Lang., 81,* 927-964.

[17] Ryalls, B.O., & Pisoni, D.B. (1997). The effect of talker variability on word recognition in preschool children. *Dev. Psych., 33,3,* 441–452.

[18] Sommers, M.S. (1997). Stimulus variability and spoken word recognition II: The effects of age and hearing impairment. *J. Acoust. Soc. Am., 101,* 2278-2288.

[19] Sommers, M.S. & Barcroft, J. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Lang. Acq., 27,* 387-414.

---

[i] I am focusing on the roles that abstract and episodic representations of L1 play during L1 perception, and the roles that abstract and episodic representations of L2 play during L2 perception. I have not (for example) addressed the role that L1 representations play during L2 perception, such as the effects of L1 phonology on L2 perception. However, the more robust abstract representations in L1 are likely to affect processing in both L1 and L2.