

Introduction to the special issue on exemplar-based models in linguistics

SUSANNE GAHL and ALAN C. L. YU

Recent years have witnessed the emergence of a cluster of approaches in linguistic theory known as exemplar-based models. Such exemplar-based models are being developed in research domains as diverse as phonetics, phonology, morphology, historical linguistics, semantics, syntax, and language acquisition. This special issue brings together, for the first time, articles applying exemplar-based models to a wide range of linguistic subdisciplines. We hope that this collection will give readers a sense for the challenges and new research tools that exemplar-based models offer.

The central idea behind exemplar-based models is that mental representations consist of memory traces of specific tokens. This idea runs counter to the goal of developing maximally simple, redundancy-free representations, a goal that has been central to many proposals within linguistic theory. For example, much phonological work in the 1980s went into the study of underspecification theory, which banned non-distinctive or predictable feature values from underlying representations. Similarly, highly influential models of syntax have taken economy to be a guiding consideration in their conceptions of the lexicon and its grammatical interfaces (Chomsky 1993, 1995).

Inspired by studies of exemplar-models of categorization (Goldinger 1996; Hintzman 1986; Kruschke 1992; Nosofsky 1986), some researchers have started exploring new models of the lexicon and its interface with other levels of linguistic competence which call into question the goal of redundancy-free representations. A category in exemplar-based models is defined by a collection of memorized tokens, or “exemplars”. Each exemplar may belong to many categories simultaneously. An exemplar-based speech processing system recognizes inputs and generates outputs by analogical evaluation across a lexicon of distinct memory traces of remembered tokens of speech. While specific exemplar-based models differ on how new experiences are assigned to relevant categories and integrated with the stored exemplars, all exemplar-based mod-

els assume that each experience alters the entire category system slightly. As the perceptual memories associated with a category accumulate and are incrementally updated, the distribution of these forms may shift. As a consequence, effects of frequency and recency of use provide crucial testing grounds for such models.

Exemplar-based models have received empirical support from many sources in recent years. Studies in perception have shown that listeners may rely on fine-grained phonetic details in word recognition (Goldinger 1996; Johnson 1997). Pierrehumbert (2002) reviews a large body of literature showing that lexical representations must include rich, subphonemic detail. Several contributions in this volume offer further support for this information-rich conception of linguistic representations. *Ernestus* shows that subphonemic information affects the production and recognition of words, based on evidence from nonce words and existing ones. *Comine and Pinnow* offer evidence in support of the idea that the frequency of phonological variants affects spoken word recognition.

With their assumption that each experience alters category representations, exemplar-based models lend themselves naturally to the study of language change. For example, Bybee (2001) reviews evidence from historical changes arguing that lexical representations of words must include incrementally updated information about the phonetic distribution of each word. Kirchner (to appear) offers a computational simulation of phonologization using an exemplar-based speech processing system. Batali (2002) investigates the emergence of recursive syntactic structures using an exemplar-based computational simulation of language acquisition. In a similar vein, *Wedel* articulates an evolutionary model of language change which takes individual exemplars as the unit of information passed on from generation to generation.

The study of language change is intimately connected with the study of language acquisition. *Abbot-Smith and Tomasello* offer an exemplar-based perspective on children's acquisition of syntax. On their view, learning proceeds from individual tokens in experience, not from an innate universal grammar. In fact, what is acquired is not a system of abstract rules, but an inventory of concrete examples, from which patterns are only gradually extracted.

The notion that what children acquire initially is not a grammar, but a set of fully specified constructs raises the issue of how to model productivity. Indeed, the claim that exemplar-based accounts cannot account for productivity is perhaps the most commonly raised objection to such models. This is particularly true in the area of syntax, where speakers' "infinite generative capacity" and their ability to make well-formedness judgments of novel sentences represent foundational observations (Chomsky 1957). *Bod* reviews a class of exemplar-based syntactic models without "rules", showing that they successfully model both productivity and meta-linguistic judgments.

The notion that representations consist of detailed memory traces implies that there is no particular reason to expect a strict separation of different types of linguistic information: One and the same memory trace may instantiate phonetic/phonological categories and syntactic constructs, for example. *Hay and Bresnan* explore the consequences of assuming that syntactic phrases may be stored along with phonetic detail. Their article combines insights from exemplar-based models of phonetics with those from exemplar-based models of syntax.

Exemplar-based models have also proven to be useful in bridging the gap between sociolinguistics and discussions of phonetics and speech perception. Hay, Jannedy and Mendoza-Denton (1999), for example, found that lexical frequency and ethnicity of addressee may influence the phonetic implementation of certain phonetic features. Clopper and Pisoni (2004) show that dialect perception lends itself to an exemplar-based account as well. *Hay, Nolan and Drager* further explore the implications of exemplar-based approaches for speech perception, by demonstrating that listeners' vowel-identification judgments depend on the presumed dialect origin of the speaker. Such findings are difficult to explain without assuming that individual words not only have associated phonetic distributions but are also associated with social indexical information.

Preparing this collection has been highly enjoyable. We are immensely grateful to the contributors and to the reviewers, for their time and their thoughtfulness. Several of the articles in this collection were presented at a workshop at the 2005 meeting of the Linguistic Society of America. We are grateful to the LSA for their support of the workshop, and to the many people we encountered along the way who share our goal of fostering interaction among subdisciplines within Linguistics and between Linguistics and neighboring disciplines.

University of Chicago

References

- Batali, John (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In *Linguistic Evolution through Language Acquisition*, Ted Briscoe, Ted (ed.), 111–172. Cambridge: Cambridge University Press.
- Bybee, Joan (2001). *Phonology and Language Use*. Cambridge, UK: Cambridge University Press.
- Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.
- (1993). A minimalist program for linguistic theory. In *The View from Building 20*, Kenneth Hale and Samuel Jay Keyser (eds.), 1–52. Cambridge, MA: MIT Press.
- (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Clopper, Cynthia and David B. Pisoni (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32: 111–140.

- Goldinger, Stephan D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166–1183.
- Hay, Jennifer, Stefanie Jannedy, and Norma Mendoza-Denton (1999). Oprah and /ay/: Lexical frequency, referee design and style. In *Proceedings of the 14th International Congress of Phonetic Sciences*, John J. Ohala et al. (eds.) 1389–1392. Berkeley: University of California.
- Hintzman, Douglas L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review* 93: 411–428.
- Johnson, Keith (1997). Speech perception without speaker normalization: An exemplar model. In *Talker Variability in Speech Processing*, Keith Johnson and John Mullenix (eds.), 145–166. San Diego: Academic Press.
- Kirchner, Robert (to appear). Preliminary thoughts on ‘phonologization’ within an exemplar based speech processing model. *UCLA Working Papers* 6.
- Kruschke, John K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review* 99: 22–44.
- Nosofsky, Robert (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115: 39–57.
- Pierrehumbert, Janet (2002). Word-specific phonetics. In *Laboratory Phonology VII*, Carlos Gussenhoven and Natasha Warner (eds.), 101–140. Berlin: Mouton de Gruyter.