It's not how many dimensions you have, it's what you do with them: Evidence from speech perception

Bob McMurray[a] and David Gow[b]

[a]*Department of Psychology, University of Iowa, Iowa City, IA 52242;* [b]*Neuropsychology Laboratory, Massachusetts General Hospital, Boston, MA 02114.* **bob-mcmurray@uiowa.edu** **gow@helix.mgh.harvard.edu**

Abstract: Contrary to Pothos, rule- and similarity-based processes cannot be distinguished by dimensionality. Rather, one must consider the goal of the processing: what the system will do with the resulting representations. Research on speech perception demonstrates that the degree to which speech categories are gradient (or similarity-based) is a function of the utility of within-category variation for further processing.

Pothos attempts to derive computational distinctions between rule- and similarity-based processes from the observation that similarity processes operate over many dimensions, whereas rules operate over very few. He argues that dimensionality can account for the intuitive characterization that similarity operations tend to allow gradient mapping whereas rules operations require strict (discrete) mapping. We argue, however, that whether the mapping is graded or discrete is a primary distinguishing characteristic of the two systems and is not derivative upon high or low dimensionality.

Dimensionality is actually a relatively weak discriminator. It is based on a decontextualized notion of cognitive processing that pays too little attention to the goals of a given computation. Cognitive processes like categorization must directly support either action or further computation. Pothos's error is to overlook the significance of this aspect of the computational problem. The usefulness of gradient information for carrying out a specific cognitive task is a more useful guide for understanding why a given process allows more or less (similarity-like) gradient mapping.

Our research in speech perception and spoken-word recognition clearly demonstrates the primacy of processing goals over dimensionality in discriminating rule and similarity operations. Work in speech perception and phonology has long assumed rule-based operations operating on discrete categories (e.g., Categorical Perception, Liberman et al. 1957). Voicing perception is an important case because much of the evidence for categorical perception of voicing comes from work in which a *single* stimulus dimension, voice onset time (VOT), is manipulated (e.g., Liberman et al. 1961).

With its basis in a single dimension, the perception of voicing should, in Pothos's terms, be a clear case of rule-based processing. However, recent work shows that voicing perception is gradient for some purposes and categorical for others. McMurray et al. (2002) presented subjects with arrays of pictures and asked them to use a mouse to select a given picture identified by a spoken word. VOT was manipulated to make continua between words like beach and peach. Examination of subjects' eye movements during this task showed that subjects were increasingly likely to look at a lexical competitor (e.g., the peach after hearing "beach") as VOT approached the category boundary. In other work using similar VOT continua in another eye-tracking task, McMurray et al. (2003) and McMurray et al. (in preparation) found very little systematic sensitivity to such within-category variation in VOT when the task was to identify initial phonemes designated by orthographic letters rather than pictures.

Why do perceptual categories appear continuous (suggesting similarity-based operations) in the first experiment, but discrete (suggesting rule-based operations) in the second? The reason seems to be that in the context of speech recognition (which is better tapped by the picture-matching task), variation in VOT is correlated with future and past acoustic events (like vowel length, pitch, and prosodic strength) that may be useful for processes including rate normalization and the use of prosody to disambiguate syntactic structure. While these processes may be essential to the interpretation of connected speech, they play no role in the categorization of isolated CV syllables. Thus, in the phoneme categorization task, within-category variability along the same single dimension is noise and is best discarded.

This observation is consistent with a growing literature showing that within a category, variation affects the dynamics of lexical activation during word recognition (Andruski et al. 1994; Dahan et al. 2001; Gow & Gordon 1995). Strict matching in the form of categorical perception of speech sounds discards variation. This may make higher-level processing more efficient, but only if the variation that is lost is uninformative.

Work in the perception of lawful phonological variation is a case in point. In English, linguists describe a rule for coronal place assimilation that states that a segment with coronal place of articulation (e.g., /t/, /d/, or /n/) takes the place of articulation of a following non-coronal segment (e.g., /b/, /p/, /m/, /g/, /k/, or /ng/). As a process operating on a single dimension (place of articulation), place assimilation would appear to be a clear example of a rules process. However, close analysis of the acoustic consequences of coronal place assimilation suggests that it tends to create graded variation. For example, in the phrase "right berries," assimilation gives the [t] properties that are intermediate between those of a [t] and a [p] (Gow 2001; 2002; 2003).

This has fundamental consequences for perception. If modification were complete, listeners would need to rely on higher-level inferential processes to determine if the phrase "ripe berries" referred to the *right* berries or the *ripe* berries. Such a modification would be perceptually destructive in that it would neutralize any contrast between right and ripe. However, if modification is graded, listeners may potentially rely on perceptual processes to both recover underlying form and also anticipate upcoming context. That type of process is perceptually enhancing because it preserves information about underlying form while potentially encoding information about upcoming context. Evidence from explicit online behavioral tasks such as phoneme monitoring and form priming (Gow 2002; 2003), and implicit measures including spontaneous eye tracking and ERP (Gow et al. 2003; Gow & Holcomb 2002) shows that graded assimilatory modification creates a range of context effects that are not found when modification is discrete. These include regressive effects enabling the recognition of the underlying form and progressive effects facilitating perception of the context that follows assimilated segments. Gradient modification encodes information about multiple segments, and listeners' sensitivity to that gradiency allows them to exploit it fully in perception.

In summary, processes operating on very few stimulus dimensions may show either the strict matching associated with rule-based processes, or graded matching associated with similarity-based processing. The distinction between these types of processes reflects the potential value of subcategorical information for further processing, rather than the number of dimensions under consideration.

Rule versus similarity: Different in processing mode, not in representations

Rolf Reber

Department of Psychosocial Science, University of Bergen, N-5015 Bergen, Norway. **rolf.reber@psysp.uib.no**
http://www.uib.no/psyfa/isp/ansatte/reber.htm

Abstract: Drawing on an example from artificial grammar learning, I present the case that similarity processes can be computationally identical to rules processes, but that participants in an artificial grammar learning experiment may use different processing modes to classify stimuli. The number of properties and other representational differences between rule and similarity processes are an accidental consequence of strategies used.

If rule processes and similarity processes were separated by the number of properties involved, as Pothos suggests, it would be im-