1397A

## Event coding as feature guessing: The lessons of the motor theory of speech perception

Bruno Galantucci,[a] Carol A. Fowler,[b] and M. T. Turvey[a]
[a]*University of CT, Box U-1020, Storrs, CT 06269;* [b]*Haskins Laboratories, New Haven, CT 06511.* **bruno.galantucci@uconn.edu**
**carol.fowler@haskins.yale.edu** **turvey@uconnvm.uconn.edu**

**Abstract.** The claim that perception and action are commonly coded because they are indistinguishable at the distal level is crucial for theories of cognition. However, the consequences of this claim run deep, and the Theory of Event Coding (TEC) is not up to the challenge it poses. We illustrate why through a brief review of the evidence that led to the motor theory of speech perception.

We agree with **Hommel et al.** that unification of theories of perception and action is necessary to understand cognition. Equipping animals with efficient perceptual capabilities is expensive for nature. The enterprise would make no sense unless it was motivated by equivalent gains in action capabilities. Conversely, complex elaborations of the action capabilities of an animal would be pointless if not accompanied by perceptual systems that support them. Evolution, unlike modern cognitive science, cannot loosen the functional connections among its products. Thus, perception and action should not be studied in isolation; most likely perceiving-acting is not a decomposable system.

The main assumption of the target article is a logical corollary of this broader claim and, as such, we agree with it. However, the Theory of Event Coding (TEC) does not fulfill its promise, and it makes the very mistake it is meant to correct.

We will focus on the authors' brief, dismissive comments about the motor theory of speech perception. The authors provide two reasons for their dismissal. First, they do not wish to claim that their account applies to language. Second, in the opinions of Jusczyk (1986) and Levelt (1989), empirical support for the motor theory is weak. We address what the authors prefer not to because we consider the reasons for their dismissal mistaken.

However, we will not reply to the second reason here. Evidence from half a century of research cannot be summarized as tersely as a dismissal can be offered (for a review, see Liberman 1996). As for the first reason, speech is an exquisite example of perceptually guided action, and thus it must be addressed by a theory of perception-action codings. Moreover, precisely because it cannot be reduced to the stimulus/response experimental settings the authors choose to use, it is the right place to look for a better understanding of the couplings between perception and action. And the empirical facts that led to the motor theory reveal the flaws of TEC to which we have alluded.

In the tasks that support TEC, experimenters devise stimuli that can be described by sets of arbitrarily-chosen, arbitrarily-combined features (e.g., a letter is red or green; a rectangle is on the right or left side of a computer screen). In the tasks, some of the features are relevant, and so the participants are encouraged to make use of them to do the task, and they do. However, these sets are not up to the task of constituting percepts or action plans in nature. Proponents of TEC have to answer, among others, the following questions:

1) Can such feature sets compose real perception-action codings?

2) Are percepts linear combinations of features?

3) If they are to refer to the distal world, perception-action codings must be grounded. How are features grounded?

The motor theory of speech perception was designed to answer these questions, which arose when Alvin Liberman and colleagues sought the features that capture our perceptions of syllables (Liberman 1957; Liberman et al. 1967; Liberman & Mattingly 1989).

They started where Hommel and colleagues would have liked them to start. During the fifties, as a part of a project to build a reading machine for the blind, they tried several ways to teach people to perceive syllables as strings of featurally distinct sounds. They devised "acoustic alphabets": acoustic features corresponded to single letters, and syllables were sequences of discrete acoustic letters. After years of effort, the project failed. Speech is not an acoustic alphabet, and people could not perceive acoustic alphabetic sequences at practically useful rates.

This realization provides an answer to our first question: arbitrarily-chosen, arbitrarily-combined features do not correspond to perception-action codings.

The next move of Liberman and colleagues was also aligned with TEC. Perhaps the sounds created by the vocal tract have a special character for which arbitrary sounds cannot substitute. Using the sound spectrograph, Liberman and colleagues searched for acoustic features that, they hoped, would characterize syllables. Again, their hopes were frustrated. Spectrograms, far from clarifying the picture, presented a new puzzle. Depending on the context, the same acoustic feature could specify different phonemes and, conversely, different acoustic features could specify the same phoneme. Acoustic features do not capture the dimensions of the space where syllables live.

We can now answer our second question: percepts are not necessarily linear combinations of acoustic features. But, more importantly, this second failure taught Liberman and colleagues a lesson. Once they rejected features as components of speech percepts, they recognized that there is an invariant among different tokens of the same phoneme; it is in the gestures that produced them. This discovery led to a major revision of their scientific assumptions. Motor competence, not the feature set conceived by the scientist, underlies speech perception. The motor theory was born, and it provided a surprising answer to our third question: speech perception is not grounded in the realm of perception, but in the very place where, according to the common coding of distal events, it should be grounded: in the actions of the vocal tract!

To understand the nature of perception-action codings, a major revision of the conceptual apparatus used to implement TEC is needed, beginning with careful decisions about what should and should not count as an observable. Liberman, and with him Roger Sperry (1952), suggested that the use of two different sets of observables – one for perception and one for action – is misleading; motoric observables capture cognition. Although we do not entirely accept this full motorization of cognition, the hard learned lessons that led to the motor theory tell us that the observables conceived by a theory of disembodied cognition will not work. Here lies the major limitation of TEC. It rejects the unreasonable division between action and perception at the level of distal events, while claiming that "late action" and "early perception" are not necessary to explain cognitive common coding. If not there, among the nuts and bolts of the contact between the nervous system and the physical world, where are we to find the grounding of perception-action distal identity?