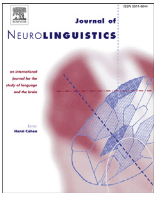
Journal of Neurolinguistics 26 (2013) 511–525

Contents lists available at SciVerse ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling



Theoretical article

Neuroethology in the service of neurophonetics

Harvey M. Sussman[a,b,*]

[a] Department of Linguistics, University of Texas, Austin, TX 78712, USA
[b] Department of Communication Sciences & Disorders, University of Texas, Austin, TX 78712, USA

A R T I C L E I N F O

Article history:
Received 3 January 2013
Accepted 28 February 2013

Keywords:
Neuroethology
Neural columns
Combination-sensitive neurons
F2 variability
Locus equations

A B S T R A C T

Single-neuron recording methods, as commonly used in neuroethology studies, provide the needed spatial and temporal resolution capacities to generate explicit hypotheses addressing the 'how' of language processing. The goal of this article is to describe two well documented neural processing mechanisms that can provide insights into (1) the auditory decoding of speech sounds, and (2) disambiguation of context-induced variability in stop place perception. The neural unit underlying speech sound processing is the combination-sensitive neuron, and the neural entity best suited to resolve context-induced variability in the speech signal is the neural column. The 'absorption' of stimulus variability via signal-specific columnar encoding is contrasted to exemplar-based treatments of stimulus variability in neural systems.

1. Introduction

Single linguistic events, such as generating a verb past tense, making a semantic association, or monitoring a phoneme target sound, occur across intervals of time best measured in milliseconds, and within neural spaces best measured in microns. A predominant methodological tool, however, used to study these language events, functional magnetic resonance imaging (fMRI), far exceeds the temporal and spatial resolutions needed to monitor these events in real time. The basic analysis unit of fMRI—the 3D-voxel, customarily varies between 3 mm^3 and 4 mm^3. This volume of neural tissue encompasses activity spanning 100,000s of neurons interconnected by millions of synapses. The resultant blood oxygenation level dependent (BOLD) signal, thought to be correlated to the linguistic event, reaches its peak 6–7 s *after* the event is over. Despite these methodological shortcomings, fMRI

---

* Department of Linguistics, University of Texas, Austin, TX 78712, USA. Tel.: +1 512 471 9002; fax: +1 512 471 4340.
E-mail address: sussman@austin.utexas.edu.

0911-6044/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jneuroling.2013.02.004

investigations have greatly enhanced our knowledge of 'where' in the brain linguistic operations might occur (e.g. Hickok & Poeppel, 2004). The more difficult challenge is understanding '*how*' these brain structures do what they do. The brain is, after all, an alien structure, in the sense that we have a very limited understanding of how cognitive-based events happen within this amazingly complex and other worldly tissue.

The purpose of this article is not to criticize the use of fMRI in language studies, but rather to enlighten a subset of linguists to the existence of another source of brain-related data that can possibly provide insights into '*how*' language processes might be carried out by neural entities. The field of neuroethology—the study of animal communication, primarily uses single-neuron electrophysiological recordings to discover the acoustic elements of species-specific input sounds a neuron is specialized to detect and process. Data based on single-cell recordings are the gold standard in terms of maximizing both temporal and spatial resolutions in monitoring neural activity.

1.1. Rationale for using an animal model to study the human model

A neuroethological-based approach to study human language processing has limitations of its own to resolve. There must be compatibility between the level of language structure examined and the data emanating from neuroethology laboratories. For example, single-cell electrophysiological recordings from animal brains cannot inform us as to how syntactic or semantic operations are carried out in human brains. Compatibility does exist, however, at the level of auditory processing. The laws of physics dictate the structure of sound, whether they be shaping a species-specific call, biosonar echos, or modulating a second formant transition.

At the outset, I would like to dispel potential skepticism from those who frown upon the notion of using an animal model to provide a theoretical springboard for conceptualizing language in the human model. It is safe to conclude that the human brain is a product of evolution. It may not be elegantly designed, or operate as an energy efficient device, but, in spite of its many shortcomings, it gets the job done (Linden, 2007). Secondly, and most importantly, evolution tends to have similar solutions for similar problems. The 'conservation of mechanism' principle (Gerhart & Kirschner, 1997) states that once mechanisms are invented by nature and serve successfully, those mechanisms are retained during evolution and may be modified and thereby adapted for new and higher-order processing. A third point relates to the fact that there are more similarities than differences in both the structure and function of auditory neural systems across species biologically specialized for processing sounds. Us humans can take our rightful place alongside crickets, rats, bats, frogs, birds, barn owls, and monkeys.

2. Neuroethology findings with phonetic-related implications

2.1. Combination-sensitive neurons: synchronous processing of multiple sound parameters

Neuroethology studies have uncovered, across a wide variety of species, a ubiquitous type of auditory neuron specialized for spectral integration—the combination-sensitive neuron. These neurons are specifically "*tuned to coincidence (synchronization) of impulses from different neurons in time, frequency and/or amplitude domains*" (Suga, 1994, p. 135). The frequency differences between two arriving signals can span up to three octaves, and the temporal discrepancy between the two inputs can vary from zero to tens of milliseconds (Suga, 1989). The earliest stage of auditory processing that has recorded activity of combination-sensitive neurons in sound processing is the midbrain central nucleus of the inferior colliculus (Mittman & Wenstrup, 1995; Portfors & Wenstrup, 1999). The afferent inputs to these inferior collicular cells have been shown to project from the ipsilateral ventral and intermediate nuclei of the lateral lemniscus (Yavuzoglu, Schofield, & Wenstrup, 2011).

Combination-sensitive neurons have been documented across: (i) frogs (Fuzessery & Feng, 1983; Mudry, Constantine-Paton, & Caprnica, 1977); (ii) birds (Margoliash, 1983; Margoliash & Fortune, 1992; Takahashi & Konishi, 1986); (iii) mammals: mustached bats (Olsen & Suga, 1991a, 1991b; Suga, O'Neill, Kujirai, & Manabe, 1983; Suga, O'Neill, & Manabe, 1978); brown bats (Neuweiler, 1983; Neuweiler, 1984); mouse (Hoffstetter & Ehret, 1992); cat (Sutter & Schreiner, 1991); and primates (Kadia & Wang, 2003; Olsen, 1994; Olsen & Rauschecker, 1992). The response characteristics of combination-