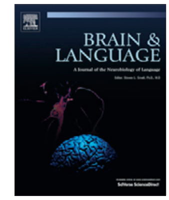
Brain & Language 125 (2013) 60–76

Contents lists available at SciVerse ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l



Reconciling time, space and function: A new dorsal–ventral stream model of sentence comprehension

Ina Bornkessel-Schlesewsky [a,*], Matthias Schlesewsky [b]

[a] Department of Germanic Linguistics, University of Marburg, Marburg, Germany
[b] Department of English and Linguistics, Johannes Gutenberg-University, Mainz, Germany

ARTICLE INFO

Article history:
Accepted 15 January 2013
Available online 26 February 2013

Keywords:
Language comprehension
Dorsal stream
Ventral stream
Hierarchical organisation
Cognitive control
Syntax
Semantics
Inferior frontal gyrus
Anterior temporal lobe
Posterior temporal lobe

ABSTRACT

We present a new dorsal–ventral stream framework for language comprehension which unifies basic neurobiological assumptions (Rauschecker & Scott, 2009) with a cross-linguistic neurocognitive sentence comprehension model (eADM; Bornkessel & Schlesewsky, 2006). The dissociation between (time-dependent) syntactic structure-building and (time-independent) sentence interpretation assumed within the eADM provides a basis for the division of labour between the dorsal and ventral streams in comprehension. We posit that the ventral stream performs time-independent unifications of conceptual schemata, serving to create auditory objects of increasing complexity. The dorsal stream engages in the time-dependent combination of elements, subserving both syntactic structuring and a linkage to action. Furthermore, frontal regions accomplish general aspects of cognitive control in the service of action planning and execution rather than linguistic processing. This architecture is supported by a range of existing empirical findings and helps to resolve a number of theoretical and empirical puzzles within the existing dorsal–ventral streams literature.

1. Introduction

The literature on the neuroscience of language has recently seen an increasing interest in the dorsal and ventral streams as possible, neurobiologically plausible streams of speech and language processing.[1] Evidence for this perspective has been gleaned from a number of different domains, ranging from speech perception and production (Hickok & Poeppel, 2004, 2007; Rauschecker & Scott, 2009) over word-level production and comprehension (Ueno, Saito, Rogers, & Lambon Ralph, 2011) to syntactic processing (Friederici, 2009). An inherent appeal of the dual streams perspective is that it may help to provide a neurobiological grounding for functionally motivated models of the language architecture. In particular, as dual streams of processing are well established within the literature on the auditory system of non-human primates, they open up the possibility for highly appealing cross-species comparisons between human speech and language and more general properties of auditory processing (Rauschecker & Scott, 2009).

However, in spite of the relatively unified neuroanatomical perspective underlying these current dual streams approaches to language (but see below for differing assumptions regarding possible neuroanatomical sub-pathways and the characterisation of posterior temporal regions), their interpretations of dorsal and ventral stream functions in language processing are quite different from one another. For example, based on studies of pseudoword production versus sentence comprehension, Saur et al. (2008, p. 18035) proposed that the dorsal stream mediates the "sensory-motor mapping of sound to articulation", while the ventral stream is involved in the "linguistic processing of sound to meaning". By contrast, Friederici (2009, 2012) draws upon results from sentence comprehension to posit that part of the dorsal stream (specifically, one dorsal sub-pathway) is crucial for the processing of "hierarchical" or "complex" syntax, whereas part of the ventral stream (one ventral sub-pathway) is assumed to be involved in the processing of "local" or "simple" syntax. Clearly, these alternative functional proposals have very different implications for the interpretation of the dorsal and ventral streams during language processing and, thereby, for models of the neurobiology of language. However, beyond these specific interpretations, are there possible unifying (and meaningful) functional generalisations that dissociate one stream from the other, irrespective of the possible existence of neuroanatomical sub-pathways?

* Corresponding author. Address: Department of Germanic Linguistics, University of Marburg, Deutschhausstrasse 3, 35032 Marburg, Germany. Fax: +49 (0)6421 2824558.

E-mail address: iboke@staff.uni-marburg.de (I. Bornkessel-Schlesewsky).

[1] Here and in the following, we use the term "stream" to denote a functional route of information processing and the term "pathway" to denote neuroanatomical connectivity. As will become clear throughout the remainder of the paper, we view the correspondence between the two as correlative in nature. Crucially, this implies that there need not be a 1:1 correspondence between streams and pathways (including, for example, a many-to-one mapping of pathways to streams).

0093-934X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandl.2013.01.010

Here, we approach this question from a novel perspective. Specifically, we attempt to bring together some basic neurobiological design principles regarding information processing within the two streams (Rauschecker & Scott, 2009) with insights on the functional architecture of sentence comprehension. We will argue that the assumption of hierarchical processing – the sensitivity for increasingly complex sets of feature combinations within neurons or neuronal assemblies – as a basic principle of brain function within the auditory system (as suggested by Rauschecker, 1998)[2] can be fruitfully combined with well-established assumptions regarding the timing of language comprehension. This assumed correspondence between a neuroanatomical hierarchy and a temporal hierarchy in information processing will be used as a basis for a new spatio-temporal model of language processing within a dorsal and ventral streams perspective.

The remainder of the paper is organised as follows. Section 2 begins by introducing some background assumptions from the neurobiological domain and sentence comprehension in time and space. Section 3 subsequently goes on to describe some puzzles that arise if these background assumptions are adopted. Section 4 offers a possible solution to the puzzles described in Section 3 in the form of a novel proposal regarding the neuroanatomical locus of syntactic structure building and the form-to-meaning mapping at the sentence level. Finally, Section 5 offers some conclusions.

2. Background assumptions

In this section, we will describe the assumptions on which our line of argumentation will be based in the following sections. While each of these assumptions can, presumably, be contended at some level, it seems to us that they are all established sufficiently to warrant their use as premises of the account to be developed here.

2.1. Hierarchical organisation as a basic property of functional neuroanatomy

On the basis of research on the visual (e.g. Felleman & Van Essen, 1991) and, more recently, auditory systems (Rauschecker, 1998), we follow Rauschecker and Scott (2009) in assuming that the functional neurонanatomy of information processing in the brain is hierarchically organised:

> Hierarchical organization in the cerebral cortex combines elements of serial as well as parallel processing: 'lower' cortical areas with simpler receptive-field organization, such as sensory core areas, project to 'higher' areas with increasingly complex response properties, such as belt, parabelt and PFC regions. These complex properties are generated by convergence and summation [...]. Parallel processing principles in hierarchical organization are evident in that specialized cortical areas ('maps') with related functions (corresponding to sub-modalities or modules) are bundled into parallel processing 'streams'. (Rauschecker & Scott, 2009, p. 719)

Evidence for hierarchical organisation within the auditory system stems from a variety of different sources. Using single cell recordings in non-human primates (rhesus monkeys), Rauschecker and colleagues found increasing sensitivity to more complex "auditory objects" – from neurons responding mainly to specific frequency bandwidths in lateral auditory belt areas to neurons responding increasingly to species-specific vocalisations in more anterior portions of the superior temporal gyrus (Rauschecker, Tian, & Hauser, 1995). From these "increasing proportions of call-selective neurons [...] from A1 to lateral belt to more anterior superior temporal areas" (Rauschecker, 1998, p. 518), Rauschecker proposed a hierarchical organisation of auditory processing that is compatible with what is known about hierarchical processing within the visual system. This perspective was recently corroborated by a meta-analysis of neuroimaging studies on language processing, in which DeWitt and Rauschecker (2012) found evidence for an anterior-directed processing gradient within temporal cortex. Across 115 studies, phoneme versus word processing engendered increasingly anterior activation within the superior temporal gyrus (STG), and phrase-level processing correlated with activation in the anterior superior temporal sulcus (STS). From these findings, DeWitt and Rauschecker (2012) argue for a concordance between the results on human language processing and the literature on primate auditory processing, with both providing evidence for hierarchical processing of auditory objects within a ventral processing stream in superior temporal cortex.

Applying these basic assumptions to language processing – and, for present purposes, sentence processing in particular – we arrive at the following hypothesis: the functional neuroanatomy of the form-to-meaning mapping should be characterised by neuroanatomical gradients originating in primary auditory areas, which correlate with the processing of successively more complex linguistic units.

2.2. Time–space correspondence

If we accept the premise that language processing is supported by a hierarchically organised auditory system (see Section 2.1), this also has implications for the temporal organisation of the form-to-meaning mapping in sentence processing. Of course, connectivity within the brain is inherently bidirectional. Nevertheless, the assumption of hierarchical organisation implies that there is a certain asymmetry in the "flow" of information, since "lower" areas with simple feature sensitivity project to "higher" areas with a sensitivity to more complex stimuli, resulting from the convergence and summation of properties from a number of "lower" areas (see the quote by Rauschecker & Scott, 2009, in Section 2.1 above). DeWitt and Rauschecker (2012, p. E509), too, refer to "a processing cascade emanating from core areas, progressing both laterally, away from core itself, and anteriorly, away from A1" in describing their ventral stream of linguistic pattern recognition (i.e. language comprehension from the phonemic to the phrasal level). We thus propose that insights on the organisation of the neuroanatomical processing hierarchy should be compatible with findings on the temporal organisation of sentence processing and vice versa.

This hypothesis can be exemplified using the gradient of phonemic processing to word processing that was observed by DeWitt and Rauschecker (2012): electrophysiological (i.e. scalp EEG) studies of auditory word recognition in sentence context have provided evidence for two mismatch-related negativities that occur when the current input is incongruent with the prior sentence or discourse context, an N200 and a following N400 (e.g. Connolly & Phillips, 1994; van den Brink, Brown, & Hagoort, 2001). Based on these results, even proponents of a highly interactive "one-step" model of sentence-level interpretation (Hagoort, 2005; Hagoort & van Berkum, 2007) have argued for a cascade of information processing during sentence comprehension (Hagoort, 2008; van den Brink, Brown, & Hagoort, 2006; van den Brink et al., 2001). According to this view, word recognition comprises the activation of a cohort of word candidates (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978) in a strictly bottom-up manner, with N200 effects emerging whenever a form-based lexical candidate is not supported by the current

[2] Note that hierarchical processing in this sense is not to be confused with hierarchical syntax in the sense of Friederici (2009). Friederici uses the term "hierarchical" to refer to particular types of syntactic structures, in contrast to the neurobiological sense that is central here. For more detailed discussions of the two senses of the term "hierarchical", see Sections 2 and 3 for the neurobiological and Friederician sense, respectively.