The Cross-Linguistic Prevalence of SOV and SVO Word Orders Reflects the Sequential and Hierarchical Representation of Action in Broca’s Area

David Kemmerer*
Departments of Speech, Language, and Hearing Sciences, and Psychological Sciences, Purdue University, Department of Neurology, Division of Cognitive Neuroscience, University of Iowa College of Medicine

Abstract

Despite the increasingly interdisciplinary nature of the language sciences, so far relatively little effort has been devoted to exploring potential connections between typology and neuroscience. To illustrate some of the insights that can be gained from pursuing such an integration, this paper focuses on one of the most well established and frequently cited typological generalizations, namely that in the vast majority of human languages, the basic word order is either SOV (about 48%) or SVO (about 41%). It has been suggested that these strong tendencies can be explained cognitively in terms of the prototypical transitive action scenario, in which an animate agent acts forcefully on an inanimate patient to induce a change of state. Two forms of iconicity are especially relevant: first, because the agent is at the head of the causal chain that affects the patient, subjects usually precede objects; and second, because it is the agent’s action, rather than the agent per se, that changes the state of the patient, verbs and objects are usually adjacent. The purpose of this paper is to show that this account converges with, and hence receives further support from, recent research on how actions are represented in the brain. Specifically, several lines of evidence are reviewed which suggest that Broca’s area plays a pivotal role in schematically representing the sequential and hierarchical organization of goal-directed bodily movements, not only when they are performed and perceived in the real world, but also when they are conceptualized and symbolically expressed as transitive clauses. Taken together, these findings support the hypothesis that the most cross-linguistically prevalent word order patterns reflect the most natural ways of linearizing and nesting the core conceptual components of actions in Broca’s area.

1. Introduction

Ever since the “cognitive revolution” took place in the 1950s, one of the most ambitious goals of the language sciences has been to understand how universal aspects of phonology, morphology, syntax, and semantics can be related to – and, to some extent, explained in terms of – universal aspects of the human brain. Despite the rapid growth of neurolinguistics during the past few decades, however, little headway has been made toward achieving this aim, largely because the challenges that must be overcome are extremely daunting (Poeppel and Embick 2005). One problem that has not received as much attention as it probably deserves is that, as the language sciences have advanced, it has actually become harder, rather than easier, to determine which aspects of the uniquely human capacity for language are most likely to have species-typical neurobiological substrates. This uncertainty stems from the fact that extensive typological research on the roughly 7,000 languages in the world has generated increasing evidence that the degree
of diversity is far greater than previously assumed (Evans and Levinson 2009). As a consequence, there appear to be few universal properties of language that one could reasonably hope to link with universal properties of the brain. Nevertheless, a number of strong cross-linguistic tendencies have been identified – e.g., the CV syllable, recursive syntax, words like red and arm, etc. – and while some of them may be due to cultural/historical factors, others may be due instead to powerful cognitive biases or “attractors” with robust neural underpinnings. So far, however, relatively few attempts have been made to use these specific sorts of typological considerations to guide empirical and theoretical work in neuroscience (e.g., Giraud et al. 2007; Bornkessel-Schlesewsky and Schlesewsky 2009a,b; Kemmerer 2006, 2010; Kemmerer and Eggleston 2010).

The purpose of this paper is to explore one particular way in which the gap between typology and neuroscience might be bridged. The central phenomenon is as follows: Although there are six logically possible sequences for the subject, object, and verb in a transitive clause, it is well established that languages overwhelmingly gravitate toward just two: SOV and SVO. Section 2 first summarizes the evidence for these word order preferences, and then it describes a semantically oriented account which maintains that the dominant linearization patterns reflect, in an iconic or isomorphic manner, the natural flow of energy from agent to patient in the prototypical transitive action scenario. For example, in the sentence Bill crushed the can, the precedence of the subject over the object mirrors the temporal structure of the causal chain that is expressed, and the contiguity of the verb and the object mirrors the tight bond between the action and its effect. Next, Section 3 shows how this cognitive account of the word order data dovetails with a substantial body of research on the high-level representation of action in the brain. Specifically, a number of recent studies are reviewed which suggest that one of the major computational hubs for language, namely Broca’s area, is integrally involved in representing, at a relatively abstract level of analysis, the sequential and hierarchical organization of goal-directed bodily movements, not only when they are performed and perceived in the real world, but also when they are conceptualized and symbolically expressed as transitive clauses. These findings invite the inference that the cross-linguistic prevalence of SOV and SVO word orders derives from the adaptive role of Broca’s area in capturing certain skeletal aspects of action concepts – most importantly, the ways in which the nested causal relationships among the core participants unfold in time.

2. The Cross-Linguistic Prevalence of SOV and SVO Word Orders

2.1. Typological Patterns

Word order has figured prominently in typology ever since the seminal work of Greenberg (1963). Numerous discoveries about cross-linguistic tendencies have been made (for a recent review see Dryer 2007), and some of those discoveries have inspired innovative theories about the evolution and processing of syntax (e.g., Hawkins 1994, 2010; Kirby 1999; Newmeyer 2000; Gell-Mann and Ruhlen 2011). As indicated above, this paper focuses on one of the most fundamental topics in word order research: the sequencing of the subject, object, and verb in a transitive clause. In ordinary language use, at least one of the two clausal arguments, either the subject or the object, is often a pronominal element (sometimes expressed as a verbal affix). Here, however, the emphasis is on those cases in which both arguments are full-fledged nominal elements.

Before presenting the key data, it is worthwhile to briefly mention some methodological and definitional points. It is generally assumed that if a given language has a basic or
dominant pattern for sequencing the subject, object, and verb in a transitive clause, that pattern will satisfy the following criteria, among others (Dryer 2007). (1) **Frequency**: The basic word order is the one that is used most often, as revealed, for instance, by corpus analyses. (2) **Markedness**: The basic word order is the one with the least amount of function-indicating phonological, morphological, or syntactic marking. (3) **Pragmatic neutrality**: The basic word order is the one that carries no special pragmatic information apart from declarative mood.

By far the most comprehensive typological analysis of the sequencing of subject, object, and verb was conducted by Dryer (2011), who combined data from a global sample of 1377 languages. 1188 (86%) of those languages reportedly have a basic word order, and all six of the logically possible linearizations of subject, object, and verb are attested. As shown in Table 1, however, there is by no means an even distribution across the six types of order, since the vast majority of languages are, in roughly equal proportions, either SOV (565, 48%) or SVO (488, 41%). The other four types are much less common, ranked in descending frequency from VSO (95, 8%) to VOS (25, 2%), OVS (11, 1%), and OSV (4, 0.5%). Notably, more or less similar breakdowns across the six ordering patterns were documented in previous studies with much smaller databases (Greenberg 1963; Tomlin 1986).

Additional evidence that SOV and SVO word orders are overwhelmingly preferred comes from several other sources. First, Kimmelman (forthcoming) found that in a sample of 24 sign languages, 21 (88%) have SOV and/or SVO as the dominant sequencing pattern(s). Second, within the last 70 years, Al-Sayyid Bedouin Sign Language (ABSL) has gradually arisen in an isolated community with a high incidence of genetically based prelingual deafness, and in the space of a single generation, it assumed a grammatical structure characterized by SOV order (Sandler et al. 2005). Given that none of the neighboring spoken or signed languages are SOV, this property of ABSL presumably developed spontaneously. Third, Goldin-Meadow et al. (2008) asked speakers of three SVO languages (English, Spanish, and Mandarin) and one SOV language (Turkish) to perform two non-verbal tasks: first, describe events using manual gestures without speech; and second, reconstruct events illustrated in pictures. The investigators found that in both tasks all of the participants were strongly inclined to use the same sequencing strategy – specifically, agent-patient-action, which is analogous to the SOV pattern in spoken languages. Taken together, these three sets of results support the view that SOV and SVO word orders – perhaps especially the former – reflect the most cognitively natural ways of linearizing the fundamental elements in a transitive clause.

Table 1. Basic order of subject, object, and verb in a sample of 1188 languages. Note that these languages are drawn from a larger sample of 1377, and that the remaining 189 languages lack a basic order (from Dryer 2011).

<table>
<thead>
<tr>
<th>Basic order</th>
<th>Number</th>
<th>Percentage</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject–object–verb (SOV)</td>
<td>565</td>
<td>48</td>
<td>Japanese</td>
</tr>
<tr>
<td>subject–verb–object (SVO)</td>
<td>488</td>
<td>41</td>
<td>Mandarin</td>
</tr>
<tr>
<td>verb–subject–object (VSO)</td>
<td>95</td>
<td>8</td>
<td>Irish</td>
</tr>
<tr>
<td>verb–object–subject (VOS)</td>
<td>25</td>
<td>2</td>
<td>Nias; Austronesian, Sumatra</td>
</tr>
<tr>
<td>object–verb–subject (OVS)</td>
<td>11</td>
<td>1</td>
<td>Hixkaryana; Carib, Brazil</td>
</tr>
<tr>
<td>object–subject–verb (OSV)</td>
<td>4</td>
<td>0.5</td>
<td>Nadêb; Vaupés-Japurá, Brazil</td>
</tr>
</tbody>
</table>
Several typologists have observed that the cross-linguistic prevalence of SOV and SVO word orders can be explained by two general principles: subject salience, which states that subjects tend to precede objects; and verb–object contiguity, which states that verbs and objects tend to be adjacent (e.g., Greenberg 1963; Tomlin 1986; Comrie 1989; see also Whaley 1997, 83–5). Both principles are clearly preserved by the two most common ordering patterns. It has been suggested, however, that the first principle – subject salience – may carry more weight than the second principle – verb–object contiguity – because the third most common ordering pattern, VSO, preserves the former principle at the cost of violating the latter (Song 1991). In the same vein, Greenberg (1963) regarded the subject salience principle as being of such paramount importance that he ascribed it the status of Universal #1: “In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.”

It is quite likely that the two principles have semantic, and ultimately iconic, bases. Transitive clauses can express a tremendous range of situation types, including many that have nothing to do with agentive behavior, as in This line parallels that line (for additional examples see Jackendoff 2002, 135–6). However, it is often assumed that their most basic function is to encode the so-called prototypical transitive action scenario, in which an animate agent performs an action that causes an inanimate patient to undergo a change of state, as in the sentence invoked earlier, Bill crushed the can (e.g., Hopper and Thompson 1980; Næss 2007). In the literature on event lexicalization and argument realization, several closely related theories maintain that this sort of scenario is conceptualized in terms of a causal chain that links the core participants through force–dynamic energy transmission. All of these theories decompose the whole event into two causally related subevents, the first of which involves the agent’s activity, and the second of which involves the patient’s transformation. The theories differ, however, in various ways (Levin and Rappaport Hovav 2005). For instance, some of them place more emphasis on hierarchical than sequential organization and favor a formal representational scheme like the following: [[Bill ACT] CAUSE [BECOME [crushed, can]]] (e.g., Jackendoff 1990; Rappaport Hovav and Levin 1998; Van Valin and LaPolla 1997; Van Valin 2005). In contrast, other theories place more emphasis on sequential than hierarchical organization and favor a more depictive representational scheme like that shown in Figure 1 (e.g., Croft 1991, 1998; Langacker 1991, 2008). Transcending these differences between the two approaches, however, are certain shared features that seem to provide a solid cognitive grounding for the typological principles of subject salience and verb–object contiguity.

First, the subject salience principle may derive in large part from the fact that in the prototypical transitive action scenario, the agent is at the head of the causal chain that affects the patient. For instance, the sentence Bill crushed the can most naturally describes

![Fig 1. A schematic illustration of the prototypical transitive action scenario, in which an animate agent (AG) transmits energy (double-lined arrow) to an inanimate patient (PAT) and thereby induces a telic change of state (single-lined arrow followed by square). IS = immediate scope of conceptualization (from Langacker 2008, 357).](image-url)
an event in which Bill initially has the intention to crush the can, and then realizes that intention by moving part of his body toward the can, contacting it, and forcefully compressing it. In this context, it is also noteworthy that the human species is evolutionarily equipped with powerful neurocognitive mechanisms dedicated to rapidly detecting and understanding the willful behavior of animate entities (Epley and Waytz 2009; Frith and Frith 2010). For present purposes, however, the essential point is that the temporal precedence of the agent’s volitional action over the patient’s resultant transformation is captured, in an iconic or isomorphic way, by the temporal precedence of the subject nominal over the object nominal.

Second, the verb–object contiguity principle may originate from another prominent aspect of the prototypical transitive action scenario, namely that it is the agent’s action, rather than the agent per se, that changes the state of the patient. Note, for instance, that operations like passivization (e.g., The can was crushed) and compounding/ incorporation (e.g., Can-crushing can be fun) allow speakers to downplay the agent so as to highlight the effect on the patient. And even when the agent is explicitly referred to, many transitive action verbs seem to provide more information about the effect on the patient than about the movement of the agent. This is illustrated by the fact that one can crush a can by stepping on it, squeezing it with one’s hands, or acting on it in countless other ways. More generally, in a survey of approximately 1900 instrumental verbs in English, Koenig et al. (2008) found that “verbs that require or allow instruments constrain the end states of the situations they describe more than they constrain the agent’s activity” (175; see also the cross-linguistic study of “cutting” and “breaking” verbs by Majid et al. 2008). The upshot is that although the agent is obviously an essential participant in the transitive action scenario, there seems to be an especially strong conceptual bond between the force-dynamic event and the patient that is transformed by it. Moreover, this close semantic connection between the action and its consequences is captured, in an iconic or isomorphic way, by the close syntactic connection between the verb and its object. Importantly, these semiotic considerations do not predict any particular order of the verb and its object, but they do predict that these elements should be adjacent (Haiman 1985).

In sum, the overwhelming cross-linguistic preference for SOV and SVO word orders is arguably the outcome of two main principles that iconically reflect certain conceptual aspects of the prototypical transitive action scenario. According to the subject salience principle, subjects usually precede objects – a syntactic pattern that mirrors the temporal profile of the flow of energy from agent to patient. And according to the verb–object contiguity principle, verbs and objects are usually adjacent – a syntactic pattern that mirrors the tight causal relationship between the agent’s action and its effect on the patient. This cognitive account of the word order data clearly has many virtues. But the next section shows that it can be strengthened and enriched even more by being integrated with corresponding neuroscientific discoveries about the central role that Broca’s area plays in representing the sequential and hierarchical organization of goal-directed actions.

3. Linking the Cognitive Account with Neuroscientific Research on Broca’s Area

3.1. Some Background on Broca’s Area

Anatomically, Broca’s area is traditionally regarded as comprising Brodmann areas (BAs) 44 and 45 in the left inferior frontal gyrus (IFG; Figure 2). Although these areas are defined cytoarchitectonically (i.e., in terms of the presence/absence, packing, and layering of specific cell types), they correspond roughly to two macroscopically conspicuous
components of the left IFG – in particular, BA44 occupies the pars opercularis, while BA45 occupies the pars triangularis. Hagoort (2005) has argued that another sector of the left IFG – BA47, which occupies the pars orbitalis – should be grouped together with the other two to form “Broca’s complex,” but this proposal remains controversial. All three areas lie anterior to the ventral primary motor (BA4) and premotor (BA6) cortices, and they differ from each other not only cytoarchitectonically, but also in terms of their connectivity with other regions in the frontal, temporal, and parietal lobes (e.g., Anwander et al. 2007; Xiang et al. 2010). Moreover, recent analyses of neurotransmitter receptor distributions have begun to disclose complex chemical parcellations within and between the three areas (Amunts et al. 2010).

Functionally, Broca’s area has been associated with language processing since 1861, when the famous French neurologist for whom the region is named, Paul Broca, first reported that severe speech production deficits could be linked with damage to this territory. During the past 150 years, Broca’s area has been found to contribute not just to speech production, but to many other aspects of language as well (for a recent survey see Grodzinsky and Amunts 2006). Even more intriguing, however, is that the latest wave of research on Broca’s area has been progressively demonstrating that it is also involved in several non-linguistic domains, including imitation (e.g., Iacoboni et al. 1999), music (e.g., Maess et al. 2001), and visuospatial perception (e.g., Bahlmann et al. 2009). These findings have prompted a search for a common functional denominator, and some of the major candidates currently being debated are cognitive control (e.g., Novick et al. 2010), sequence processing (e.g., Gelfand and Bookheimer 2003), and hierarchical processing (e.g., Koechlin and Jubault 2006).

Because my aim is to provide a neurobiological foundation for the cognitive account of the word order data described above, I will draw from the abundant literature on Broca’s area in a selective and systematic manner. The next four subsections review

---

**Fig 2. Cytoarchitectonic map of Brodmann areas (BAs) on the lateral surface of the left hemisphere.** BAs 44 and 45 are traditionally regarded as comprising Broca’s area; some scholars argue, however, that BA47 should be grouped together with the other two to form “Broca’s complex.” Note that the deep sulcal portions of these areas are not visible here. Abbreviations: ab = ascending branch of the lateral fissure; cs = central sulcus; hb = horizontal branch of the lateral fissure; ifs = inferior frontal sulcus; lf = lateral fissure; prcs = precentral sulcus (from Amunts et al. 2010, 2).
several closely related lines of work which suggest that the posterior part of Broca’s area – i.e., BA44 – is essential for (1) representing actions during both execution and observation, (2) representing actions at a conceptual level that is body-part-independent, (3) representing specifically the sequential and hierarchical organization of action concepts, and (4) representing the canonical linearization of action concepts in transitive clauses. Then the concluding section of the paper steps back from the empirical details to elaborate the idea that these remarkable properties of BA44 capture precisely those aspects of the prototypical transitive action scenario that motivate the cross-linguistic preference for SOV and SVO word orders.

3.2. MIRROR NEURONS

Mirror neurons are a unique class of brain cells that discharge not only when certain kinds of actions are executed by the self, but also when they are seen or heard being performed by someone else (Fogassi and Ferrari 2011). Thus, mirror neurons appear to represent behavioral patterns per se, regardless of the self/other distinction. These types of cells have been found in many sectors of the frontal and parietal lobes of the macaque monkey, but they were first identified in a region called area F5c, which, importantly, is assumed by many researchers to be homologous with human BA44 (for a summary of supporting evidence see Arbib and Bota 2006, 153; for a critique see Toni et al. 2008). As indicated by Rizzolatti and Sinigaglia (2008, 46), the mirror neurons in area F5c fall into several classes, with some coding

the general goal of the act (holding, grasping, breaking, etc.), others the manner in which a specific motor act can be performed (precision grip, finger prehension, etc.), and lastly, there is a group that designates the temporal segmentation of the motor act into its elementary movements (opening and closing of the hand).

Although mirror neurons have not yet been found at the cellular level in BA44, there is already substantial evidence that this component of Broca’s area contributes to both the production and the perception of actions. The relevant data come from multiple brain mapping techniques, including positron emission tomography (PET; e.g., Rizzolatti et al. 1996), functional magnetic resonance imaging (fMRI; e.g., Kilner et al. 2009), magnetoencephalography (MEG; e.g., Nishitani and Hari 2000), transcranial magnetic stimulation (TMS; e.g., Pobric and de C. Hamilton 2006), and neuropsychology (e.g., Pazzaglia et al. 2008). To take a concrete example, consider the fMRI study by Kilner et al. (2009). In each trial of this experiment, the participants either executed or observed one of two goal-directed hand actions – a precision grip or an index finger pull. When the brain activity associated with executing both actions was conjoined with that associated with observing both actions, significant spatial overlap was found in BA44 bilaterally (Figure 3a). This outcome suggests that the same neuronal populations in BA44 were engaged during execution and observation, but it does not necessarily force such a conclusion, because separate neuronal populations within the volume of spatial overlap could have been engaged instead. To overcome this limitation, the researchers took advantage of a neurophysiological phenomenon called adaptation or repetition suppression, since it is one way in which the fMRI signals evoked by different tasks can be confidently attributed to the same neuronal populations. The basic assumption, which is well supported, is that if a given neuronal population codes for a specific type of information, its response will decrease when that information is repeated (Grill-Spector et al. 2006). Using this approach, the researchers discovered that the signals in several patches of BA44 bilaterally
were significantly suppressed when an executed action was immediately followed by the same vs. a different observed action, and when an observed action was immediately followed by the same vs. a different executed action (Figure 3b). These cross-modal adaptation results are quite impressive and valuable, since they constitute powerful evidence that certain neuronal populations in BA44 are in fact tuned to certain kinds of actions, irrespective of whether those actions are produced or perceived.
3.3. ACTION CONCEPTS

Several studies suggest that BA44 contributes to the processing of actions not only at the level of execution and observation, but also at the level of conceptual knowledge. Some of the strongest evidence for this view comes from neuropsychology (e.g., Tranel et al. 2003; Saygin et al. 2004; Kemmerer et al. forthcoming). For instance, in one of the largest investigations of this topic to date, Kemmerer et al. (forthcoming) administered a battery of six standardized tasks to 226 brain-damaged patients with widely distributed lesions in the left and right hemispheres. The tasks probed conceptual knowledge of actions in a variety of verbal and nonverbal ways, including picture naming, word-picture matching, attribute judgments involving both words and pictures, and associative comparisons involving both words and pictures. The relationships between the patients’ behavioral performances and their lesion sites were carefully analyzed, and a key finding was that one of the few lesion sites that was consistently linked with impaired vs. normal performances across all six tasks was Broca’s area. The most significant effects appeared in BA47, but BA45 and, more importantly, BA44 were also clearly implicated (Figure 3c).

What sort of functional role does BA44 play in the conceptual processing of actions? Some tantalizing hints come from an influential fMRI study by Tettamanti et al. (2005) in which the participants listened passively to four types of Italian sentences that were syntactically equivalent but described different kinds of situations: leg/foot actions (e.g., Calcio il pallone “I kick the ball”); arm/hand actions (e.g., Afferro il coltello “I grasp a knife”); mouth actions (e.g., Mordo la mela “I bite an apple”); and psychological states (e.g., Apprezzo la sincerita “I appreciate sincerity”). Two main results are relevant here. First, when the investigators contrasted the three types of action sentences against the psychological sentences, they found that the left premotor cortex was engaged in a somatotopic manner, such that the leg/foot sentences activated a dorsal area associated with leg/foot movements, the arm/hand sentences activated a lateral area associated with arm/hand movements, and the mouth sentences activated a ventral area associated with mouth movements. These results suggest that understanding action sentences involves body-part-specific motor simulations of the designated behaviors, and similar findings have emerged in several other studies employing diverse methods (for reviews see Kemmerer and Castillo 2010 and Willems and Casasanto 2011). Second – and, in the current context, even more interestingly – the investigators discovered that all three types of action sentences engaged BA44 significantly more than the psychological sentences. Because the former sentences had the same syntactic structure as the latter sentences, the only distinguishing factor was semantic content; and because the former sentences varied with regard to which body parts were used to carry out the designated behaviors, the only common semantic content was body-part-independent. In fact, what the three types of action sentences shared, and what separated them from the psychological sentences, was that they all exemplified the prototypical transitive action scenario, in which an animate agent induces a change of state in an inanimate patient. Hence, Tettamanti et al.’s (2005) fMRI study suggests that BA44 may contribute to the processing of action concepts by capturing the skeletal structure of this general scenario. It is worth emphasizing that the key factor does not appear to be animacy per se, since that was equivalent across the four sentence types. Instead, what BA44 seems to be sensitive to is the schematic structure of goal-directed action.

Another fMRI study, this one conducted by Baumgaertner et al. (2007), both reinforces and refines this idea. The participants in this experiment made sensibility judgments about goal-directed arm/hand actions that were presented in two ways – linguistically as spoken German sentences, and visually as dynamic video clips. Two additional conditions
involving sentences and videos about inanimate motion events were included to control for action specificity. The investigators found that BA44 was engaged when the action sentences were contrasted against the non-action sentences, when the action videos were contrasted against the non-action videos, and, finally, when the results of the former contrast were conjoined with those of the latter (Figure 3d). As in Tettamanti et al.’s (2005) study, the action and non-action sentences were syntactically similar but semantically different. In particular, all of the action sentences described situations in which an animate agent causes an inanimate patient to change state by operating on it with tool (e.g., Er rührt mit einem Löffel “He is stirring with a spoon”), whereas all of the non-action sentences described situations in which an inanimate entity moves as a result of natural forces (e.g., Die Blätter wirbeln durch die Luft “The leaves are swirling through the air”). It is noteworthy that although the former sentences were not grammatically transitive, they could easily have been made transitive through the insertion of direct object nominals. This would have been possible because all of them encoded purposive actions characterized by a complex sequential and hierarchical structure in which the agent’s ultimate goal is achieved by first performing an embedded or intermediary action involving tool manipulation. It is also worth emphasizing that BA44 responded significantly more to these sorts of action stimuli than to the non-biological motion stimuli not only when they were presented as spoken sentences, but also when they were presented as video clips. Taken together, these findings support the hypothesis that BA44 is, as the investigators put it, “endowed with polymodal capabilities, allowing the processing of higher-level conceptual aspects of action understanding” (Baumgaertner et al. 2007, 881).

3.4. THE SEQUENTIAL AND HIERARCHICAL ORGANIZATION OF GOAL-DIRECTED MOVEMENTS

More recently, two “companion” neuropsychological and TMS studies have further illuminated the specific role that BA44 plays in the processing of action concepts. In particular, these studies show that BA44 is essential for representing the spatiotemporal structure of complex actions in which certain movements are performed in a certain sequence in order to achieve an overarching goal. The neuropsychological study was conducted by Fazio et al. (2010) and involved six brain-damaged patients whose lesions overlapped maximally in BA44 (Figure 3e). These patients were given an action comprehension task that had almost no linguistic requirements. On each trial, they first watched a video clip of either a goal-directed human action (e.g., a man reaching for and grasping a bottle) or a nonhuman physical event (e.g., a bicycle falling over). Then they were shown four randomly ordered photographs that were snapshots of different stages of the video clip that they had just seen. The task was to re-order the sequence of photographs so that they were lined up in a way that reflected the natural unfolding of the action or event. The striking discovery was that, compared to a group of healthy control subjects, the patients were significantly impaired in the human action condition, but not in the nonhuman event condition. Furthermore, in a parallel study that involved repetitive TMS, Clerget et al. (2009) demonstrated that temporarily disrupting the operation of BA44 when healthy subjects performed a slightly modified version of the very same task led to significantly slower response times in the human action condition than in the nonhuman event condition. In contrast, stimulation at a control site – specifically, leg-related somatosensory cortex – did not differentially affect response times in the two conditions. Overall, then, the studies by Fazio et al. (2010) and Clerget et al. (2009) provide strong convergent evidence that BA44 is necessary for appreciating specifically the sequential and hierarchical organization of goal-directed movements.
Returning to the language domain, there is mounting evidence that when it comes to sentence processing, BA44 is highly sensitive to the syntactic linearization of the nominal elements that designate the core participants in events – most notably, the agent and patient. Not surprisingly, most if not all of the available data indicate that this region is biased toward clauses with canonical mappings between syntax and semantics – that is, mappings in which the subject > object syntactic prominence hierarchy corresponds isomorphically to the agent > patient semantic prominence hierarchy (for detailed discussion see Bornkessel-Schlesewsky and Schlesewsky 2009a,b; and for a computer model see Dominey et al. 2006). For instance, numerous PET and fMRI studies have shown that when clauses with canonical mappings, like actives (e.g., *The boy chased the girl*) and subject-relatives (e.g., *The boy who chased the girl was fast*), are compared with clauses with non-canonical mappings, like passives (e.g., *The boy was chased by the girl*) and object-relatives (e.g., *The boy who the girl chased was fast*), the latter engage BA44 significantly more than the former (e.g., Stromswold et al. 1996; Peelle et al. 2004; Bornkessel et al. 2005; Grewe et al. 2005, 2007; Bahlmann et al. 2007; Kinno et al. 2008).

An especially compelling demonstration of this effect was provided by Grewe et al. (2007), who conducted an fMRI study that required the participants to make acceptability judgments for German sentences. One condition consisted of sentences with blatant grammatical violations; indeed, these sentences were included in the experiment so as to elicit definite decisions of “unacceptable.” Four other conditions consisted of matched sets of sentences that were identical except for orthogonal variations along two dimensions: the order of the subject and object; and the animacy of the object (Table 2). As expected, the participants rated the sentences with object > subject order to be worse than the sentences with subject > object order, but not as bad as the sentences with blatant grammatical violations. Regarding the neuroimaging results, two main results are relevant here. First, when the investigators contrasted the two conditions in which both arguments were animate (S_A > O_A and O_A > S_A) against the two conditions in which only the subject was animate (S_A > O_I and O_I > S_A), they found significant activation in the left posterior superior temporal sulcus (pSTS), which is part of Wernicke’s area. This outcome is consistent with several other studies which suggest that the left pSTS, and the surrounding

<table>
<thead>
<tr>
<th>Condition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_A &gt; O_I</td>
<td>Wahrscheinlich hat [der Mann]_NOM [den Garten]_ACC gepflegt. probably has the man the garden taken care of “The man probably took care of the garden.”</td>
</tr>
<tr>
<td>S_A &gt; O_A</td>
<td>Wahrscheinlich hat [der Mann]_NOM [den Direktor]_ACC gepflegt. probably has the man the director taken care of “The man probably took care of the director.”</td>
</tr>
<tr>
<td>O_I &gt; S_A</td>
<td>Wahrscheinlich hat [den Garten]_ACC [der Mann]_NOM gepflegt. probably has the garden the man taken care of “The man probably took care of the garden.”</td>
</tr>
<tr>
<td>O_A &gt; S_A</td>
<td>Wahrscheinlich hat [den Direktor]_ACC [der Mann]_NOM gepflegt. probably has the director the man taken care of “The man probably took care of the director.”</td>
</tr>
</tbody>
</table>
temporoparietal territory, may be essential not only for representing thematic roles like agent and patient (Thompson et al. 2007; den Ouden et al. 2009), but also for linking them with grammatical relations like subject and object—a process that, during sentence comprehension, is more difficult to accomplish when both entities are animate and hence capable, in principle, of volitional behavior (Wu et al. 2007; Thothathiri et al. forthcoming; see also Blakemore et al. 2001). Second—and, for present purposes, more importantly—when the investigators contrasted the two conditions in which the object preceded the subject (O_I > S_A and O_A > S_A) against the two conditions in which the subject preceded the object (S_A > O_A and S_A > O_I), they found significant activation in BA44 (Figure 3f). Although some researchers might wish to interpret this result as reflecting syntactic movement operations (e.g., Grodzinsky 2000; Grodzinsky and Santi 2008) or the recruitment of syntactic/phonological working memory (e.g., Caplan et al. 2000; Rogalsky and Hickok 2011), Grewe et al. (2007) interpret it instead as reflecting the processor’s response to the violation of a fundamental linearization principle which stipulates that subjects encoding agents typically precede objects encoding patients. And this principle, it should be noted, is essentially the same as the subject salience principle discussed above. (For other linearization principles and their relevance to Broca’s area, see Bornkessel-Schlesewsky and Schlesewsky 2009a,b, forthcoming; Bornkessel-Schlesewsky et al. 2009, forthcoming.) In short, these findings are consistent with the view that while the left temporoparietal cortex may subserve much of the thematic content of linguistically represented goal-directed actions, BA44 may capture the sequential and hierarchical organization of such actions.


Based on the foregoing considerations, it is now possible to elaborate the following proposal: The strong typological tendency for transitive clauses to have either SOV or SVO word order ultimately derives from the adaptive capacity of BA44 to represent the sequential and hierarchical organization of goal-directed actions. As indicated above, BA44 is critically involved not only in executing and observing such actions (Section 3.2.), but also in processing them at a more schematic conceptual level (Section 3.3.). Moreover, this region is especially sensitive to how the major segments of such actions unfold in time (Section 3.4.). Several scholars have suggested that, from an evolutionary perspective, once BA44 became adept at extracting the skeletal structure of goal-directed actions, it could then apply that ability to other cognitive domains (e.g., Fiebach and Schubotz 2006; van Schie et al. 2006; Tettamanti and Weniger 2006; Fadiga et al. 2009; Pulvermüller and Fadiga 2010). And in fact there is growing evidence that BA44 does play a pivotal role in identifying the hierarchical patterns that are latent in many different kinds of sequential events, including music (e.g., Maess et al. 2001), visuospatial stimuli (e.g., Bahlmann et al. 2009), artificial grammars (e.g., Friederici et al. 2006), and, as described in the immediately preceding section (i.e., Section 3.5.), natural grammars.

With specific regard to natural grammars, the original specialization of BA44 for capturing the linear, nested organization of goal-directed actions may have provided the neurocognitive platform that gave rise to the powerful cross-linguistic preference for SOV and SVO word orders (Figure 4). As indicated above, it seems plausible that this region represents the sequential and hierarchical structure of the prototypical transitive action scenario, in which an animate agent transmits energy to an inanimate patient and thereby changes its state. And if that assumption is correct, then the two semantically based explanatory principles that are arguably rooted in that scenario—subject salience and
verb–object contiguity – emerge rather straightforwardly along the lines described earlier. Thus, the linguistically manifested syntax of transitive clauses may be a reflection of the motorically manifested syntax of goal-directed actions, and both types of syntax may have overlapping neural substrates in BA44.

Although this proposal is grounded in a wealth of empirical studies, it is admittedly a bold and speculative attempt to bridge the gap between linguistic typology and cognitive neuroscience. Whether it is on the right track remains to be seen, but hopefully it will stimulate further research on the interaction between the cross-linguistic representation of action and the functional architecture of Broca’s area.

Short Biography

David Kemmerer’s research focuses on how different kinds of linguistic meaning are mediated by different neural systems, drawing on behavioral and lesion data from brain-injured patients as well as behavioral, functional neuroimaging, and electrophysiological data from healthy subjects. He has authored or co-authored papers for *Brain and Language, Cognitive Neuropsychology, Neuropsychologia, Cortex*, and the *Journal of Neurolinguistics*. He has also contributed chapters to several books, including *Action to language via the mirror neuron system* (edited by M. Arbib), *Words in the mind* (edited by B. Malt and P. Wolff), and *Language, cognition, and space* (edited by V. Evans and P. Chilton). He is currently writing a textbook called *The cognitive neuroscience of language*. Before coming to Purdue University, where he presently teaches, he was a postdoctoral fellow in the Department of Neurology at the University of Iowa. He holds a PhD in linguistics from SUNY Buffalo.

Note

* Correspondence address: David Kemmerer, Department of Speech, Language, and Hearing Sciences, Purdue University, Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-1353, USA. E-mail: kemmerer@purdue.edu

Works Cited


Bornkessel-Schlesewsky, Ina, Tanje Grewe, and Matthias Schlesewsky. forthcoming. Prominence vs. aboutness in sequencing: a functional distinction within the left inferior frontal gyrus. Brain and Language.


——, and ——. forthcoming. Linguistic sequence processing and the prefrontal cortex. The Open Medical Imaging Journal.


Frith, U., and C. Frith. 2010. The social brain: allowing humans to boldly go where no other species has been. Philosophical Transactions of the Royal Society B 365. 165–76.


——, and Javier Gonzalez Castillo. 2010. The two-level theory of verb meaning: an approach to integrating the semantics of action with the mirror neuron system. Brain and Language 112. 54–76.


——, David Rudrauf, Ken Manzel, and Daniel Tranel. forthcoming. Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. Cortex.


© 2012 The Author
Language and Linguistics Compass © 2012 Blackwell Publishing Ltd