Innate Mechanisms for Acquiring Syntactic Displacement

Misha K. Becker

1. Introduction

The central arguments within Lenneberg’s thesis of a biological basis for language are the species-specific nature of the physiological and neurological structures that make language possible, the cross-species uniformity of language development (the fact of its acquisition as well as its developmental path, irrespective of culture, race, etc.; excepting cases of pathology), and the transformational nature of syntax. Transformational syntax forms an important piece of support for Lenneberg’s discontinuity theory of the evolution of language, meaning that human language has not descended directly from communication systems found in non-human animals (i.e., our shared ancestors). This is because transformational syntax is also species-specific, i.e. not found in the communication systems of other animals.

Transformational syntax allows us to convey complex and abstract meanings, rather than being limited to the here-and-now (e.g. alarm calls) or to simple semantic relations, and it enables us to transform our expressions through syntactic displacement, or movement. In this short paper I will address some questions about how human children come to acquire the meanings of semantically abstract predicates, how they figure out which strings of words are generated by displacing operations, and the sense in which the tools that allow children to acquire both these things are innate. The inspiration for this research can be traced to some of the central themes in Lenneberg’s important work.

2. Displacement

The examples Lenneberg drew on to illustrate the transformational nature of human language syntax involved syntactic ambiguity: the surface form They are boring students is compatible with two underlying structures, neither of which is given ostensively to the child learner. Such types of sentences present an interesting puzzle and a demonstration of the necessity for learners to determine the underlying structure of a sentence which may not be straightforward to determine from its surface properties. Another example of this puzzle, and one that intersects with the puzzle of children’s word learning (in particular, predicate learning), concerns what Lees called “multiply ambiguous adjectival constructions,” such as John is tough to please (Lees 1960). It is not that the sentence itself is structurally ambiguous, but rather it has a “constructional homonym” (Chomsky 1964) in surface-similar sentences like John is eager to please. Thus, while not ambiguous in the adult grammar, the string of words cannot be mapped unambiguously onto its underlying structure without
prior knowledge of what the main predicate means. In the case of children, however, we can’t assume that the lexical meaning of such abstract predicates is known a priori.

Within movement-based approaches to syntax, such as Government–Binding Theory or Minimalism, phrases are generated in one position, as determined by their argument structure relations, and then they can move to a different position in the sentence (driven and constrained by structural requirements) where they are pronounced on the surface. When NPs are displaced like this there is a difference between where the phrase is spoken and where it is interpreted. In order to represent these sentences correctly, children have to figure out where the displaced NP is interpreted.

I will limit the discussion here mostly to raising (1) and tough-constructions (2).

(1) Donald, seems \([t, to \text{ be lying}])
(2) Melania, is tough \([\text{PRO}, to \text{ please}])

The well-known structure in (1) presents a case in which the matrix subject Donald has raised into the subject position of the main clause from the subject position of the embedded clause. Thus, although Donald is local to seem on the surface, its selecting predicate (the predicate it is semantically related to) is the non-local predicate be lying. Tough-constructions have evaded a clear syntactic analysis within the GB and Minimalist frameworks (Chomsky 1977, Lasnik & Uriagereka 1994; but see Hicks 2009 for one recent Minimalist account), but they likewise involve a structure in which the matrix subject is not interpreted in its surface position. Instead, it is interpreted as the object of the embedded verb (in this case, the theme of please). Thus, in both of these constructions the subject of the main clause is not an argument of the (surface-local) matrix predicate (see or be tough).

The question for acquisition, then, is how learners figure out the underlying structure and interpretation of these sentences. A possible answer comes from the Semantic and Syntactic Bootstrapping hypotheses, two accounts of how children use inductive biases about language structure in order to “bootstrap” themselves into the grammar. In a nutshell, Semantic Bootstrapping is the idea that since there is a regular and (mostly) reliable mapping between semantics and the syntactic category or function of constituents (nouns often label objects (DP in VP), verbs (VP/VP) often label events and states, subjects (DP in vP) are often agents), children could use the expectation of this mapping to figure out some basic aspects of syntax, such as word order (Grimshaw 1981, Pinker 1982). If you hear the sentence The dog is chasing the cat while viewing a scene in which a dog is chasing a cat, and if you know which objects the nouns dog and cat map onto and that the agent (the chaser) should be the syntactic subject, and if you can infer that chase means something like “pursue,” then you could determine that your language has Subject-Verb-Object (SVO) word order. In this way, the child uses expectations about semantic patterns to bootstrap into syntax. The expectations themselves are taken to be innate (not learned from experience) and specific to language (that is, domain specific).

Evidence for the bias to map animate or agentive NPs to the most prominent syntactic position comes not only from psycholinguistic studies with children and adults (Clark & Begun 1971, Trueswell & Tanenhaus 1994, Mak et al. 2002, Traxler
et al. 2005, Becker 2005, Becker & Estigarribia 2013), but also from argument hierarchies in language that are based on broad typological patterns. Well-known hierarchies such as the Animacy Hierarchy (Silverstein 1976, Dixon 1979) and the Thematic Hierarchy (Jackendoff 1972), as well as the Promotion to Subject Hierarchy (Keenan 1976) show how widespread the preference is in human language to associate animacy and agency with syntactic prominence, and inanimacy with lower positions. While this does not in itself necessitate an appeal to innateness, many linguists believe that typological universals are good candidates for innate components of language. Furthermore, there is recent neurolinguistic evidence for anatomically distinct cortical areas for representing agents and patients in sentences (Frankland & Greene 2015). Once again, by itself this does not necessarily point to innate knowledge of agents vs. patients, but combined with typological universals and psycholinguistic evidence, it is strongly suggestive of a biological basis or predisposition for this knowledge—as Lenneberg already recognized (see Piattelli-Palmarini, this issue for discussion).

There is wide support for the biases underlying Semantic Bootstrapping aiding early parsing of what Keenan (1976) called “basic sentences.” One problem in applying this type of approach to sentences like (2) and (1), however, is that, as mentioned above, these sentences have “constructional homonyms,” surface-similar sentences that are generated by very different underlying structures. They are exemplified in (3) and (4), respectively.

(3) Donald claims [PRO, to be tremendous]
(4) Melania is eager [PRO, to please e]

Thus, these constructions provide another example of the challenge Lenneberg alluded to with structurally ambiguous sentences. Prior to knowing what claim or seem means, the sentence is structurally ambiguous. A second shortcoming of Semantic Bootstrapping is that this procedure depends upon learners’ ability to infer the meanings of individual words from observing what is going on in the world when words are uttered. Rarely are individual words used to label events or properties in isolation, outside of the context of a sentence. Thus, while it may be reasonable to assume that children could figure out the meanings of some concrete nouns (dog, cat) from observing the world, most words, verbs in particular, are not learned that way (parents might say “Look at the dog!” but probably not “Look at chasing!”). If this criticism holds with verbs like chase, consider how much more problematic it would be to have to figure out the meaning of seem based on observation of the world.

Gleitman (1990) proposed a solution that exploited the predictable relationships between the argument structure frames a verb may participate in (that is, the number and category of “arguments,” or phrases a verb selects as the participants in the verb’s event) and the semantics of that verb. Her idea was that if children anticipated such relationships, they could use argument structure to bootstrap into verb meanings. Argument structure frames do not tell you exactly what an individual verb means (certainly, transitive sentence frames admit thousands of verb meanings), but they can help discriminate verbs that denote an individual action or change of state (1 argument; 5a) from verbs of contact or causation (2 arguments; 5b) from verbs of transfer (3 arguments; 5c; Landau & Gleitman 1985, Gleitman
1990). Crucially, this procedure works particularly well for certain verbs that denote unobservable states such as mental states and beliefs, like (5d; Gleitman et al. 2005).

(5) a. Jeff slept.  
   (1 NP argument: intransitive, individual action/change of state)  
b. Jim hugged Melissa.  
   (2 NP arguments: transitive; contact, causation)  
c. Tommie offered [his constituents] [options]  
   (3 arguments: ditransitive; transfer)  
d. Theresa regrets [that she called for an election]  
   (NP subject and clausal complement; mental verb)

According to this procedure, known as Syntactic Bootstrapping, learners use expectations about syntax to map whole sentences onto events in the world and thereby bootstrap into verb meanings. While Semantic and Syntactic Bootstrapping aim to explain different aspects of how children learn syntax and argument structure, they both rely on the fact that argument structure relations must be discernable from the surface forms of sentences; that is, they are surface-local. More must be said, however, when thematic relations among words in a sentence span long distances. What happens when the noun next to a verb is not the verb’s argument? Precisely this situation arises in raising and tough-constructions.

In brief, both Semantic and Syntactic Bootstrapping employ inductive bias to help children determine the syntactic and semantic properties of “basic sentences.” But both approaches fall short when we consider displacing constructions, which illustrate one of the core properties of human language and present a similar puzzle to what Lenneberg cited as evidence for the transformational nature of syntax. The solution to how children figure out these constructions will rest on a number of innate biases.

3. Toward a Solution

3.1. Inanimate Subjects as Cues to Structure

We just saw that both Semantic and Syntactic Bootstrapping rely on inductive biases to associate certain semantic roles with certain syntactic roles or grammatical categories, and to associate certain argument structure patterns with certain subcategories of predicate meanings. Both the semantic roles and the argument structure relations are projected locally and, in the canonical case, computed locally. These biases are taken to be innate in the sense that they are specific to language (which in turn is species-specific), and apply cross-linguistically (and thus species-wide; see Lee & Naigles 2008 for cross-linguistic evidence).

With these biases in place, the learner could identify sentences that break with the canonical pattern if there is some feature of these non-canonical sentences that could serve as a trigger. In my previous work I have argued that hearing an inanimate subject in sentences like (1) and (2) can provide such a trigger by signaling that the subject and main predicate are not thematically related (i.e. not in an argument structure relationship), and therefore that the subject is displaced from elsewhere in the sentence (Becker 2005, 2006, 2009, 2014, 2015, Mitchener & Becker 2011, Becker
Inanimate subjects provide a helpful clue because inanimate NPs cannot be agents or experiencers and therefore can’t be the subject argument of a control predicate (the ones found in (3) and (4)). All languages which contain both raising and tough constructions, to my knowledge, allow inanimate NPs as subjects of raising and tough predicates and disallow inanimate NPs as subjects of control predicates (various Indo-European languages, plus Tongan, Samoan, Niuean, Chamorro and Maori for raising, and Finnish, Mandarin, Labrador Inuttut, Niuean and Bahasa Indonesian for tough constructions). Moreover, inanimate subjects are more generally restricted: there are a number of languages that simply ban inanimate subjects in monoclausal transitive contexts (Japanese, Jacaltec, Navajo, Tlapanec, Blackfoot), and in languages that allow inanimate subjects in such contexts there are restrictions on their distribution that do not apply to animate subjects (Chung 1983, Comrie 1989, Dahl & Fraurud 1996).

In short, inanimate subjects are well tolerated as NPs that have undergone the kind of displacement in raising and tough-movement, but not as well tolerated as NPs that are underlyingly generated as external arguments—that is, as underlying subjects.

In terms of the learning procedure, if children are biased to expect sentence subjects to be agents of the main predicate, and if they further assume that inanimate things cannot be agents (supported by both research on conceptual development, e.g., Woodward et al. 1993, Poulin-Dubois et al. 1996, and linguistic development, e.g., Corrigan 1988, Scott & Fisher 2009), then an inanimate NP in subject position should indicate that there is something out of the ordinary going on with the sentence. It is either a syntactic object, implying an alternative word order in just these sentences, or it is a displaced subject. My contention is that learners take it to be a displaced subject. In my experimental work with adults and children, I have found that by manipulating the animacy of subjects of sentences with infinitive complements we can see asymmetries in how people parse these sentences: an animate subject tends to lead people to interpret the subject as an argument of the main predicate, as it would be in a control structure, and an inanimate subject tends to lead people to interpret the subject as being an argument only of the lower predicate, as it would be in a raising or tough construction.

It is important to note that tough- and raising constructions have been argued to be acquired relatively late in development. Chomsky (1969), Cromer (1970) and, more recently, Anderson (2005) have shown that children go through a long stage of misinterpreting the subject of a tough-construction as coreferent with the subject of the embedded clause rather than the object (i.e., they take the matrix subject to refer to the “pleaser” in (4)—note that this doesn’t mean children cannot interpret the subject as displaced, rather they haven’t figured out exactly where it is displaced from); Hirsch & Wexler (2007), Orfitelli (2012) and others have argued that certain types of raising constructions are acquired late in development, as late as age 7.

1 There is an interesting question about what this would mean for languages that are considered topic-prominent, meaning that they contain topics instead of subjects. As discussed in Becker (2014:p. 290ff), although topics, unlike subjects, are not selected by predicates, and therefore are not as much semantically limited as true subjects are, there are nevertheless strong tendencies both for topics to be associated with agenthood (Li & Thompson 1976, Givón 1976, Schachter 1976) and for topic-prominent languages to also admit subjects when no topic is available, e.g. in Lisu (Li & Thompson 1976).
Although my own empirical work has shown that children may comprehend these structures as early as age 4 it is clear that their knowledge of them is not fully adult-like until later. The relatively late acquisition of these constructions is not in itself problematic for a biological account of language acquisition. Lenneberg notes that the gradual development of language along a specific trajectory is in fact to be expected on the view that it is biologically based; biological structures typically mature along a developmental timeline.

3.2. Other Displacing Constructions

I’ve focused here on inanimate subjects as a cue to raising and tough-constructions, but there are other constructions involving displacement, as well as other cues to displacement. I’ll touch on some of these constructions briefly.

There are several other types of A(rgument)-movement, the type of displacement I’ve discussed above, such as passivization, Raising-to-Object, and subject raising with unaccusative verbs (intransitive verbs whose sole argument is internal, rather than external, e.g. *fall, arrive*). There is some evidence that inanimate subjects can serve as a cue to distinguishing these subcategories of verbs as well. In adult grammar unaccusative verbs (fall, arrive) freely permit inanimate subjects (see 6), while unergative verbs (laugh, jump) typically require animate subjects (see 7).

(6) a. The tree, fell $t_i$.
    b. The package, arrived $t_i$.
(7) a. #The tree laughed.
    b. #The package jumped.

There is room for debate about how strictly unergative verbs require an animate subject; Folli & Harley (2008) in fact argue that the relevant feature here is not animacy or agency, but “teleological capability,” and that this accounts for the fact that (8a) is well-formed (since trains are capable of emitting a whistling sound) but (8b) is not, unless you add the directional PP into the room, since bullets are not capable of sound emission except through their movement.

(8) a. The train whistled.
    b. The bullet whistled *(into the room).

There are some other cases in which unergative verbs allow inanimate subjects (e.g. The machine ran for hours). Nevertheless, when we look at the distribution of unaccusative and unergative verbs with respect to subject animacy, we see a striking asymmetry: while children produce unergative verbs almost exclusively

---

2 Space limitations prevent taking up the issue of expletive subjects, which serve as another cue to learners that a given predicate is a raising or tough type of predicate (e.g. It’s tough to please John vs. *It’s eager to please John*). Expletives undoubtedly serve as an important cue to a predicate’s selectional properties in languages that have them.

3 I am putting aside constructions involving what is known as A-bar movement (e.g. wh-movement, scrambling, topicalization), which does not move a phrase into a canonical argument position.
with animate subjects (over 93 %), they produce unaccusative verbs with both animate and inanimate subjects in roughly equal proportions (Becker & Schaeffer 2013, Becker 2014). A similar asymmetry is found in parental speech to children (Becker 2014). In addition, experimental work has shown that children interpret an inanimate subject of an intransitive verb as a patient (and the verb as an unaccusative verb), but they interpret an animate subject as an agent (and the verb as unergative; Bunger & Lidz 2008, Scott & Fisher 2009).

The passive construction, like unaccusatives, involves displacement of a verb’s internal argument to the subject position on the surface. Like unaccusatives and the other displacing constructions discussed, passives are quite compatible with inanimate subjects. Languages that disallow inanimate subjects of active transitive verbs (noted above, e.g., Japanese, Navajo) freely allow inanimate subjects under passivization. Thus, inanimate subjects could in theory be used as a cue for acquiring, or at least identifying passive constructions. Unfortunately, it is not so clear that inanimate subjects are used by children as cues to passives. To give just one example, Lempert (1989) trained children who were not yet producing passives spontaneously to imitate passive sentences that had either animate or inanimate patients. She then tested these children a few days later to see how many passives they would produce on their own in a picture description task. Lempert found that children who had been trained on sentences with animate patients produced the most passives, while children trained with inanimate patients produced more active sentences.

An important consideration about passives is that the lexical meaning of a verb does not change according to whether it is used with passive or active voice (X ate the sandwich and The sandwich was eaten both denote a sandwich-eating event), but the subcategories of predicates that participate in either raising vs. control, tough vs. control or unaccusative vs. unergative structures have very different kinds of lexical meanings. Table 1 shows a partial list of the meanings of raising and control predicates found in various languages that contain these constructions.

Perlmutter (1978) defined unergative verbs as “predicates describing willed or volitional acts” and unaccusative verbs as denoting “non-voluntary emission of stimuli that impinge on the senses” as well as verbs denoting aspectual and in-

<table>
<thead>
<tr>
<th>Language</th>
<th>Raising meanings</th>
<th>Control meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chamorro</td>
<td>begin, stop</td>
<td>be afraid</td>
</tr>
<tr>
<td>English</td>
<td>seem, appear, turn-out, tend, used-to, gonna</td>
<td>want, try, decide, claim</td>
</tr>
<tr>
<td>German</td>
<td>seem, used-to, must</td>
<td>try, forget, forbid</td>
</tr>
<tr>
<td>Italian</td>
<td>seem, turn-out</td>
<td>try, claim, pretend</td>
</tr>
<tr>
<td>Maori</td>
<td>not, don’t</td>
<td>want, decide, go, be able, agree, prepare</td>
</tr>
<tr>
<td>Niuean</td>
<td>be possible, begin, not, usual, almost, nearly</td>
<td>try, want, choose</td>
</tr>
<tr>
<td>Samoan</td>
<td>be able, be necessary, begin, be (done) thus</td>
<td>want, try, encourage, go, think, be tired of, (dis)like</td>
</tr>
<tr>
<td>Tongan</td>
<td>be able, be possible</td>
<td>go, come, stand up, turn, hit</td>
</tr>
</tbody>
</table>

Table 1: A partial list of raising and control predicate meanings in different languages (in alphabetical order top to bottom).
choative properties of events. In languages that have *tough*-adjectives (e.g. English, Finnish, Mandarin, Labrador Inuttut, Niuean and Bahasa Indonesian) the lexical meanings of these predicates display a remarkably tight range of meanings exclusively revolving around ease and difficulty. If an inanimate subject serves as a cue not only for identifying a displaced subject but also for narrowing down the lexical meaning of the predicate in these constructions, then the claim about inanimate subjects should not in fact extend to the case of passives. Space limitations prevent a more thorough exploration of this issue.

3.3. **Animacy**

The ability to use cues from animacy in the service of decoding syntax, whether for canonical or non-canonical structures, depends upon young children’s ability to distinguish conceptually between animate and inanimate entities, and to know that animate entities can be agents while inanimate entities cannot. There is a wealth of research on cognitive development showing children’s very early ability to make these distinctions (Spelke 1991, Woodward et al. 1993, Poulin-Dubois et al. 1996, Woodward 1998). Much of this research reports infants’ and toddlers’ ability to distinguish between inanimate objects and humans, but there is strong evidence that preschoolers reason quite differently about the internal properties and potentialities of inanimate objects and non-human animals as well (Carey 1985, Massey & Gelman 1988, Keil 1989, Rakison & Poulin-Dubois 2001), and that they have a sophisticated understanding of the propensity for animals, but not inanimate objects, to have agentive properties (contrary to the common wisdom, due to Piaget, that children are “animistic” in their beliefs). Thus, there is good evidence that young children have reliable knowledge of the animate–inanimate distinction that is not limited to human vs. inanimate entities.

Could knowledge of the animate–inanimate distinction be innate? It is difficult to answer this question, but I’ll offer some ways to think about how to approach it. First, there is evidence that from the earliest moments of life, infants attend to human faces but not to rearranged components of a human face (Johnson et al. 1991). This suggests a hard-wired ability to distinguish humans via a salient and prototypically animate property. Cognitive systems for recognizing faces vs. objects are functionally and anatomically distinct (Farah 1995), as are the systems for recognizing animate vs. inanimate objects (Warrington & Shallice 1984), suggesting a differentiation of these categories on a neurological level. The way people reason about animals vs. objects, according to a folk biology or folk taxonomy, is remarkably uniform across the world (Atran 1998), speaking to the species-wide nature of this distinction. Taken together, these facts suggest that the animate–inanimate distinction is profound and fundamental to our conceptualization of the world, and they are consistent with a view that this distinction, or the basis for it, is innate.

Second, there is evidence that the ability to discriminate animates from inanimates and to attribute agency only to animates is found across species, not only among primates but also distantly related species such as dogs and birds (Hare et al. 2000, Flombaum & Santos 2005, Carey 2009), suggesting a shared trait with our common ancestors. Here I depart from Lenneberg’s focus on species-specific traits as evidence for biological sources of abilities or behaviors. However, using
cross-species evidence for the ability to distinguish animates from inanimates as a rationale for its biological roots is not at odds with Lenneberg’s overarching thesis; rather, if we view an innate ability to distinguish animates from inanimates as external to language, but used in the service of decoding grammar by individuals who have language (i.e. humans), this should not run counter to any of Lenneberg’s arguments. Indeed, he noted that “No biological phenomenon is without antecedents,” (Lenneberg 1967: 234). Surely, a great many physiological and cognitive traits are shared with our ancestors, encoded in our genetic makeup, and a shared ability to distinguish animates from inanimates could be one of these. It seems reasonable to acknowledge that this ability is likely outside of language and merely used within the process of acquiring language, as many other cognitive, motoric and perceptual abilities are. In the case of the constructions discussed here this conceptual distinction, in combination with inductive biases about grammatical structure and syntax–semantics mappings, provides language learners with a crucial means of achieving what no other communication system permits and is one of the hallmarks of human language: computation of non-local semantic dependencies.

4. Conclusion

In this brief article I extended some of Lenneberg’s claims about the biological basis for language by exploring the acquisition of some syntactic constructions that epitomize the transformational nature of syntax, namely raising and tough-constructions. The primary claim is that in order to acquire the structure of these types of sentences children must rely on biases specific to language as well as an innate (but not necessarily specific to language) ability to distinguish animate from inanimate entities and to associate animates with (some degree of) agency. The biases specific to language include the expectation that subjects will tend to be animate and/or agentive (as in Semantic Bootstrapping), that argument structure frames are projected locally prior to any syntactic displacements and that these frames restrict the lexical meanings of predicates (as in Syntactic Bootstrapping), and that deviations from the canonical alignment of semantic properties with syntactic roles and categories within a sentence provides a clue to learners that semantic relations may need to be computed non-locally.

References


Piattelli-Palmarini, Massimo. 2017 [= this issue]. From zero to fifty: Considerations on Eric Lenneberg’s Biological Foundations of Language and updates. Biolinguistics.


