

# Ergativity and Semantic Bootstrapping \*

Laura Siegel  
University of Pennsylvania  
*siegel@ling.upenn.edu*

## 1 Overview

The purpose of this paper is to raise some learnability issues that arise when languages with relatively rich morphology are investigated. Specifically, I investigate whether semantic bootstrapping (e.g. Pinker 1984) can account for the acquisition of case-marking inflection in ergative and split ergative languages.

This question has been debated in the literature with respect to ergative languages, with Pinker (1984, and others) arguing that it can and Pye (1990) arguing that it cannot. I argue that semantic bootstrapping can account for the acquisition of case inflection in ergative languages. However, I argue that semantic bootstrapping cannot be extended to the acquisition of case-marking in split ergative languages (languages that use a combination of ergative and accusative patterns), and therefore, that it cannot explain the acquisition of case-marking in general.

In this paper, I examine some linguistic patterns with respect to case-marking that arise in these languages. I will investigate the ramifications of these case-marking patterns for semantic bootstrapping as a way to account for their acquisition, and assess whether semantic bootstrapping is a desirable or even possible mechanism to explain their acquisition. I will conclude that a better account of the acquisition of case-marking is one that places a greater emphasis on syntactic and distributional information.

---

\*I have benefitted from discussions with Robin Clark, Lila Gleitman Toby Mintz, Laura Wagner, Ed Kako and Jesse Snedeker, as well as from comments from the audience of the LSA 2000 Meeting, Chicago. Thanks also to Rajesh Bhatt and Rashmi Prasad for help with the Hindi data, and to Martha McGinnis for help with the Georgian data.

## 2 Background

Before applying semantic bootstrapping to ergativity, I will digress briefly to review semantic bootstrapping, and then to review ergativity.

### 2.1 Semantic Bootstrapping

Pinker (1984, 1989, 1994), following Grimshaw (1981) and Macnamara (1982) has suggested “semantic bootstrapping” to account for children’s beginning acquisition of syntax. Pinker postulates that children are born with linkings between some semantic categories (e.g. agent of action) and syntactic categories (e.g. subject of active sentence). Learners then use real world/contextual information to identify the semantic category and then “bootstrap” their way into the syntax.

Of course, this cannot be used for every aspect of acquisition of syntax, as there are cases where the linkings between grammatical roles and syntactic positions are not straightforward (for example passive voice), and cases where contextual information is insufficient to identify the roles of the participants (for example, in cases where it is necessary to identify the speaker’s perspective on an event to know his or her role e.g. the distinction between *chase* and *flee*, etc.), but it has been proposed as a mechanism that could be used in the early stages to aid acquisition of basic syntax.

### 2.2 Ergativity

Generally speaking, languages mark grammatical role with either word order or nominal inflectional morphology (case-marking). Cross-linguistically, there are two basic patterns in how languages divide up the three basic roles; subject of intransitive, subject of transitive, and object of transitive. ACCUSATIVE languages (e.g. English, Latin, Spanish) group together subjects as a class, and distinguish them from objects. ERGATIVE languages (e.g. Basque, Inuktitut (Eskimo), K’iche’ Maya) group together the subjects of intransitive verbs and the objects of transitive verbs, and distinguish these from the subjects of transitive verbs (see Dixon 1994).

The examples in (1) from Inuktitut (Eskimo) show typical case marking patterns in an ergative language.<sup>1</sup> The example in (1a) shows an intransitive verb and the example in (1b) shows a sentence with a transitive verb. These examples show that Absolutive case (ABS) is assigned to the subjects of intransitive verbs and the

---

<sup>1</sup>It has been argued that many languages which are generally analyzed as being ergative are actually split-ergative (for example, see Bok-Bennema (1991) for Inuktitut). Split-ergative languages sometimes show an ergative case-marking pattern and sometimes an accusative pattern. This will be discussed further in §4. In any case, nothing here hinges on this at all and the examples given clearly show an ergative pattern.

objects of transitive verbs, and Ergative case (ERG) is assigned to the subjects of transitive verbs.

(1) Inuktitut

- a. Jaani            tikit   -juq  
Johnny(**ABS**) arrive -MOOD&AGR  
'Johnny arrived.' (from Allen (1994: 15))
- b. arnaq   -up   qimiq   taku -vaa  
woman -**ERG** dog(**ABS**) see -MOOD&AGR  
'A/the woman saw the dog.' (from Kalmár (1979: 87))

### 2.3 Importance of Ergativity for Semantic Bootstrapping

Ergative languages present an apparent problem for semantic bootstrapping because the linking rules between grammatical role and structure are seemingly disrupted.

Specifically, the semantic category of agent is not treated uniformly by the morphosyntax. That is, there is no unique mapping between agents of actions and particular case-marking inflection since agents might be marked with one of two possible cases. Furthermore, in some situations, the marking on subjects is the same as case marking that appears on objects.<sup>2</sup>

#### 2.3.1 Previous Work

This question has been discussed in the literature to some extent. Pinker (1984) specifically claims that semantic bootstrapping could be used to account for the acquisition of case-marking in ergative languages. In fact, Pinker (1984) uses acquisition data from case-marking in ergative languages as evidence for semantic bootstrapping. He discusses data from Kaluli from Schiefflin (1985, as discussed in Slobin (1985)) that indicate that children pay attention to semantic categories when marking case, and initially place semantic restrictions on the use of case-marking.

---

<sup>2</sup>It is theoretically possible that even though case-marking allows for free word order in ergative languages, the input to learners uses a generally constant word order. If this were true, then learners could use word order to learn case-marking (as in languages like English). I was unable to obtain data bearing on this issue for many of the languages under consideration here. However, there is evidence that learners of free word order languages do, in fact, hear and understand a variety of word orders (e.g. Slobin and Bever (1982) on Turkish). This, combined with the variety of word orders found in adult language, indicates that learners could not just use word order to map grammatical role with case marking.

Pye (1990), on the other hand, argues that ergative languages are a problem for Pinker's (1984) semantic bootstrapping mechanism because this mechanism relies on the *agent* relation, which should lead the child to group together subjects of both transitive and intransitive verbs, failing to distinguish them for case-marking purposes.

However, I will argue that nonetheless, this is not actually a problem for semantic bootstrapping, but that semantic bootstrapping is not an adequate mechanism to explain the acquisition of case-marking in all languages.

### **3 Semantic Bootstrapping and Ergativity**

I argue, contra Pye (1990), that the fact that ergative languages do not mark all agents identically is not a problem for semantic bootstrapping. The reason is that this difference in the linking rules between subjects and case-marking is tied to the verb's transitivity. Crucially, for semantic bootstrapping, there are linguistic cues (other than case) to the transitivity of the verb in ergative languages. These cues include the nature of the verbal agreement (e.g. in Inuktitut (Fortescue 1984)), the presence of a transitive marker (e.g. in Ki'che' Maya (Pye 1985)), as well as there often being a differing number of arguments. This means that there will always be unique syntax into which children can bootstrap.

### **4 A Problem for Semantic Bootstrapping**

However, I will show that there are case-marking situations in which the linking rules are disrupted without corresponding linguistic marking. This evidence comes from languages called split ergative languages.

So, I will digress briefly and describe split ergativity, and then investigate issues that arise with semantic bootstrapping as an account of the acquisition of case-marking in these languages.

#### **4.1 Split Ergativity**

In addition to accusative languages (which treat subjects in a structurally similar manner and distinguish them from objects), and ergative languages (which group together subjects of intransitives and objects of transitives, and distinguish them from subjects of transitives), there is cross-linguistically a third way of treating these basic grammatical relations.

SPLIT ERGATIVE languages (e.g. Georgian, Hindi) combine the two systems. These languages sometimes use the nominative-accusative pattern and sometimes

the ergative-absolutive system. This is a common pattern in languages of the world.

In split ergativity, the split is conditioned by some other property or properties of the sentence, such as tense/aspect, lexical semantics of the verb, and possibly main vs. subordinate clause, or some property of the noun (e.g. full noun vs. pronoun, animacy), etc. (Dixon 1994).

The examples in (2) and (3) show the case-marking patterns in a split ergative language, Georgian. Georgian case-marking generally splits according to the tense/aspect of the verb. The examples show two different tense/aspects; the present tense (in (2)) and the aorist (a past tense for actions that are neither complete nor incomplete) (in (3)). In these examples the (a) sentences show the same intransitive verb, and the (b) sentences show the same transitive verb.

These examples show that sentences with the verb in the present (PRES) tense (shown in (2)) have an accusative case-marking pattern (i.e. subjects are treated similarly) while sentences with the verb in the aorist (AOR) tense (shown in (3)) have an ergative case-marking pattern (i.e. subjects of intransitives and objects of transitives are treated similarly).<sup>3</sup>

(2) Georgian - Present

- a. Student -i midis  
student -**NOM** goes(PRES)  
'The student goes.'
- b. Student -i ceril -s cers  
student -**NOM** letter -**ACC** writes(PRES)  
'The student writes the letter.'

(3) Georgian - Aorist

- a. Student -i mivida  
student -**NOM** went(AOR)  
'The student went'
- b. Student -ma ceril -i dacera  
student -**ERG** letter -**NOM** wrote(AOR)  
'The student wrote the letter.'

---

<sup>3</sup>The actual case-marking situation in Georgian is considerably more complicated, with many more splits, but these examples serve to show how split ergativity works.

## 5 Semantic Bootstrapping and Split Ergativity

The acquisition of case-marking in split ergative languages clearly presents a greater challenge than that of ergative languages to semantic bootstrapping.

There is, first of all, the problem that particular roles are associated with several different case markers. It would be possible for a proponent of semantic bootstrapping to argue that although the linking rules are not consistent in split ergative languages, they are correlated with other linguistic elements, so are, in fact, retrievable by the learner, assuming that the learner is sensitive to these other factors (tense, animacy, etc.).

The more factors there are conditioning the split, the more complicated the situation becomes, because the child has to correlate several factors simultaneously. But, assuming that there is linguistic marking correlated with the split, the use of semantic bootstrapping to learn case inflection in split ergative languages becomes significantly more difficult, although perhaps not impossible.

### 5.1 The Problem

However, there is not always some other structural difference correlated with the split in linking rules. I'll present examples of this situation from two split ergative languages, Hindi and Georgian.

#### 5.1.1 Hindi

The first such situation I'll discuss comes from experiencer subjects in Hindi.

**Experiencer subjects** The examples in (4) and (5) show two sentences that are structurally the same, but with different case-marking on the subject. As background, in Hindi, the split ergativity is conditioned by aspect. These sentences are in Perfective (PFV) aspect, which normally triggers the ergative case-marking pattern. So, the typical pattern would be for transitive subjects to have ergative case.

The example in (4) has dative (DAT) case on the subject, while the example in (5) has ergative case on the subject. If they were in the Habitual aspect (which triggers the accusative pattern), the dative-marked subject would stay the same and the regular transitive subject would be nominative).

It is true, though, that this difference in case-marking is correlated with a semantic distinction. The subject in sentences of the type in (4) is an "experiencer" subject, the action cannot be volitional, while, sentences of the type shown in (5) are not subject to this restriction.

- (4) Hindi - Dative-marked (experiencer) subject

Ram -ko ek chiRiyaa dikh-ii

Ram -**DAT** a bird(NOM) see1-PFV/AGR

'Ram saw a bird.' (but not volitionally, cannot be combined with adverbs like "deliberately")

- (5) Hindi - Ergative subject (typical transitive)

Ram -ne ek chiRiyaa dekh-ii

Ram -**ERG** a bird(NOM) see2-PFV/AGR

'Ram saw a bird.' (can be volitional)

Recall that semantic bootstrapping requires that the learner use contextual and perceptual information to understand the semantics, and then use innate linkings between semantics and syntax to bootstrap into the syntax. So, semantic bootstrapping might have a way out here, but **only** if the volitionality was recoverable from the scene, and the bootstrapping mechanism included the ability to perceive and track this information.

**Implied objects?** The Hindi examples in (6) and (7) illustrate another complication.

These two sentences are both intransitives with exactly the same overt structure, but different case-marking on the subject. The sentence in (6) has nominative case on the subject, which is the expected case for an intransitive subject here, while the subject in (7) is marked with ergative case.

This is exactly the situation that is problematic for semantic bootstrapping, a disruption of the linking rules between agent and subject case, with no other correlated linguistic marking.

- (6) Hindi intransitive

Ram gayaa thaa

Ram(**NOM**) go(PFV&AGR) be(PST&AGR)

'Ram had gone.'

- (7) Hindi intransitive - Ergative subject

Ram -ne nahaayaa thaa

Ram -**ERG** bathe(PFV&AGR) be(PST&AGR)

'Ram had bathed.'

There is, however, something that does distinguish the two sentence types shown in (6) and (7). The sentence in (6) is a true intransitive, with only one

argument structurally, and only one participant in the action. The sentence in (7) has a verb that is an inherent reflexive. Even though there is structurally only one argument, the interpretation includes an object, or patient of the action. That is, it means “Ram had bathed himself.”

The fact that the presence of a non-overt but implied object affects the case-marking can be seen with a further example. The sentence in (8) shows another intransitive sentence with the subject in the ergative case, not the typical (for a verb in this aspect) nominative case. Although a surface intransitive, there is an implied object here, i.e. The sentence means “Ram had eaten something.”

(8) Hindi intransitive - Ergative subject

Ram -ne khaa liyaa  
Ram -**ERG** eat(PFV/AGR) take(PST/AGR)  
'Ram had eaten.'

So, the apparent generalization is that sentences that look like intransitives such as (7) and (8) have case-marking on the subject that we would expect on transitive subjects, but not intransitive subjects. These intransitives though, are transitive at some level.

If this is, in fact, the correct generalization, in order for examples of this type not to be a problem for semantic bootstrapping, the bootstrapping mechanism would also have to include the ability to extract the presence of an implied object from the scene.

**Summary of Hindi** The data from Hindi indicate that in order for semantic bootstrapping to account for the acquisition of case-marking, the bootstrapping mechanism needs to be sensitive to semantic notions that are more subtle than *agent* of action.

It also becomes crucial that notions such as *non-volitional* vs. *possibly volitional* and sensitivity to a direct object that is implied but not expressed are apparent to the learner from the input context.

At best, this indicates that the bootstrapping process is significantly more difficult and complicated than originally proposed.

### 5.1.2 Georgian

Georgian provides examples of a situation that seems even more problematic for semantic bootstrapping, because there are examples of the linking rules between grammatical role and case-marking being disrupted with no corresponding difference in linguistic marking, nor any kind of semantic difference.



There is a class of intransitive verbs, called, in the descriptive literature, MEDIAL verbs which occur with the subject in ergative case when there is a third person subject (Merlan 1985). This verb class includes the verbs meaning ‘cry’, ‘sing’, ‘dance’, ‘work’, ‘play’ and ‘speak’ (Nash 1995). An example of a sentence with a verb of this type is shown in (10).

Comparing the intransitive aorist sentence with a “medial” verb in (10) with the intransitive aorist sentence ((3a), repeated here as (9)), it is clear that there are cases in Georgian where intransitive verbs have exactly the same structure and morphological marking, but different case-marking on the subject.

(9) Georgian Aorist Intransitive

Student -i mivida  
 student -**NOM** went(AOR)  
 ‘The student went.’ (Aronson 1990)

(10) Georgian Aorist Intransitive – ‘medial’ verb class

gogo -m itira  
 girl -**ERG** cried(AOR)  
 ‘The girl cried.’ (Merlan 1985: 342)

**Summary of Georgian** This situation is problematic for a semantic bootstrapping explanation for the acquisition of case-marking, because there is simply no way for a learner to bootstrap from the knowledge that the subject in each sentence above is the agent to the correct case.

## 5.2 Suggestive evidence

It is difficult to find data bearing on this question in the literature, and there are clearly many open empirical questions. However, it has been reported that children do in fact learn Georgian case-marking early and perfectly, including in precisely the situation that is problematic for semantic bootstrapping, medial verbs (Imedadze & Tuite 1985), one example is given below in (11).

(11) Correct use of ERG case with intransitive medial verb in 2;10 year old child

am bavš -am galaxa da am -am itila  
 this child -**ERG** beat(AOR) and this -**ERG** cry(AOR)  
 ‘This child beat [him] up, and this one cried.’ (Imedadze & Tuite 1985: 88)

## 6 Syntactic Bootstrapping

Another learnability issue raised by case-marking in (split-)ergative languages concerns SYNTACTIC BOOTSTRAPPING. This is the proposal that children can use the syntax of their language to learn meanings of words (Gleitman 1990, and others). Specifically, it is argued that learners use syntactic structure to guide and constrain their hypotheses about the meanings of words, particularly verbs.

The motivation and evidence for syntactic bootstrapping facilitating word learning comes partially from cases where real world context is not available to facilitate bootstrapping into syntax (e.g. blind children, mental state verbs) or is insufficient (e.g. verbs which require knowledge of the perspective of speaker, e.g. *chase/flee, give/get*, etc.).

The structural information most relevant for verbs is argument structure. This requires knowing about the syntactic “frame” of the verb; how many arguments it has, and to some agree what kind of arguments, nature of their participation, etc.

So, although syntactic bootstrapping is proposed for different aspects of acquisition that those for which semantic bootstrapping has been proposed, syntactic bootstrapping also involves mappings between structure and meaning. Therefore, one might expect that the examples from split ergative languages of the linking rules between grammatical role and case-marking being disrupted with no overt linguistic marking to signal this disruption would also be a problem for syntactic bootstrapping.

The question would be whether or not learners could use the information about grammatical role gleaned from case-marking on a verb’s arguments to bootstrap their way into the meaning of the verb. Of course, this assumes that children have acquired case-marking, a non-trivial assumption that should ideally be further explained.

However, assuming that learners have acquired the case-marking in their language, the question for syntactic bootstrapping is whether examples of the type shown from Georgian in (9) and (10) allow for a syntactic bootstrapping account.

Even though the linking rules are disrupted in examples of this type, they do not actually pose a problem for syntactic bootstrapping the way that they do for semantic bootstrapping. This is because the sentences have intransitive verbs. The fact that they involve intransitive verbs means that there is only one argument,<sup>4</sup> so it is clear that the child can understand the frame of the verb, regardless of any apparent mismatch in the case-marking.

If there were a case, similar to the Georgian case, that is, a case of the same

---

<sup>4</sup>The intransitive verbs involved are not, as far as I can determine, verbs like *eat* or *break* which, even when intransitive, can be construed as involving more than one actor/participant.

role being marked differently in the absence of any other linguistic cue, with a verb with more than one argument (i.e. where the frame of the verb is not apparent), then this would be a problem for syntactic bootstrapping, as well as for semantic bootstrapping. I have not been able to find such a case, and it is an interesting question whether such a case exists or not.

So, it appears that syntactic bootstrapping could potentially be used in these cases, although there are details which would need to be fleshed out.

## 7 Conclusion

In this paper, I have shown that case-marking in ergative and split ergative languages raises interesting learnability and acquisition questions. I have argued that semantic bootstrapping might help to account for the acquisition of case-marking in ergative languages. However, I have also shown that semantic bootstrapping runs into difficulties in accounting for the acquisition of case-marking in split ergative languages. These difficulties are of two basic types.

One problem is that the nature of the semantic information that the bootstrapping mechanism must be sensitive to is more refined than notions such as *agent* (e.g. volitionality). Moreover, these subtler semantic distinctions must be recoverable from the scene. It is not clear that this is, in fact, possible (cf. Gleitman (1990); Gillette, Gleitman, Gleitman & Lederer (1999)).

The second issue is that it is striking that the linking rules must take into account a wide range of, and considerably more specific, syntactic information than notions like *subject* or even *subject of transitive* vs. *subject of intransitive*. Since the goal of semantic bootstrapping is to use semantics to bootstrap into syntax, this seems like a questionable addition to the theory.

In summary, I have shown that semantic bootstrapping would need to be significantly modified to account for split ergative languages, and even with these modifications, would not be able to account for all of the data from these languages. Moreover, the nature of the modifications that would be necessary argues for an account of the acquisition of case-marking that gives a greater role to syntactic and distributional information.

## References

- Allen, S. E. M. (1994) *Acquisition of Some Mechanisms of Alternation in Arctic Quebec Inuktitut*, Doctoral dissertation, McGill University.
- Allen, S. E. M. (1996) *Aspects of Argument Structure Acquisition in Inuktitut*, John Benjamins, Amsterdam.
- Aronson, H. I. (1990) *Georgian: A Reading Grammar*, Slavica, Columbus. Corrected Edition.
- Bloom, P., ed. (1994) *Language Acquisition*, The MIT Press, Cambridge, MA.
- Bok-Bennema, R. (1991) *Case and Agreement in Inuit*, Studies in Generative Grammar, 38, Foris, Dordrecht.
- Dixon, R. M. W. (1994) *Ergativity*, Cambridge University Press, Cambridge.
- Fortescue, M. (1984) *West Greenlandic*, Croon Helm, London.
- Gillette, J., L. R. Gleitman, H. Gleitman, and A. Lederer (1999) "Human simulations of vocabulary learning," *Cognition* 73:2, 135–176.
- Gleitman, L. (1990) "The structural sources of verb meanings," *Language Acquisition* 1:1, 3–55. Reprinted in Bloom (1994).
- Gleitman, L., and H. Gleitman (1997) "What is a language made out of?" *Lingua* 100, 29–55.
- Gleitman, L., and E. Wanner (1992) "Language acquisition: the state of the state of the art," in L. Gleitman and E. Wanner, eds., *Language Acquisition: The State of the Art*, Cambridge University Press, New York, 3–48.
- Imedadze, N., and K. Tuite (1985) "The Acquisition of Georgian," in D. I. Slobin, ed., *The Crosslinguistic Study of Language Acquisition*, volume 3, Lawrence Erlbaum Associates, Hillsdale, NJ, 39–109.
- Merlan, F. (1985) "Split Intransitivity: Functional Oppositions in Intransitive Inflection," in J. Nichols and A. C. Woodbury, eds., *Grammar Inside and Outside the Clause: Some Approaches to Theory from the Field*, Cambridge University Press, Cambridge, 324–362.
- Mithun, M. (1989) "The Acquisition of Polysynthesis," *Journal of Child Language* 16, 285–312.
- Nash, L. (1995) *Portée argumentale et marquage casuel dans les langues SOV et dans les langues ergatives: l'exemple du géorgien*, Doctoral dissertation, Université de Paris VIII.

- Pinker, S. (1984) *Language Learnability and Language Development*, Harvard University Press, Cambridge, MA.
- Pinker, S. (1989) *Learnability and Cognition*, The MIT Press, Cambridge, MA.
- Pinker, S. (1994) "How Could a Child Use Verb Syntax to Learn Verb Semantics?" *Lingua* 92, 377–410. Reprinted in Gleitman, L. and B. Landau (1994).
- Pye, C. (1985) "The Acquisition of K'iche' Maya," in D. I. Slobin, ed., *The Crosslinguistic Study of Language Acquisition*, volume 3, Lawrence Erlbaum Associates, Hillsdale, NJ, 221–308.
- Pye, C. (1990) "The acquisition of ergative languages," *Linguistics* 28, 1291–1330.
- Saville-Troike, M. (1996) "Development of the Inflected Verb in Navajo Child Language," in E. Jelinek, S. Midgette, K. Rice, and L. Saxon, eds., *Athabaskan Language Studies: Essays in Honor of Robert W. Young*, University of New Mexico Press, Albuquerque, 137–185.
- Slobin, D. I. (1985) "Crosslinguistic Evidence for the Language-Making Capacity," in D. I. Slobin, ed., *The Crosslinguistic Study of Language Acquisition*, volume 2, Lawrence Erlbaum Associates, Hillsdale, NJ, 1157–1256.
- Slobin, D. I., and T. G. Bever (1982) "Children use canonical sentence schemas: A crosslinguistic study of word order and inflections," *Cognition* 12, 229–265.