

Mark C. Baker
Rutgers University

1. Introduction

Some prominent recent literature has challenged the notion of language universals and a universal grammar (Bickel 2007; Evans and Levinson 2009; Dunn, Greenhill et al. 2011). This literature claims that, as we collect more data from a wider range of languages and use it to construct large typological databases, we find that there are counterexamples to essentially every substantive and interesting universal property of languages that has been proposed. Indeed, we find (they say) that language seems to vary without clearly defined limits, subject primarily to cultural and historical factors. This has been taken to undermine the foundations of the formal-generative-syntactic approach to linguistics in particular, because that has long been associated with the idea that natural human languages are constrained by a rich, powerful, and innate universal grammar (UG), which is essentially constant across the human species (Chomsky 1965; Chomsky 1981). Also implicit or explicit in this work is the claim that further progress will be made primarily by increasing the size of our typological databases, by applying more sophisticated statistics to that data, and by emphasizing cultural and historical modes of explanation.

In the light of these claims, it is reasonable to ask questions like the following. First, are there bona fide linguistic universals that stand up to scrutiny over a wide range of languages? Second, can formal generative grammar play any distinctive role in discovering such universals? Third, for any such universals, can generative grammar play a distinctive role in explaining them? These are questions of great interest and significance for cognitive science anyway, and would be worth considering even if they had not been posed in critical fashion in articles like the ones mentioned.

In this work, I explore these questions within the specific domains of agreement and case-marking, claiming that the answer to all three questions seems to be yes, despite some bold claims to the contrary. I show that there are ideas and results in this domain that researchers can make use of in their own investigations with reasonable confidence.

2. Characterizing the domain of agreement and case

* For comments and discussion that led to improvements in this paper, I thank Jonathan Bobaljik, Steve Wechsler, [and ...]. For passages that cite previous works of mine, the acknowledgements in those works also apply to this one. Remaining errors are my responsibility.

Abbreviations used in this article include: ABS, absolutive; ACC, accusative; ACT, active; AOR, aorist; CAUS, causative; CIS, cislocative; DAT, dative; DJ, disjunctive; ERG, ergative; F, feminine; FAC, factitive; FACT, factual; FUT, future; FV, final vowel (Bantu); GEN, genitive; GER, gerund; HAB, habitual; IMPF, imperfective; IND, indicative; INDEF, indefinite; INF, infinitive; M, masculine; NE, prenominal particle (Mohawk); NEG, negative; NOM, nominative; NOML, nominalizer; OM, object marker; PASS, passive; PAST, past; PERF, perfective; PL, plural; POSS, possessor; PP, past perfective; PRES, present; PTPL, participle; PUNC, punctual; SG, singular; SM, subject marker; STAT, stative; T, tense marker; TH, translocative. The glosses of an agreement marker may include a number (1, 2, 3, ...) expressing the person or (in a Bantu language) the gender/noun class, a letter (s or p) expressing the number, a letter (M, F, N) expressing the gender, and a letter (S, O) expressing the grammatical function of the agreed-with nominal.

My primary emphasis is on the phenomenon of agreement. This can be characterized roughly as morphological marking on one word in a clause or other syntactic unit that reflects the features of another expression within that unit. A prototypical example is the finite verb agreeing with a noun phrase functioning as its subject. This happens in a limited way in English ((1)), and much more robustly in Bantu languages like Kinande ((2)):

- (1) a. The woman buys fruit each day in the market.
 b. The women buy fruit each day in the market
- (2) a. Abakali ba-a-gul-a eritunda. (Kinande (Baker 2008))
 woman.2 2S-T-buy-FV fruit.5
 ‘The women bought a fruit.’
- b. Omukali a-gul-a eritunda.
 woman.1 1S.T-buy-FV fruit.5
 ‘The woman bought a fruit.’

A secondary interest of this article is the phenomenon of case. This can be characterized roughly as morphological marking on a noun phrase (NP) or similar expression that reflects its grammatical relationship to the central verb of the clause, or to another key word in the syntactic unit that the NP is found in (e.g., to the preposition that heads a prepositional phrase). For example, pronouns in English take a different form depending on whether they function as the subject of a finite verb or as its object, as in (3). Similar differences are seen more robustly on NPs of all sorts in languages like Japanese ((4)).

- (3) a. I usually find **him** in the park.
 b. **He** usually finds me in the park.
- (4) a. John-no imooto-ga sinda. (Kuno 1973:6)
 John-GEN sister-NOM died
 ‘John’s sister died.’
- b. John-ga Mary-ni hon-o yatta (Kuno 1973:5)
 John-NOM Mary-DAT book-ACC gave
 ‘John gave Mary a book.’

If case and agreement are defined relatively narrowly, they are opposites: agreement is morphology on (say) the verb that is determined by features of a nearby NP, and case is morphology on an NP that is determined by properties of a nearby verb. Speaking more broadly, they are similar topics in that both are instances of the morphology of one linguistic expression being determined by its relationship to another expression within the same syntactic unit. This is one reason why the two topics are often treated together. A more specific reason is that many languages seem to follow rules that relate the two directly, such as “A verb agrees with the NP has nominative case.” The status of such connections is one of the issues that I consider below.

Although it also poses some special challenges,¹ this topic presents special opportunities, in that it has been studied fairly intensively within traditional-descriptive grammar, within generative grammar, and within functional-typological approaches. Since agreement and case are relatively accessible and one or the other is present on the surface in most languages of the world, they are perhaps the topics for which the interests of these three traditions overlap the most. Hence, we have better than average opportunities to compare the results of these traditions in this domain.

In what follows, I concentrate on telling a connected and accessible story that touches on some highlights of generative research in this area. I do not concentrate on the historical development of the relevant ideas, or on some of the narrower technical issues that arise, but rather on some points of contact with typological research. The story as I tell it has a personal dimension, in that I use some of my own research to connect the dots. However, it includes many points that most researchers would agree have broad and continuing significance. (For a more neutral handbook-style exposition of some of this same territory, see Baker in press-a.)

3. Agreement on verbs versus adjectives

3.1 A difference in agreement

Indeed, I begin the story proper in a somewhat idiosyncratic place. At the center of most generative work on case and agreement has been the fact that the finite verb agrees with the nominative subject—not with the direct object, or with a subject that bears some case other than nominative. But rather than starting there, I approach the topic by way of a more accessible and less controversial fact: that verbs and adjectives have systematically different agreement properties in languages that distinguish the two. The difference is shown in (5) and (6) for Spanish.

- (5) a. Est-as mujer-es com-en manzanas.
 these-F.PL women.F-PL eat-3pS apples
- b. Los hombre-s com-en manzanas.
 the.M.PL man.M-PL eat-3pS apples
- c. Nosotras com-emos manzanas.
 we.F.PL eat-1pS apples
- (6) a. Est-as mujer-es son gord-*as*.

¹ The flipside of their relative accessibility is the fact that case and agreement may include a degree of irregularity that is typical of topics that have a morphological component: languages sometimes have “exceptions” which do not agree or mark case according to the general pattern, and one may need to see beyond these to discern the universals. More generally, the generative notion of UG was introduced primarily to deal with “poverty of stimulus” issues, where the data the child has to work with seems to underdetermine the grammar he or she internalizes (for recent discussion, see Crain and Pietroski 2001). Since case and agreement are so visible and so common in languages that have them, one might expect there to be no serious poverty of stimulus issues in this domain, so UG would play less of a role. It would not be too surprising, then, if universals in this domain would be fewer and/or noisier. Nevertheless, I show that universals can apparently be found even in this domain, and I take that to be an a fortiori argument that the impact of UG on particular languages is indeed deep.

these-F.PL women-F.PL are.3sS fat-**F.PL**

b. Los hombre-s son gord-**os**
the.M.PL man.M-PL are.3sS fat-**M.PL**

c. Nosotras somos gord-**as**. (*gord-**amos**)
we.F.PL are.1pS fat-**F.PL** fat-**1p**

In Spanish, verbs and predicate adjectives both agree with the subject of the clause, but they do so in different ways. Verbs agree in person and number, but not gender, whereas adjectives agree in number and gender but not person. As a result, the verb form in (5c) is different from the one in (5a), even though both subjects are feminine and plural. In contrast, the adjective form in (6b) is different from the one in (6a), showing that adjectives agree with their subjects, but the adjective form in (6c) is the same as the one in (6a), showing that they do not agree in person. Crosslinguistic comparison shows that the failure of the verb to agree in gender is an accidental property of Indo-European languages: verbs also agree with their subjects in gender in many languages, including Semitic, Bantu, Mohawk, etc. But the failure of the adjective to agree in person is crosslinguistically robust; it can be seen in languages from around the world, including Semitic, Bantu, New Guinean, Australian, and Amazonian languages (Baker 2008:ch.2).

Sometimes this fact about agreement might be used as part of the definition of adjective, but that is not general enough. For example, the adjective-verb distinction exists also in languages without agreement, such as Japanese. There are also independent tests for distinguishing adjectives and verbs, depending on the language. For example, adjectives can often modify nouns directly, whereas verbs only do so with the help of participial or relative heads, and verbs are inflected for a full range of tenses, moods, and aspects, whereas adjectives often are not. The question arises, then, what this cluster of properties that distinguish adjectives from verbs has to do with the agreement difference shown in (5) and (6).

3.2 A formal-syntactic account of the difference

In Baker 2008, I offered the following answer. At the heart of the adjective-verb distinction is the fact that adjectives are less directly connected to their subjects than verbs are. The relationship between a predicate adjective and its subject is mediated by a copular verb or particle; the relationship between a verb and its subject is not ((Baker 2003), with many precedents in the work of others). The difference can be expressed in the schematic structures in (7), where VP is not necessarily distinct from clause, and PredP is not necessarily distinct from VP.

- (7) a. [_{VP} we fall]
b. [_{PredP} We are [_{AP} fat]]

Now suppose we say that agreement in number and gender is allowed to happen at a nontrivial syntactic distance, but agreement in first and second person is not. In Baker 2008, I formulate this idea as the *Structural Condition on Person Agreement* (SCOPA), stated as follows:

- (8) A word F can agree with an NP in the features +1 or +2 if and only if a projection of F merges directly with that NP (F being taken as the label for the resulting phrase).²

The difference between adjectives and verbs follows from this together with the schematic structures in (7). Simply put, the subject is part of the VP headed by the verb but not part of the AP headed by the adjective, and this accounts for why the verb can agree with it in person but the adjective cannot. Note also that the copular verb that is in construction with the predicate adjective can agree in person according to (8), and in fact it does, as seen in (6).

What is the advantage of looking at the agreement difference between verbs and adjectives in this structure-based way? Why not simply stipulate it as a fundamental difference between verbs and adjectives, as traditional descriptive treatments do? The big advantage is that (8) predicts that when verbs are somehow forced to agree with something further away than usual, then they too will agree in number and perhaps gender but not in first or second person. This is confirmed in some interesting cases. For example, in Icelandic when the subject of the clause is in dative case rather than nominative, the verb agrees with the object rather than the subject, as shown in (9a,b). However, this agreement can only be for third person, number, and (if the verb is a passive participle) gender; if the object is first or second person, such examples are ruled out, as in (9c) (Taraldsen 1995; Sigurðsson 1996).³

- (9) a. Henni leiddust þeir. (Icelandic)
 She-DAT was.bored.by.3p they.NOM
 ‘She was bored with them.’
- b. Henni leiðist bókin sín.
 her.DAT was.bored.by.3s book self’s
 ‘She was bored with her own book.’
- c. *Henni leiddumst við.
 She-DAT was.bored.by.1p we.NOM
 ‘She was bored with us.’

A similar effect can be found in some languages when the verb agrees with something trapped in an embedded clause, as shown in (10) for the Nigerian language Lokaa. Verbs in Lokaa normally agree with their subjects in person, as well as in gender and number, as shown in (10a). (10b) shows that if the subject is a nominalized verb, the agreement on the main verb is *ke-*, matching the nominalizing prefix *ke-*. (10c) shows the crucial complication: if the nominalized verb in subject position is transitive, then the main verb agrees with the object of the nominalized verb, not with the nominalized verb itself. Thus the prefix on the main verb in (10c)

² The parenthesized phrase in (8) is intended to prevent an attributive adjective from agreeing in first or second person with the NP that it modifies. But since adjectives rarely modify first or second person pronouns anyway, we can ignore this detail here.

³ More precisely, this follows from (8) once one recognizes that the agreeing element is Tense (T); see section 4 for discussion. The subject then moves to Spec, TP, but the object does not, explaining why T agrees with the subject in person in (2) and (5), but not with the object in person in (9) (see (27b)).

There are other languages in which the equivalent of (9c) is grammatical, including Russian. However, these seem to be languages in which the dative NP has fewer subject properties than in Icelandic, opening up the possibility that the dative NP does not occupy Spec, TP, so that the object can; see Baker 2008:90-94 for discussion.

is *e-* matching *e-sau* ‘fish’, not *ke-* matching *ke-dei* ‘buying’. In structural terms, this is a kind of agreement at a distance, since the verb does not agree with the phrase that is combined directly with it as its subject, but rather with something properly contained within that phrase. (8) thus predicts that person agreement is not possible on the main verb if the object of the nominalized verb is first or second person. (10d) shows that this is correct.⁴

- (10) a. **Ami n-tum n-dam.**
 I 1sS-be.very 1sS-be.big
 ‘I am very big.’
- b. **Ke-paala ke-tum ke-tawa.**
 GER/5-fly 5S-be.very 5S-be.difficult
 ‘Flying is very difficult.’
- c. [**E-sau ke-dei**] **e-tum e-tawa.** (*...ke-tum ke-tawa)
 7-fish GER/5-buy 7S-be.very 7S-be.difficult
 ‘Buying fish is very difficult.’
- d. *[**Min ke-funna**] **n-tum n-tawa.**
 me GER/5-surprise 1sS-be.very 1sS-be.difficult
 ‘Surprising me is very difficult.’

Note that we cannot say that (9c) or (10d) is bad because the verbs lack first person forms in their inflectional paradigms, parallel to what one might say for the Spanish adjective in (6c). The verbs clearly do have first person forms in Icelandic and Lokaa; they simply cannot use those forms in these situations, given the SCOPA.

The SCOPA thus finds a common cause as to why first and second person agreement is possible in (5c) and (10a) but not in (6c), (9c) and (10d), and this unified explanation cuts across the verb-adjective distinction. It is not the part of speech that is crucial, but the syntactic configuration—although this is related to the part of speech in many instances, because a word’s part of speech largely determines what syntactic configurations it can appear in. This line of reasoning can be extended to explain why verbs show person agreement with their indirect objects but not with their direct objects, and other more specific results in particular languages (see Baker 2008 and in press-b).

3.3 Comparison with typological research

This result of generative syntax is related to a result of typological research in an instructive way. Stassen’s (1997) extensive study of intransitive predication led him to the following generalization (see also Croft 1991:82):

⁴ Some Algonquian languages seem to have long distance agreement in person as well as in number and gender, but Bruening (2001) shows that putatively long distance agreement with first and second person pronouns has some subtly different properties which reveal it to be local agreement; see Baker 2008:106-107 for an interpretation that is consistent with the SCOPA. See also Bobaljik and Wurmbrand 2005 for discussion of some cases that look like long distance agreement in person but are not, since the agreed-with NP has moved into the matrix clause, as shown by scope tests.

(11) The Agreement Universal

If a language has person agreement in intransitive main clauses, this agreement will be used in sentences with event predicates.

One interesting fact about this generalization is that it is exceptionless in Stassen's fairly large sample (410 languages). A second interesting thing about it is that there was no obvious explanation for it within Stassen's functionalist framework. Thus, Stassen writes (p. 38):

As far as I am aware, no known principle of linguistic theory prevents us from imagining a possible natural language with some form of person agreement where this agreement can be used only with predicate adjectives or only with predicate nouns, but not with predicate verbs. However, in my sample I have not found a single instance of a language in which such a situation can be observed. In this book, no attempt at an explanation for the Agreement Universal will be made.

So there are substantive universals that hold up under investigation—at least they remain valid in samples that are an order of magnitude larger than Greenberg's (1963) original sample, and only one order of magnitude smaller than the total number of languages spoken in the world today, by *Ethnologue* standards (Lewis 2009). Some recent writings by functionalist-typologists have forgotten this.

Furthermore, this universal *can* be given a sort of explanation in formal generative terms, using (8). Characteristically, this is an internal explanation, explaining one linguistic fact in terms of other linguistic facts, not an external explanation that explains a linguistic fact in terms of nonlinguistic facts, whether cognitive, social, pragmatic, or historical. However, it counts as a genuine explanation inasmuch as it relates one linguistic fact to a broader pattern of linguistic facts in an interesting way. One can always hope for a still deeper explanation, one that would explain why the SCOPA itself should be true (see Baker 2008:ch. 4 for an effort in this direction), and perhaps that would take us beyond the bounds of the narrowly syntactic, where generative linguists are at their best. But one should not downplay language-internal explanations of this sort even if one looks to go beyond them. On the contrary, such explanations can tell us more accurately where to look.

Although they are clearly related, Stassen's Agreement Universal is not identical to the generalization about agreement on adjectives and verbs that I set out to explain. There are two prominent differences. First, Stassen's universal distinguishes state-denoting predicates from event-denoting predicates, rather than adjectives from verbs. This is a semantically-oriented distinction rather than a syntactically-oriented one, and the two are not equivalent. It may be true that all adjectives are state-denoting predicates, but there are verbs that express states rather than events. This is true even in English, but it is more striking in Mohawk, where even a predicate like 'big' counts as a verb. For example, 'big' in Mohawk inflects for tense and aspect ((12a)), and it can incorporate its nominal argument the way many verbs can in this language ((12b)) (Baker 2003:4, 249).

- (12) a. Ra-kowan-v-hne' ne Sak. cf: t-yo-ya't-v'-v-hne'
 MsS-be.big-STAT-PAST NE Sak CIS-NsO-body-fall-STAT-PAST
 'Sak used to be big.' 'It has fallen.'
- b. w-a'shar-owan-v. cf: t-a'-ka-wis-v'-ne'

NsS-knife-be.big-STAT
'The knife is big; it is a big knife.'

CIS-FACT-NsS-glass-fall-PUNC
'The glass fell.'

The second difference between Stassen's universal and mine is that Stassen phrases his as an implicational statement rather than as a biconditional one. He does not rule out person agreement on state-denoting predicates; he simply says that these cannot agree in person unless event-denoting predicates (i.e., canonical verbs) also agree in person. Stassen's universal thus permits the three kinds of language in (13a-c), but not the kind in (13d).

- (13) a. a language that has person agreement on both event predicates and state predicates
b. a language that has person agreement on event predicates but not on state predicates
c. a language that has person agreement on neither event predicates nor state predicates
d. *a language that has person agreement on state predicates but not event predicates

Spanish exemplifies (13b), as discussed above. Chinese, Japanese, and other languages that do not have agreement at all are (13c)-type languages; these are also consistent with (8), since this puts a limit on agreement but does not say that languages must have agreement.⁵ (13d) would include languages in which adjectives agree with their subjects in person but eventive verbs do not; such languages would violate (8), and Stassen and I agree that they do not exist. The one type of language where there might be a substantive mismatch between what I have derived and Stassen has discovered is (13a). If we translate this as saying that some languages have first or second person agreement on both adjectives and verbs, then Stassen allows for such languages and my SCOPA does not.

But we already that this is not a good translation, because languages like Mohawk have state-denoting predicates that qualify as verbs on grounds other than agreement. It is perfectly consistent with my generalization that a predicate like 'big' in Mohawk agree in first person with its subject, although predicates like 'fat' in Spanish do not. And this is true, as shown in (14).

- (14) i:ʔi k-kowan-v (Deering and Delisle 1976:197)
I 1sS-be.big-STAT
'I am big.'

I believe, then, that there is a good degree of match between my analysis and Stassen's results, and that my somewhat stronger statement is vindicated if one does the work of distinguishing systematically between predicates that are verbs and predicates that are adjectives.

But this is not the whole story regarding type (13a). There are also languages in which adjectives are distinct from verbs, but they nevertheless bear person agreement under some specific circumstances. One that I have studied is the Turkic language Sakha, illustrated in (15). Here both the verbal predicate in (15a) and the adjectival predicate in (15b) bear the same first person plural suffix *-BYt* (with regular phonological changes), distinct from the third person plural suffix *-lAr* seen in (15c),

⁵ Another sort of language in this category would be one that has only number and gender agreement on both verbs and adjectives, such as Tsez (Polinsky and Potsdam 2001) and other Caucasian languages. These are also consistent with both (8) and (12)—although how best to represent verbal agreement in formal synchronic terms in such languages remains an interesting question.

- (15) a. Bihigi bil-e-bit.
 we know-AOR-1pS
 ‘We know.’
- b. Bihigi bytaam-myt.
 we slow-1pS
 ‘We are slow.’
- c. Kini-ler bytaan-nar.
 they-PL slow-3pS
 ‘They are slow.’

Nevertheless, we know that predicates like *bytaan* ‘slow’ are adjectives because they can modify nouns directly, without the participial morphology or relativizing particles that are needed for verbs to modify nouns (*bytaan balyksyt* ‘a slow fisherman’). Moreover, a full range of tense suffixes can attach to verbs, but not to an adjective like ‘slow’ (Vinokurova 2005:221, 232).⁶

- (16) a. En aaq-yaq-yn.
 You read-FUT-2sP
 ‘You will read.’
- b. *En bytaan-yaq-yn; En bytaan buol-uoq-yn.
 *You slow-FUT-2sP You slow be-FUT-2sP
 ‘You will be slow.’ ‘You will be slow.’

See Baker and Vinokurova 2009 for other, more subtle expressions of the adjective-verb distinction in Sakha.

But when we look at the full range of environments that show that a word like ‘slow’ is an adjective in Sakha, we also see that the contexts in which these words bear person agreement are quite limited. The adjective does not bear person agreement in the grammatical version of (16b); only the copular verb ‘be’ inserted to bear the tense suffix does. Nor does an adjective bear agreement when it modifies a noun in Sakha: ‘slow fishermen’ is *bytaan balyksyt-tar*, not **bytaan-nar balyksyt-tar* (‘slow-(PL) fisherman-PL’). Nor does an adjective bear agreement with its subject when it is the complement of a verb like ‘become’ or ‘make’:

- (17) Bu miigin djolloox-(*pun) oŋor-but-a.
 this me.ACC happy-(1sS) make-PAST-3sS
 ‘This made me happy.’

Spanish adjectives do show (number and gender) agreement with an NP in all these environments. Overall, then, adjectives in Sakha agree with an NP if and only if they are the

⁶ This particular example is not ideal because the root *bytaan* can also be used as a verb in Sakha, with the eventive meaning ‘to become slow, to slow down’ (Vinokurova 2005), very much like the English pair ‘John is slow’ and ‘The train gradually slowed to a stop’. The fact that some adjectives have verbal twins in Sakha does not seriously blur the adjective-verb distinction in that language, however, especially since other adjectives do not have verbal doubles.

primary predicates in a clause that is interpreted as being in present or simple past tense, where there is no overt tense morpheme.

I take this to mean that adjectives per se do not agree in Sakha—not even in number, much less in person. Rather, the Tense element (T) agrees in person and number with the subject, and that tense element can cliticize to the adjective to form a single phonological word with it on the surface. This happens if and only if the tense marker and the adjective would find themselves linearly adjacent to each other, and if the tense marker is phonologically null. In other situations, the tense attaches to a verb root, as in (16) and (17), with the result that the inflected verb agrees in person with the subject, but the adjective does not. In Spanish sentences analogous to (16b), there are actually two agreements: person agreement on the finite copula, and number-gender agreement on the adjective, in accordance with (8) and the structure schematized in (7) (see (6)). Sakha has the analog as the first agreement, but not the second. Sakha is thus perfectly consistent with (8), when analyzed correctly. The structures are schematized roughly in (18).

- (18) a. [_{Pred P} We [_{AP} slow] Pred+PRES[1pS]] → We slow 0-0-1pS → we slow-1pS (= (15b))
 b. [_{Pred P} You [_{AP} slow] Pred+FUT[2sS]] → You slow be-FUT-2pS (= (16b))
 c. [This me-ACC [_{AP} happy] make PAST-3sS] → This me-ACC happy make-PAST-3sS

We have seen, then, two ways in which a state-denoting predicate might bear person agreement, illustrating Stassen's type (13a), and neither undermines the claim that adjectives do not agree in first or second person in the syntax. One is that the state-denoting predicate can be a verb rather than an adjective, as in Mohawk; the other is that a person-agreeing tense can cliticize to the adjective under some conditions, as in Sakha.⁷ The prediction is that once linguists do the analytical work needed to distinguish these cases—as one would have to do anyway to provide an accurate description of each language—then it will be seen that my stronger and more syntactic generalization is also true universally. No research that I know of shows this to be false, although much of the work remains to be done. To the extent that this is confirmed, generative work can offer internal explanations for typological universals, and it can offer ways of developing them into stronger and more precise statements.

4. Subject agreement associated with verbs

This provides us with some groundwork for approaching more standard topics in the generative literature on agreement and case. In fact, the kinds of the reasons for saying that Tense is a separate syntactic element from the adjective and the one that actually agrees with the subject in Sakha generalize to finite verbs in many languages, including English. The only significant difference is that it is very common for verbs to join with Tense into a single word on the surface, whereas this is less common for adjective (probably because the predicative particle intervenes between them; see Baker 2003:46-52). However, Tense and the verb can be pulled apart under some circumstances, and then agreement clearly goes with Tense, not the verb. One case in point is verb-plus-auxiliary constructions, which look very much like adjective-plus-copula constructions. (19) is an example from Spanish (compare (6c)). Notice that the subject agreement is on the tense-bearing auxiliary verb, not on the main verb, which is in invariant participial form.

⁷ See Baker 2008:59 for some discussion of the less clear case of Nahuatl, another putative type-(13a) language.

- (19) Nosotros esta-mos comiendo manzanas.
We are.PRES-1pS eat-PTPL apples
'We are eating apples.'

Similar evidence comes from the fact that subject agreement is usually not present on nonfinite verbs, which also do not have (the same kind of) tense marking, as in (20b).

- (20) a. The child go-es to school every day.
b. I made the child go-(**es*) to school every day.

There are also more language-particular considerations that point in this direction, such as the well-known English fact that, when tense is prevented from attaching to the verb by (say) the presence of negation, then agreement shows up along with the tense marking on the dummy verb *do*, rather than with the verb:

- (21) a. The child does not go to school every day.
b. *The child do not goes to school every day.

These facts do not necessarily force one to treat Tense as a separate syntactic element from the verb, but they make such a treatment plausible. And if the distinction is accepted, then it is clear that agreement is associated with Tense rather than with the verb proper.

This brings us to another point of contact between generative syntax and typological research. It is well-known in the typological literature that it is more common for a verb to agree with its subject than for it to agree with its object. Hence (22) has often been invoked as a robust if not absolute universal (Moravcsik 1974; Croft 1990; Siewierska 1999):

- (22) Verbs in a given language agree with their objects only if (some, finite) verbs in the language agree with their subjects.

All the languages cited in this article so far are consistent with (22): verbs agree with their subjects and not their objects in English, Spanish, Sakha, Lokaa, and Icelandic; verbs agree with their subjects as well as with their objects in Kinande and Mohawk; verbs agree with neither in Japanese. A small number of counterexamples to (22) have been proposed in the literature, but by all accounts they are few and far between.⁸ So it is well-worth considering why natural languages should favor subject agreement in this way.

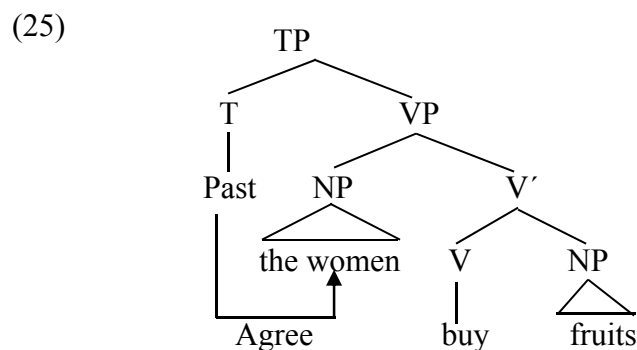
Although he does not discuss it in exactly these terms, the theory of agreement in Chomsky (2000; 2001) provides the outline of an explanation for this salient fact. His account

⁸ Siewierska (2005a) reports a relatively large number of languages—24 out of a sample of 378—which have agreement with the Patient (roughly the object) but not with the Agent (roughly the subject), although she says that this “is an infrequent phenomenon everywhere.” This may be because she considers only agreement marked on the verb, putting aside agreements that appear on a tense marker or other particle that does not attach to the verb. Indeed, the example she cites of a language with verbs that agree only with the object (Yawa) does have a separate particle that agrees with the subject. Such a language is consistent with (22) if it is generalized to say “If some head in the clause agrees with the object, then some head in the clause agrees with the subject”, leaving open the possibility that Tense may or may not combine morphologically with the verb. Siewierska also explicitly does not distinguish true agreement from clitics in her typology, and this contribute to her high count.

rests on the hypotheses about syntactic structure in (23), together with the fundamental property of agreement in (24).

- (23) a. Tense is the category in a clause that is most likely to agree with something.⁹
 b. Tense is the highest category in the syntactic representation of the clause, taking scope over the proposition as a whole.
 c. Agentive subjects are structurally higher than other NPs in the canonical organization of the clause.
- (24) An agreeing head A agrees with the closest eligible noun phrase B (technically, with the eligible NP B such that A c-commands B and there is no other eligible NP X that c-commands B and is c-commanded by A.¹⁰)

Given (23b,c), the initial syntactic representation of a typical clause like (2b) will be approximately (25).¹¹



Given (23a), it is T that will agree if anything does. Then the agreement principle in (24) applied to the representation in (25) implies that T agrees with the subject rather than the object, since the subject is lower than T but higher than the object or any other NP in the clause.

Once again, one might ask what the advantage is of casting this in syntactic terms, rather than simply stating (22) directly. One advantage is that the syntactic account uncovers interconnections among many morphosyntactic phenomena. Thus, there is substantial independent support for most or all of these assumptions. For example, Tense is one of the outermost affixes on a morphologically complex verb in most languages (Bybee 1985), and this is consistent with (23b) given that the layering of the morphemes on the complex verb mirrors

⁹ We should probably generalize “Tense” in this statement to something like “the functional category that anchors the sentence as a whole, placing it with respect to the speech act.” That category could be tense, or mood, or aspect. Stated in this way, the theory applies also to a language like Mapudungun, which arguably has mood (indicative vs. imperative vs. conditional) but not tense (cf. Smeets 2008), with Mood bearing agreement, and to Amharic, which has aspect rather than tense, with aspect bearing agreement (cf. Kramer 2010). In this article, I use Tense loosely as a cover term for this range of categories.

¹⁰ A category X c-commands a category Y if and only if the smallest category that properly contains X also contains Y (Reinhart 1976). I return to the definition of “eligible” below.

¹¹ The subject NP subsequently moves to Spec, TP in many languages to derive Subject-(Tense)-Verb-Object order; see note 3. This movement is crucial for T to agree with the subject in first and second person as well as in number and gender, given the analysis in section 3.

the layering of elements in the clause it is generated in (Baker 1988; Cinque 1999). This also coheres with some formal semantic treatments of tense. Given this, (23a) is independently motivated by data like that in (19)-(21). The fact that agentive subjects are higher than other NPs in the clause ((23c)) plays a central role in the generative account of many other phenomena, including word order (the subject comes before the object in the vast majority of languages; cf. Kayne 1994), binding relations (subjects can bind anaphors in object position, but not vice versa; cf. Reinhart 1976), the relative scopes of quantifiers (Aoun and Li 1993), and so on. Finally, the closeness condition on agreement in (24) is also independently motivated: attributive adjectives invariably agree with the closest NP—the NP they modify—not with some more remote one; adpositions agree with the NP with which they form a phrase, not with some more remote NP, and so on. So (23) and (24) provide an internal explanation of the fact that finite verbs typically agree with subjects rather than objects in that it reveals important relationships among superficially different-looking phenomena. Even if these are not ultimate explanations, they can be satisfying and useful intermediate explanations.

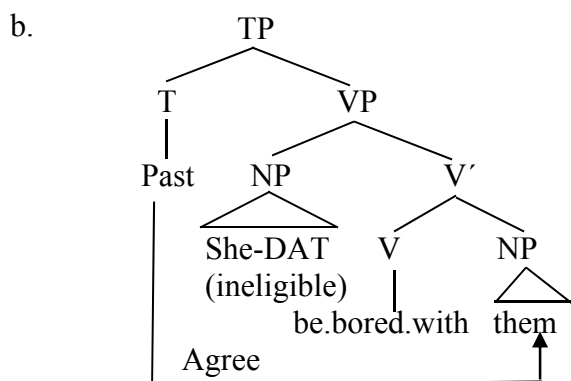
A second advantage of the structural account of the predominance of subject agreement is that it has the potential to be more precise than accounts in informal, traditional terms. Some functionalist-typological researchers try to avoid theory-laden notions like subject and object by referring to thematic relations instead: they say that in languages like English the verb agrees with its agent argument, not its patient argument (e.g., Siewierska 2005a). But this cannot be taken very literally, since in examples like (26), which have a patient argument but no agent argument, the verb does agree with the patient.

- (26) a. A vase breaks every time there is an earthquake.
 b. Mary is often criticized by the press.

In contrast, it follows from (24) that if there is no agentive NP, but only a patient NP, then the patient NP is the highest NP, and T automatically agrees with it.

Neither is the traditional notion of subject quite accurate for all languages. I mentioned above that in Icelandic the subject of certain predicates is marked with dative case rather than with nominative case. When this happens, the finite verb does not agree with the subject:

- (27) a. Henni leiddust þeir. (=9a)
 She-DAT was.bored.by-3p they.NOM
 ‘She was bored with them.’



There are many excellent reasons to say that *henni* in (27) is the subject of this clause: indeed, it behaves grammatically like more canonical subjects in virtually every way other than case marking and agreement (Zaenen, Maling et al. 1985). So the traditional statement fails in this example, but (24) succeeds, on the understanding that dative case subjects are not eligible for agreement in Icelandic (Chomsky 2000; Bobaljik 2008). This is a new stipulation in the account, of course. But the important result is that, once the highest NP is declared ineligible for agreement, it follows automatically from (24) that T will now agree with the second highest NP in the clause, in this case the object (in number only, as discussed above). The same thing can be seen in the well-known case of split ergativity in Hindi, where in some perfective tense-aspects the subject of a transitive verb is not nominative but has a special ergative case marking *-ne*. In exactly those tense-aspects, the finite verb no longer agrees with the subject, but rather agrees with the object if there is one that is itself eligible in the sense of not being marked with an incompatible case. (If the object in (28b) bears the case marker *-ko*, then the finite verb agrees with nothing, but is *becaa*, a default third masculine singular.)

- (28) a. Anil kitaabẽ becegaã (Mohanan 1995:83)
 Anil.M(NOM) book-F.PL sell-FUT.M.SG
 ‘Anil will sell (the) books.’
- b. Anil-ne kitaabẽ becĩĩ
 Anil-ERG book-F.PL sell-PERF-F.PL
 ‘Anil sold (the) books.’

Indeed, if the subject is ergative and the verb has a clausal complement rather than an object, then T can, under certain conditions, look even further down, into the embedded clause, agreeing with the highest overt NP inside that clause (the object, since the subject is either absent or controlled). This is shown in (29).

- (29) Rohan-ne aaj [phir-se mehnat ker-nii] chaah-ij.
 Rohan-ERG today again hardwork.F do-INF.F want-PREF.F.SG
 ‘Today Rohan wanted to do hard work again.’ (Bhatt 2005:766)

(24) can thus be made to cover important details of how agreement works better than more traditional notions like agent and patient or subject and object.¹² (I return below to the relationship between case and eligibility.) So we see that if there is one agreement in the finite clause, it will be *roughly* subject agreement—but the syntactic details matter, helping to sharpen up the “roughly”.

5. Object agreement associated with verbs

Now let us consider the other sort of language permitted by the universal in (22), a language that allows two agreements on or near the verb, hence one that has (roughly!) object agreement as

¹² Of course, there might be other details for which the generative account would also need elaboration to get all the facts right: for example, equational sentences like “It is I” may pose special problems in some languages, as do sentences with conjoined subjects. But there are good prospects for extending it to these domains as well.

well as subject agreement. These make up approximately 50% of languages in the world, according to Siewierska 2005a (193/378). Two examples are given in (30), one from Nez Perce (spoken in Idaho, USA), and one from Mapudungun (spoken in Chile).

- (30) a. **'In** wéet'u 'itúu-ne **'aa**-p-sá-qa. Nez Perce (Deal 2010:78)
 1sg NEG INDEF-ACC (1sS+)3sO-eat-IMPF-PAST
 'I didn't eat anything.'
- b. Llūka-l-ka-ke-fu-fī-y-ngün ti nge-n ruka.
 become.afraid-CAUS-FAC-HAB-PAST-3O-IND-3pS the have-NOMLhouse
 'They used to intimidate the house owner.' Mapudungun (Smeets 2008:228)

The proper generative analysis of object agreement is somewhat less certain, because the best-studied languages happen not to have it, but the analytical possibilities are relatively clear.

In fact, at least three kinds of analysis are possible within these terms. First, the highest head T could be permitted to agree twice, hence with the two highest NPs, by a natural generalization of (24). Second, there could be a second category in the clause, lower than T and the subject, but higher than other elements, which agrees (once) with the NP that is closest to it. Third, some putative instances of object agreement could have quite a different source: they could be phonologically weak pronouns that are attracted to the verb ("cliticized"), not true agreement on par with finite verb agreement and adjectival agreement.

In our current state of knowledge, it seems that each of these possibilities may occur in the world's languages. If so, then object agreement will turn out to be a relatively heterogeneous phenomenon. For this article, I put aside the possibility of object clitics, since those are conceptually less related to my topic. I simply note that the topic is important because such clitics may easily be mistaken for true object agreement (and vice versa), and learning to distinguish between clitics and agreement is important if we are to have realistic chances of identifying universals of agreement (see, for example, note 8). For some recent proposed diagnostics, see Nevins 2010, Preminger 2009, Kramer 2011, Baker 2008, in press-b, and Coppock and Wechsler 2010, among others.

Focusing on the other two possibilities, then, how might one tell if a language has a single head (Tense) that agrees twice, or two heads that agree once each? One way could be to consider nonfinite clauses. Recall that, in many languages, if finite tense is replaced by infinitival or nominal morphology, the ability of the verb to agree with the subject is lost (see (20)). That was one sort of evidence that subject agreement is really a characteristic of T, not of the verb itself. Now if object agreement in a given language is the result of T agreeing a second time, then we expect object agreement also to be absent in nonfinite clauses. This is what we find in Nez Perce (Amy Rose Deal, personal communication; also Chukchi, Yimas, etc.):

- (31) [weet'u'ituu-ne mic'ii-t'as] qooqoX hi-wc'ee-y-e.
 not what-ACC hear-PTPL raven 3S-become-P-PAST
 'She became a raven, unable to hear anything.'

In contrast, if object agreement is the result of a different head agreeing, then we expect that it could survive the replacement of T with a nonfinite version, hence it could be found in nonfinite clauses. This is what we find in Mapudungun (also Spanish, many Bantu languages, etc.):

- (32) Fey ayü-w-üy iñché ñi pe-me-fi-el fey.
 he love-REFL-IND.3sS I POSS.1sS see-TH-3O-INF he
 ‘He was happy at my going to see him.’ (Smeets 2008:216)

This is possible evidence that object agreement has different sources in different languages.

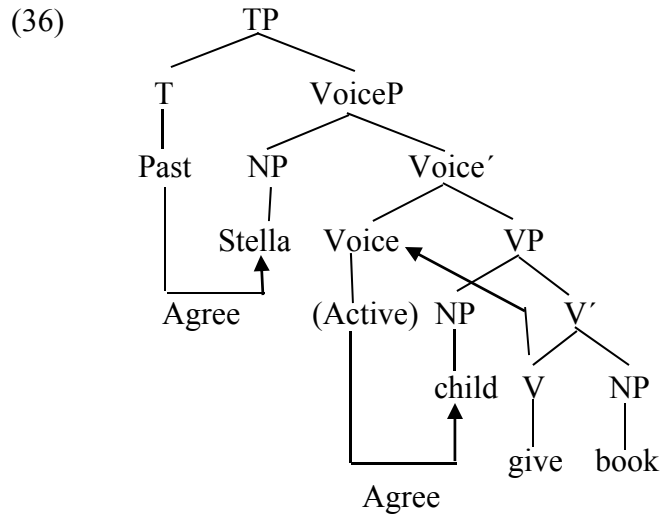
For languages in which object agreement is not associated with T, what head is it on? A reasonable answer might be the verb itself. But an even better answer might be the active voice marker—phonologically null in most languages, but assumed to exist in opposition to (say) passive or middle voice morphology. This captures the fact that, in a significant range of languages, the possibility of object agreement is lost in passives, middles, reflexives, or other kinds of nonactive verb. Mapudungun, for example, has ditransitive verbs like ‘give’, and in active sentences the verb shows object agreement with one of its two objects, as shown in (33a). Mapudungun also has a passive morpheme *-nge*, and when this combines with a ditransitive verb, one of the objects (the goal) becomes the subject and triggers subject agreement, as in (33b). However, there is still an object of sorts in (33b), and one might well imagine that the verb could show object agreement with this remaining object. However, (33b) shows that this is impossible in Mapudungun.

- (33) a. **Iñché** elu-Ø-fi-**n** kiñe metawe Roberto. (Active, with object Agr)
 I give-(ACT)-3O-IND.1sS one vessel Roberto
 ‘I gave Roberto a vessel.’ (Baker 2006:319)
- b. (**Iñché**) elu-nge-(*fi)-**n** ti metawe. (Passive, no object agreement)
 I give-PASS-(*3O)-1sS the vessel
 ‘I was given the metawe.’ (Baker 2003)

(32) shows that nonfinite verbs can have object agreement but not subject agreement in Mapudungun, whereas (33) shows that nonactive verbs can have subject agreement but not object agreement. This double dissociation suggests that subject agreement is a property of Tense and object agreement is a property of active voice in this language. This is sketched in (34a); compare (34b) as a structure for Nez Perce.¹³

¹³ The category I call Voice in (34) is known by many names in the generative literature, including v (Chomsky 2000). I assume that the verb moves into the Voice head in all the languages discussed, so that the two are pronounced as a single word. This is also important to derive the correct word order in examples like (36) and (37).

The structure of the Swahili version would be roughly (36); the structure for Nez Perce would be the same, except that T would agree twice rather than Voice agreeing once.

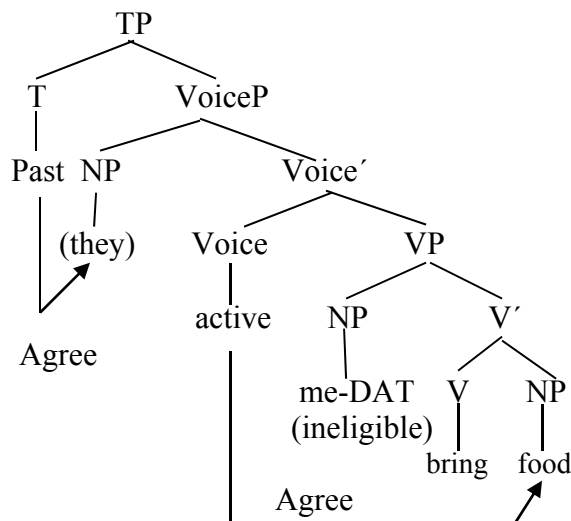


Indeed, there is a probable universal here: Haspelmath (2005:427) states (speaking specifically of verbs meaning ‘give’) that “Languages without case or adpositional marking in which only the patient and theme, but not the recipient, are indexed [=agreed with, MCB] are unattested” (in his sample of 378 languages).

6. The relationship of case and agreement

Recall, however, that subjects that are marked with dative or ergative case may count as ineligible for agreement with T. In that situation, T looks past the subject and agrees with the theme/object (see (27) and (28)). Now in many languages, the goal of a ditransitive verb can or must be marked with dative case. When it is, we might expect the goal to be ineligible for object agreement. In such a language, we will see object agreement with the theme rather than the goal—different from Swahili and Nez Perce. An example is (37) from Mangarayi (Australian).

- (37) a. Wula-niri-j ŋanju (∅)-mawuj. (Merlan 1982:65, 66)
 3pS/3sO-bring-PP me.DAT ACC-vegetable.food
 ‘They brought me vegetable food.’



b.

So the kind of case marking system a language has influences how object agreement works in that language. This parallels nicely what we saw for subject agreement, supporting the claim that the same formal syntactic theory of agreement applies to both.

There is interesting crosslinguistic variation when it comes to examples like (37), however. There are also languages in which the verb shows object agreement with the goal rather than the theme even though the goal is dative. An example is (38) from Burushaski, a language spoken in the Himalayas (Lorimer 1935).

- (38) a. (U:n) gu-yets Λ -m. (Voice agrees with ABS object)
 you-ABS 2sO-see-1sS
 ‘I saw you.’
- b. U:n $\text{\textcircled{r}}$ hik tr Λ n gu-čič Λ -m. (Voice agrees with DAT object)
 you-DAT one half 2sS-give-1sS
 ‘I shall give a half to you.’

As far as agreement is concerned, the representation of (38b) in Burushaski is more like (36) than like (37b), even though the goal is marked dative (see also Baker In press-b on Amharic).

This variation can be built into our definition of the notion “eligible”, which (24) makes reference to. We can *parameterize* the notion of eligibility as follows:

- (39) a. In some languages, only bare (uncase-marked) NPs are eligible for agreement.
 b. In some languages, NPs marked for structural case are still eligible for agreement.¹⁶

¹⁶ This statement assumes the often-made distinction between “structural” (or “grammatical”) case and “oblique” (or “semantic”) case. Structural case is case assigned to an NP by virtue of its being in a particular structural position; it includes case assigned via agreement with a functional category, case assigned by a rule of dependent case assignment (see (44) below), or case assigned by default. In particular, the category of structural case includes most instances of nominative, absolutive, accusative, and ergative case, and some instances of dative case. It does not include certain other instances of dative case, or most instances of locative, ablative, or instrumental case, etc. NPs that bear these oblique cases are usually invisible to agreement in either kind of language; indeed, they may not be syntactically different from PPs.

This is called the Case Dependency of Agreement Parameter in Baker (2008) (see (Bobaljik 2008) for a related idea). (39a) is the more familiar setting, roughly equivalent to the *activity* condition on agreement in Chomsky (2000, 2001). (39b) is the alternative setting, designed to bring object agreement in languages like Burushaski and Amharic into the fold.

All things being equal, we would expect the parameter in (39) to apply to agreement on T as well. We saw above that in Icelandic and Hindi the finite verb does not agree with a subject in dative or ergative case. This is the expected behavior for a language that follows (39a). But now we also expect to find languages in which the finite verb does agree with the ergative or dative subject: this would be a language that follows (39b) with respect to T. In fact, Burushaski itself is such a language: T agrees with ergative subjects and bare/absolute subjects indiscriminately:

- (40) a. Acaanák hilés i-ír-imi (Willson 1996:19)
 suddenly boy.ABS MsO-die-PAST.MsS
 ‘Suddenly the boy died.’
- b. Hilés-e **dasin** **mu-yeéts-imi**. (Willson 1996:17)
 boy-ERG girl.ABS FsO-see-PAST.MsS
 ‘The boy saw the girl.’
- c. Jé-e **ín** **é-s-kharan-am**. (Willson 1996:13)
 1s-ERG 3sh.ABS MsO-CAUS-be.late-PAST
 ‘I made him late.’

Burushaski is minimally different from Hindi in this respect; compare (40b) with (28b) above.

In Baker 2008, I claimed that this is no coincidence. I captured the fact that both subject agreement and object agreement are insensitive to the case of the agreed-with nominal in languages like Burushaski by interpreting (39) as a general parameter that governs languages as wholes, not individual heads or constructions in a language. So if a language has both subject agreement and object agreement, the prediction is that either they will both be case-dependent, or neither of them will be. I performed a study of 108 languages in an effort to evaluate this, and found that it was supported: of the 108 languages, only two showed inconsistent behavior in this regard (and I offered an alternative analysis for one of them). This then is another potential language universal in the area of case and agreement. It must be considered more tentative than others I have discussed, since it has been tested on a somewhat smaller sample, and has not yet been replicated by other researchers.¹⁷ However, one of this article’s goals is to show how generative theory can lead to the discovery of new universals as well as contributing to the explanation of known ones. This is a possible example of that.¹⁸

¹⁷ The survey is also smaller than it may look, since only 16 languages of the 108 have all the properties needed to investigate this fully. Nevertheless, the fact that 14 of those 16 behave consistently with respect to the parameter in (39) looks statistically significant if not truly universal: it seems that the behaviors of subject agreement and object agreement in a language are not formally independent of each other. (A precedent to my analysis of languages like Burushaski is Dryer’s (1986:842n.33) brief discussion of Warlpiri.)

¹⁸ A complication is that there seem to be languages (including Burushaski: Willson 1996:5) in which T agrees with nominative and ergative NPs but not dative NPs—a possibility allowed by Bobaljik’s (2008) version of (39) but not mine. But given that Burushaski allows object agreement with datives, one cannot simply stipulate that dative is invisible to agreement, as Bobaljik does. One somewhat technical possibility is that T is fussier about what it can agree with than Voice is because of EPP concerns: dative NPs might be embedded in null PPs and this bothers T

Bobaljik (2008) points out another typological universal that follows from this type of theory, one discussed previously by Dixon (1994:94-97). It is well-known that the case marking on NPs in a language can either follow a nominative-accusative pattern, where subjects of transitive and intransitive verbs receive the same case marking and a different case is used for the object of transitive verbs, or an ergative-absolutive pattern, where objects and the subjects of intransitive verbs receive the same case marking and a different case is used for the subject of transitive verbs. Japanese is an example of the nominative-accusative system ((4)); Burushaski is an example of the ergative-absolutive system ((40)). A similar distinction can be made for agreement systems: we call the agreement system of a language nominative-accusative if it shows the same kind of agreement with subjects of transitive and intransitive verbs but a different kind of agreement with objects; we call an agreement system ergative-absolutive if it shows the same kind of agreement with objects and the subjects of intransitive verbs, but a different kind with subjects of transitive verbs. Then we can ask if a language's case system and its agreement system always match in this respect. The answer is often yes, but not always. Burushaski is an example of a mismatch: its case system is ergative-absolutive, but its agreement system is nominative-accusative, as seen in (40). Interestingly, the opposite kind of mismatch is unattested: no known language has a (consistently) nominative-accusative case system and an ergative-absolutive agreement system. Dixon (1994) only cited a couple of examples, but the generalization is confirmed by *The World Atlas of Language Structures*. (41) gives the distribution of languages, crossing Comrie 2005 with Siewierska 2005b (188 shared languages).

(41)

Case \ Agreement	Nom-Acc	Erg-Abs	Neutral (no marking)
Nom-Acc	34	0!	14
Erg-Abs	12	3 (+6 split)	9
Neutral (no marking)	52	5	21

So there are no counterexamples to this typological asymmetry in *WALS*.

The asymmetry follows from the generative theory as developed here. Suppose that whether a language has nominative-accusative case or ergative-absolutive case is or can be determined independently of agreement (see section 7). What then are the options for agreement? We have said that T always agrees with the closest eligible NP in a (39a)-type language, or with the closest NP of any sort in a (39b)-type language. In the first type, the case system guides the agreement system, giving languages that are nominative-accusative in both respects or ergative-absolutive in both respects, the circled types in (41). In the second type, agreement ignores case and always agrees with the subject: that gives a nominative-accusative agreement system, regardless of whether the case system is nominative-accusative or ergative-absolutive (or neither). This includes the boxed type in (41). But there is no way to generate a language with nominative-accusative case and ergative-absolutive agreement. In such a language, T would be looking beyond the nearby subject to agree with the more remote object, even though the subject bears a case that T can agree with, as shown by the fact that it agrees with the subject of an intransitive clause. So T has no motive to do this; it would be a violation of (24). Therefore, the generative theory explains this typological universal with no new assumptions.

more than Voice because the null-headed PP cannot satisfy T's EPP feature. (This suggestion is modeled on my (to appear-b) analysis of Amharic.)

We can draw one additional, more tentative result from about the relationship of case and agreement. Chomsky’s (2000, 2001) version of (39a) is actually a bit stronger: he claims that a head F can only agree with an NP that is not already case marked *because* F must itself assign case to NP in the process of agreeing with it. For example, T agrees with NP if and only if T assigns NP nominative—a connection that is well-supported in many Indo-European languages, among others. Similarly, active Voice agrees with an NP if and only if it assigns NP accusative. (This connection has less robust empirical support, but Mangarayi may be a relevant language (Baker to appear-b)). In Baker 2008, I pointed out another prediction that this makes: it should be bad for two heads to agree (fully) with the same NP in (39a)-type languages, but not in (39b)-type languages. The reason is that, if a lower head agrees with a given NP in a (39a) language, then it must assign that NP case. This makes the NP ineligible for agreement with a higher head. However, this reasoning does not carry over to a (39b) language, where both the lower and the higher head can agree with a given NP without assigning it case, whether it already has case or not. Hence, we should expect to see a correlation between whether a language allows agreement with ergative and dative NPs or not, and whether it allows multiple full agreement in auxiliary constructions or not. And I claim that we do. In Spanish, for example, T agrees with nominative subjects but not dative ones, and only the auxiliary agrees with the subject, not the main verb, as shown again in (42).¹⁹

- (42) Nosotros esta-mos comiendo manzanas. (=19))
 We are.PRES-1pS eat-PTPL apples
 ‘We are eating apples.’

In contrast, in Burushaski, T agrees with both absolutive and ergative subjects, active Voice agrees with dative and absolutive objects, *and* both verbs can agree with the same subject in person, number, and gender. This is shown in (43) (see also (10a) from Lokaa).

- (43) Ja be.ʌdʌpi.en et-a b-a. (Lorimer 1935)
 I.ERG discourtesy do-1sS be-1sS
 ‘I have committed a discourtesy.’

Looking more broadly, in the 108-language sample of Baker 2008, I found 20 languages that disallowed multiple agreement and actively avoided agreement with case-marked NPs, 11 languages that allowed both multiple agreement and agreement with case-marked NPs, and no languages that allowed multiple agreement with a single argument but avoided agreement with NPs in ergative or dative NPs (if any). The numbers are not overwhelming, because not so many languages have all the relevant properties. But this is another candidate for a linguistic universal discovered by generative theory, in that it has both a plausible derivation from independently motivated principles and some nontrivial empirical support.

7. Implications for case assignment

¹⁹See Baker (2008) for certain qualifications that allow a participial main verb to have adjective-like agreement in number, gender, and case, but not verb-like agreement in person in (38a)-type languages, including some Indo-European languages.

Finally, I note briefly an implication of this for the theory of case assignment. If the parameter in (39) is more or less correct with Chomsky's interpretation of (39a), then we know how structural case is assigned in (39a)-type languages: it is assigned by relevant functional heads under agreement. In particular, nominative is assigned by T to the subject under agreement, and accusative is assigned by active Voice to the object under agreement. This is the standard Chomskian picture, and it holds for a significant range of languages. But at least two questions remain. First, where does case come from in (39b)-type languages? It cannot (always) come from agreement with a given functional head, because these languages are recognized by the fact that the same head can agree with NPs bearing different cases. Second, where do ergative and absolutive case come from in languages with ergative-absolutive case systems? Such case might (or might not) cause T to agree with the object rather than the subject, but this presupposes that ergative case is already in place before T starts to agree. Where did it come from?

Fortunately, suitable answers to both of these questions are available in the generative literature. Marantz (1991) argues for an alternative form of case assignment, where case does not come from functional heads per se, but is determined by an algorithm that simply counts the NPs present in a certain local domain, such as the clause. Marantz's idea can be expressed as follows (Baker and Vinokurova 2010):²⁰

- (44) a. If there are two distinct NPs in the same clause such that NP1 c-commands NP2, then value the case feature of NP2 as *accusative*.
- b. If there are two distinct NPs in the same clause such that NP1 c-commands NP2, then value the case feature of NP1 as *ergative*.

(44b) is sufficient to account for the ergative case in (40b,c) in Burushaski; other NPs are left bare (i.e., absolutive). (44a) is sufficient to account for the accusative case in (4b) in Japanese; other NPs are marked in other ways. Since (44) makes no reference to functional heads or to agreement, it can be the source of structural case in (39b)-type languages, where agreement functions independently of case. (See also Bobaljik and Branigan 2006 and Baker and Vinokurova 2010 for an extension of this view to dative case in certain languages.²¹) Finally, (44) accounts for ergative-absolutive case systems just as easily as it does nominative-accusative systems, as desired.

Following this logic through might make an additional typological prediction: it predicts that languages with ergative-absolutive case systems should be a subset of the (39b) languages, since (44b) is the only relevant source of structural ergative case, given that no known factor would cause a functional head to agree with the subject in transitive sentences only, except perhaps for a kind of case sensitivity (more like Bobaljik's than Chomsky's) that assumed that such subjects were already ergative. More precisely, we might expect about 50% of (39b)-type languages to be ergative, but about 0% of the (39a)-type languages to be ergative. I investigate this (briefly) in Baker to appear-a, by crossing my 2008 parameter with Comrie's (2005)

²⁰ See also Yip et al. 1987 and Comrie 1981:118-120 for similar ideas.

²¹ Baker and Vinokurova 2010 argue at length in favor of (44a) as the rule of accusative case assignment in Sakha, and discuss how this can be fit into an otherwise standard Chomskian conception of case and agreement. In fact, Sakha is a (39a)-type language, with a close relationship between nominative case and subject agreement and no double agreement with the subject in auxiliary constructions. However, Sakha happens to lack object agreement on active Voice, so (44a) applies to assign accusative. The upshot is that (44) can be the source of some structural case in certain (38a)-type languages, as well as the source of all structural cases in (38b)-type languages.

typology of overt case marking (102 shared values). The preliminary results show that, among (39b)-type languages and languages with no agreement, 16 have ergative-absolutive case and 17 have nominative-accusative case—the expected 50-50 split. In contrast, among (39a)-type languages, 12 have nominative-accusative case and at most 2 have ergative-absolutive case (Hindi and Greenlandic). This is a 6:1 ratio, with a small total number of exceptions (both of which have known complicating issues).²² So that is preliminary support for what may prove to be yet another typological interaction in this domain.

7. Conclusion

In this article, I have considered agreement on predicate adjectives with their subjects, agreement on verbs with their subjects, and agreement on verbs with their objects. I have also considered how the latter two kinds of agreement interact with the case marking of the subject and objects. I hope to have shown how three main assumptions about agreement (the SCOPA in (8), the locality condition in (24), and the eligibility parameter in (29)), together with some assumptions about clause structure that do lots of other work for the generative syntactician, yield quite broad empirical coverage. In particular, I have shown how the generative theory of case and agreement can derive four still-valid universals discovered by functionalist typologists such as Stassen, Croft, Moravcsik, Haspelmath, and Dixon. I have also proposed three new universals of my own, each of which has some support from the exercise in Formal Generative Typology reported in Baker 2008: ch.5. Therefore, I conclude that Universal Grammar does substantively constrain the behavior of case and agreement in natural languages, resulting in observable universals, that generative reasoning can help give (internal) explanations of these universals which reveal connections with other aspects of the grammatical system, and that generative reasoning can lead to the discovery of promising new universals. Any rumors of the death of this paradigm have, I believe, been greatly exaggerated.

References

- Aoun, J. and Y.-h. A. Li (1993). Syntax of scope. Cambridge, Mass., MIT Press.
- Baker, M. (1988). Incorporation: a theory of grammatical function changing. Chicago, University of Chicago Press.
- Baker, M. (2003). Lexical categories: Verbs, Nouns, and Adjectives. Cambridge, Cambridge University Press.

²²In particular, Hindi is a split ergative language, not a true ergative language, and ergative marking has lexical idiosyncrasies: a few predicates have ergative marking on the subject even if the verb is intransitive, and a few fail to put ergative marking on the subject even if it is transitive. This suggests that ergative case in Hindi is not assigned by (44b), but rather is a sort of lexical/inherent case (cf. Woolford 2006). This sort of faux-ergativity can apparently exist in a (39a)-type language. (Of course, this type of case marking must be somewhat exceptional, or one could not expect to see the relationship between ergative case marking and (39b)-type languages that I claim we do see without doing more work to sort out the cases. This is a topic for further research.)

According to Bok-Bennema (1991), Greenlandic may be a language in which T agrees twice. This would have to be taken into account in evaluating both its ergativity and its status with respect to (39).

- Baker, M. (2003). "On the loci of agreement: inversion constructions in Mapudungun." Proceedings of NELS 33.
- Baker, M. (2006). On zero-agreement and polysynthesis. Arguments and agreement. P. Ackema, P. Brandt, M. Schoorlemmer and F. Weerman. Oxford, Oxford University Press: 289-320.
- Baker, M. (2008). The syntax of agreement and concord. Cambridge, Cambridge University Press.
- Baker, M. (in press). Agreement and case. The Cambridge Handbook of Generative Syntax. M. d. Dikken. Cambridge, Cambridge University Press.
- Baker, M. (In press). "When agreement is for number and gender but not person." Natural Language and Linguistic Theory.
- Baker, M. (to appear). "On the relationship of object agreement and accusative case: evidence from Amharic." Linguistic Inquiry.
- Baker, M. (to appear). Types of Crosslinguistic Variation in Case Assignment. Linguistic Variation in the Minimalist Program. J. Brucaat and C. Picallo. New York, Oxford University Press.
- Baker, M. and N. Vinokurova (2009). Forms of predication in Sakha: will the true lexical predicates please stand up. Workshop on noncanonical predication. I. Paul. University of Western Ontario.
- Baker, M. and N. Vinokurova (2010). "Two modalities of case assignment in Sakha." Natural Language and Linguistic Theory **28**(3): 593-642.
- Barss, A. and H. Lasnik (1986). "A note on anaphora and double objects." Linguistic Inquiry **17**: 347-354.
- Bhatt, R. (2005). "Long distance agreement in Hindi-Urdu." Natural Language and Linguistic Theory **23**(4): 757-807.
- Bickel, B. (2007). "Typology in the 21st century: Major current developments." Linguistic Typology **11**(1): 239-251.
- Bobaljik, J. (2008). Where's Phi? Agreement as a post-syntactic operation. Phi Theory: Phi Features Across Interfaces and Modules. D. Adger, D. Harbour and S. Béjar. Oxford, Oxford University Press: 295-328.
- Bobaljik, J. and P. Branigan (2006). Eccentric agreement and multiple case checking. Ergativity: Emerging Issues. A. Johns, D. Massam and J. Ndayiragije. Dordrecht, Springer: 47-77.
- Bobaljik, J. and S. Wurmbrand (2005). "The domain of agreement." Natural Language and Linguistic Theory **23**: 809-865.
- Bruening, B. (2001). Syntax at the edge: cross-clausal phenomena and the syntax of Passamaquoddy. Cambridge, Mass., MIT.

- Bybee, J. (1985). Morphology: a study of the relation between meaning and form. Amsterdam, John Benjamins.
- Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, Mass., MIT Press.
- Chomsky, N. (1981). Lectures on government and binding. Dordrecht, Foris.
- Chomsky, N. (2000). Minimalist inquiries: the framework. Step by Step. R. Martin, D. Michaels and J. Uriagereka. Cambridge, Mass., MIT Press: 89-155.
- Chomsky, N. (2001). Derivation by phase. Ken Hale: A life in language. M. Kenstowicz. Cambridge, Mass., MIT Press: 1-52.
- Cinque, G. (1999). Adverbs and functional heads: a cross-linguistic perspective. New York, Oxford.
- Comrie, B. (1981). Language universals and linguistic typology. Chicago, University of Chicago Press.
- Comrie, B. (2005). Alignment of case marking of full noun phrases. The world atlas of language structures. M. Haspelmath, M. Dryer, D. Gil and B. Comrie. Oxford, Oxford University Press: 398-403.
- Croft, W. (1990). Typology and universals. Cambridge, Cambridge University Press.
- Croft, W. (1991). Syntactic categories and grammatical relations. Chicago, University of Chicago Press.
- Deal, A. R. (2010). "Ergative case and the transitive subject: a view from Nez Perce." Natural Language and Linguistic Theory **28**(1): 73-120.
- Deering, N. and H. Delisle (1976). Mohawk: a teaching grammar. Kahnawake, Quebec, Thunderbird Press.
- Dixon, R. M. W. (1994). Ergativity. Cambridge, Cambridge University Press.
- Dryer, M. (1986). "On primary objects, secondary objects, and antitativity." Language **62**: 808-845.
- Dunn, M., S. Greenhill, et al. (2011). "Evolved structure of language shows lineage-specific trends in word-order universals." Nature **473**: 79-82.
- Evans, N. and S. Levinson (2009). "The myth of language universals: Language diversity and its importance for cognitive science." Behavioral and Brain Sciences **32**(5): 429-492.
- Greenberg, J. (1963). Universals of language. Cambridge, Mass., MIT Press.
- Haspelmath, M. (2005). Ditransitive constructions: the verb 'give'. The World Atlas of Language Structures. M. Haspelmath, M. Dryer, D. Gil and B. Comrie. New York, Oxford University Press: 426-429.
- Kayne, R. (1994). The antisymmetry of syntax. Cambridge, Mass., MIT Press.

- Kramer, R. (2010). Object markers are doubled clitics in Amharic. American Conference on African Linguistics (ACAL). Montreal.
- Kuno, S. (1973). The structure of the Japanese language. Cambridge, Mass, MIT Press.
- Larson, R. (1988). "On the double object construction." Linguistic Inquiry **19**(3): 335-392.
- Lewis, M. P., Ed. (2009). Ethnologue: Languages of the World. Dallas, Texas, SIL International.
- Lorimer, D. L. R. (1935). The Burushaski language 1: Introduction and Grammar. Cambridge, Mass., Harvard University Press.
- Marantz, A. (1991). Case and licensing. The 8th Eastern States Conference on Linguistics, University of Maryland, Baltimore, Ohio State University.
- Marantz, A. (1993). Implications of asymmetries in double object constructions. Aspects of Bantu Grammar **1**. S. Mchombo. Stanford, CA, CSLI: 113-150.
- Merlan, F. (1982). Mangarayi. Amsterdam, North Holland.
- Mohanan, T. (1995). "Woodhood and lexicality: noun incorporation in Hindi." Natural Language and Linguistic Theory **13**(1): 75-134.
- Moravcsik, E. (1974). "Object-verb agreement." Working papers in language universals **15**: 25-140.
- Nevins, A. (2010). Multiple Agree with clitics: person complementarity vs. omnivorous number. University College London.
- Polinsky, M. and E. Potsdam (2001). "Long distance agreement and topic in Tsez." Natural Language and Linguistic Theory **19**(3): 583-646.
- Preminger, O. (2009). "Breaking agreements: distinguishing agreement and clitic doubling by their failures." Linguistic Inquiry **40**: 619-666.
- Reinhart, T. (1976). The syntactic domain of anaphora. Cambridge, Mass., MIT.
- Riedel, K. (2009). The Syntax of Object Marking in Sambia: a comparative Bantu perspective. Leiden, University of Leiden.
- Siewierska, A. (1999). "From anaphoric pronoun to grammatical agreement marker: Why objects don't make it." Folia Linguistica **33**: 225-251.
- Siewierska, A. (2005). Alignment of verbal person marking. The World Atlas of Language Structures. M. Haspelmath, M. Dryer, D. Gil and B. Comrie. New York, Oxford University Press: 406-409.
- Siewierska, A. (2005). Verbal person marking. The world atlas of language structures. M. Haspelmath, M. Dryer, D. Gil and B. Comrie. Oxford, Oxford University Press: 414-417.

- Sigurðsson, H. (1996). "Icelandic finite verb agreement." Working Papers in Scandinavian Syntax **57**: 1-46.
- Smeets, I. (2008). A grammar of Mapuche. Berlin, Mouton de Gruyter.
- Stassen, L. (1997). Intransitive predication. Oxford, Oxford University Press.
- Taraldsen, K. (1995). On agreement and nominative objects in Icelandic. Studies in Comparative Germanic Syntax. H. Haider, S. Olsen and S. Vikner. Dordrecht, Kluwer: 307-327.
- Vinokurova, N. (2005). Lexical categories and argument structure: a study with reference to Sakha, University of Utrecht.
- Willson, S. (1996). "Verb agreement and case marking in Burushaski." Work Papers of the Summer Institute of Linguistics North Dakota **40**: 1-71.
- Woolford, E. (2006). "Lexical case, inherent case, and argument structure." Linguistic Inquiry **37**(1): 111-130.
- Yip, M., J. Maling, et al. (1987). "Case in tiers." Language **63**: 217-250.
- Zaenen, A., J. Maling, et al. (1985). "Case and grammatical functions: the Icelandic passive." Natural Language and Linguistic Theory **3**: 441-483.