Clausal Complementation and Selection

Mark Baker and Ken Safir
Rutgers University

This research initiative would explore how predicates select for clausal complements. Natural languages most typically make available several different clause types which can function as the complements to a verbal (or adjectival) matrix predicate and it is frequently possible, in a given language, to predict which of the several clause types in the language the predicate will select based on the lexical semantics of that predicate. Native speakers appear to acquire the typology of clausal complementation in their language without explicit instruction, which raises familiar poverty of the stimulus questions that imply a great deal of tacit knowledge about the consequences of what a verb means for the syntactic form of the complement it selects. Although work on the semantic and syntactic selection by verbs that take nominal or prepositional complements has been a staple of linguistic work for many years (as, for example, in the work of Levin and Rappaport, 1995 and much work on thematic roles), barely any research has sought to use crosslinguistic contrasts to identify the key factors, or even the key generalizations, that should inform our account of how speakers arrive at the classification of clausal complementation in their native language. Our proposal is to use the Afranaph resources to explore this question. This is attractive in that the African languages permit us to investigate two types of crosslinguistic comparison: comparisons between languages that are broadly different, both in history and typology (e.g. across Afro-Asiatic, Niger-Kordofanian, and Nilo-Saharan), and comparisons among languages that are broadly similar, but differ in the details of clausal selection in ways that reveal smaller cleavages in classification (e.g., across the Bantoid languages).

Insofar as our research must identify which differences in the meaning of a predicate influence its selection, the first order of business is to track the variation in clause types to see what the verbs that select them have in common. This in turn requires that we can identify what the clause types are in a given language, that is, to identify the inventory of possible distinctions in a given language based on the clause types it contains. Some languages have very few clause types while others provide a richer inventory.

Clause types can differ from one another in at least three ways. First, they can differ in the *amount* of structure they contain: In English, for example, manner of speaking verbs take a full clause CP, raising verbs take a bare TP, causative complements appear to take a bare vP/VP, and so on. Second, complements can differ in their *major categorial properties*. Ross (1973) suggests that embedded clauses in a given language might be nominalized to various degrees, an insight that might be understood in current terms as differences in which functional head in the structure of the clause is "nominalized" (Borsley and Kornfilt, 2000, etc.). Third, clauses can be distinguished on the basis of the *types of non-categorial features* that characterize the clause (also often manifest on the clause head): a TP might be finite or nonfinite, a CP might be [+/-interrogative], the clause might be subjunctive or indicative, and so on.

Putting together these different factors, a single language could have quite a few different kinds of syntactically distinguishable clause types. For example, English has at least the following: bare VP complements, infinitival TP complements, control infinitives, *for-to* infinitives, tensed finite CPs, interrogative CPs, interrogative infinitives, Poss-*Ing* gerunds, ACC-*ing* gerunds, PRO-*ing* gerunds, and derived nominals. Furthermore, some languages have possibilities that may not be attested in English: for example, many Romance languages have a rich and robust category of subjunctive clause that is limited and moribund in English. The Turkic language Sakha has a type of participial clause that is not quite like anything in English. Conversely, some languages may have fewer clause types than English does: for example, Mohawk apparently has no nonfinite clauses of any kind (Baker 1996), whereas all embedded clauses are nominalized with structures approximately like that of English gerunds in Mapudungun and Quechua. While it is known that such differences among languages exist, the precise extent of the variation has never been determined, nor is it known whether the logical possibilities cluster into coherent systems in principled ways.

At the same time, most languages have a wide variety of verbs that select for clausal complements of one kind or another. And in many cases what sort of clausal complement a verb appears with is crucial. For example, the verb *want* in English can take an infinitival complement, with or without an overt subject, but it cannot take a finite complement or any kind of gerund:

1a) Chris wants [to arrive early]
 b) Chris wants [John to arrive early]
 c)*Chris wants that John arrives early.
 d)*Chris wants that John arrive early.   (Subjunctive)
 e)*Chris wants arriving early.
 f)*Chris wants John('s) arriving early.

Other verbs in English would take a different range of clausal complements, in a very complex pattern. Nevertheless, these patterns are acquired without instruction by every native speaker of English, even though the space of logical possibilities is so large (many different matrix verbs crossed with many different types of complements). Since adult native speakers arrive at fairly stable judgments of what is and is not possible, we can suppose that they have gotten some substantive guidance from Universal Grammar.

While we can explore these questions in part by simply considering the judgments of English speakers, we can't find out in this way how selection is achieved in languages with different inventories of clause types, nor can we know what patterns are persistent across languages. For example, we observe that verbs translatable as 'want' take infinitival complements but not tensed indicative complements in language after language, whereas verbs translatable as 'say' take tensed indicative complements but not infinitives in language after language. Subtle distinctions in English may play out differently in other languages. For example, *prefer* in English has a meaning not vastly different from that of *want*, but the complementation properties of the two English verbs differ significantly: *prefer* can select subjunctive/indicative CP complements and gerunds as well as infinitives. In another language it might not be obvious which is which (or if there is a difference). For example, in the Nigerian language Edo, the verb translated as 'want' selects a subjunctive CP complement and not (the equivalent of) an infinitival

complement. The principles by which it is determined which verbs select for which types of complements remain largely unknown for any language.

Thus it stands to reason that what kind of complement a given verb will take in language X will depend to a significant degree on the range of complement types that are available structurally in that language. Some of these effects should be quite straightforward. A verb meaning 'say' selects a gerund as its complement rather than a finite CP in Mapudungun—at first glance a surprising difference. But this becomes unsurprising once one takes into account the general fact Mapudungun has no CP in its grammar (indeed no lexical items of category C). Therefore it "makes do" with what it does have, namely gerund-like structures. Similarly, in Mohawk and the Balkan languages verbs meaning 'want' take subjunctive CPs not infinitives. This difference is less surprising when we take into account that (a) these languages do have a robust category of subjunctive clause (unlike English), and (b) they do not have a category of infinitive. We can say then that either 'want' prefers a subjunctive, but can settle for an infinitive if there is no subjunctive, or that 'want' prefers an infinitive, but can settle for a subjunctive if there is no infinitive (or both, with different senses of 'want'—cf. Romance languages). Interesting theoretical questions now arise immediately: What constitutes "settling for", how systematic is it, and how can we formalize a system that predicts the level of systematicity that is found? And we imagine that there could be much more subtle effects of this kind as well: for example, what verbs a gerund can be used with could depend on whether the language also has available participial clauses of the Sakha sort or not. In the most extreme case, we may expect that if a language has only one sort of clausal complement (Mapudungun?), every verb that takes a clausal complement will select that type of clause. Many possible comparisons are probably amongst those that will have to be considered, such as the possibility that intricate complementation patterns in a language that allows five different kinds of clausal complement could differ from a language that permits a different set of five clausal complements.

Our research hypothesis, then, is that the general patterns of what verbs take what complements in a given language will depend on a sophisticated understanding of both what the verbs actually mean and on what range of clausal complements happens to be available in that language. Crosslinguistic differences in selection that can seem quite idiosyncratic when compared point by point in isolation might become principled when studied in the context of the overall system in this way.

The research we imagine would thus have the following three components:

- Establishing a rich notion of the lexical semantics of verbs and other predicates that select clausal complements, to insure crosslinguistic comparability. This would be done by inquiring into the range of inferences each verb supports. For example, *prefer* differs from *want* in English in that it denotes a contrastive judgment between imagined alternatives, whereas *want* need not presuppose that any other alternative is under active consideration.
- Establishing a rich understanding of the structure of the different kind of clauses available in particular languages. This would be done by collecting the usual sorts of syntactic judgments over a wide range of diagnostic phenomena: does the clause appear in all grammatical functions, can it be extraposed, is it an island for

extraction, can it be case marked, etc. (A classical model is Ross 1973; a more recent one is Baker 2010 on Sakha.)

- Seeing systematically which verbs can take which complements, so as to deduce the patterns as a function of both verb meaning and the range of possible complements made available in the language.

Although the interest of these questions is in no way limited to African languages, and we have illustrated them with examples from other parts of the world as often as not, the dual track of exploring both high contrast and low contrast linguistic variation is achievable within the indigenous languages of the African continent, as mentioned above. We do know that there are some very interesting things going on in sentential complementation in African languages—both things that look quite familiar and are possible instances of UG at work, and things that look quite unfamiliar and constitute interesting puzzles for comparative syntactic theory. For example:

- Embedded clauses (e.g. gerunds) have strange word order properties in some Nigerian languages but not others, showing OV order in languages that are otherwise VO (Lokaa). Does this have any effect on the overall structure of the clauses and what predicates can selects them?
- Infinitives are somewhat nominalized in Bantu languages. Are these more like infinitives or gerunds in other languages? What implications does this have for selection? In some languages, infinitives are part of a noun class that contains simple nouns as well (*eri*- Kinande) whereas in others they are part of a unique noun class that contains no simple nouns (*ku*- in Kirundi and Iklalanga; see data below). Does this morphological difference have any further implications for structure and for selection?
- Complementizers in some Bantu languages bear agreement with the matrix subject (see Kinande below; also Lubukusu), and others don't (see Kirundi). What verbs select that special complementizer? Is having an agreeing complementizer only an option for finite indicative clauses, or could it happen with infinitival or subjunctive clauses as well?
- African languages like Yoruba have a distinction between complements that can contain logophoric pronouns and complements that cannot. This seems to depend on whether there is a CP projection or not, and perhaps also on the particular C that is used. What is the structure of the complement, and what can select it and why?
- Some Bantu languages use the subjunctive final vowel –e in the CP complement of a verb like 'want', whereas others seem to use the ordinary indicative final vowel –a (see Ikalanga versus Kirundi and Kinande below).
- Amharic has the possibility of Shlenker-type indexical shift in the clausal complements of some verbs but not others. There are also interesting structural questions about complementizers and gerund prefixes in Amharic which arise from the apparent mixed headedness in the language.
- Machobane 1988 found fascinating and unexpected effects of adding causative and applicative morphology to verbs that take clausal complements in Sesotho (Bantu). Such changes affect complementation, control, and so on. Very little

replication or follow up has been done on this issue in other languages, whether related to Sesotho or not.

And so on.

In fact, the Afranaph database already has some information about complementation in it, because of the relationship between complementation and binding (e.g. long distance anaphors being possible in infinites and subjunctives but not tensed clauses, Tensed Sentence Condition effects, logophoricity). Simple tests performed for other reasons reveal, for example, that the verb translated as 'want' in KiRundi, IKalanga, and KiNande takes both clauses introduced by a complementizer and noun class 14 (infinitive?) complements, whereas Yoruba, which does not assign noun classes to clauses, appears to make a similar distinction for bare verbs that may or may not be analyzed as infinitives, but must have whatever it takes to fill the lexical space that c14 does in Narrow Bantu (see the "Data from Afranaph" below). The "whatever it takes" property is then what we are aiming to characterize and then, hopefully, explain.

So the ice has already been broken in this general domain. But even more importantly, this is a project that would require the distinctive Afranaph methodology. It could not be pursued by looking at a single language, since this would not allow one to distinguish what is principled but variable from what is invariant or what is truly idiosyncratic. It also clearly needs the participation of linguistically trained native speaker linguists. Only such collaborators would be able to draw the crucial inferences to establish what a given verb actually means, and only they would be able to reliably carry out the syntactic tests needed to establish in detail the structure of the various clause types. They would also be sensitive to the possibility that there might be distinctions that are not revealed by the superficial morphology, such as the different types of gerunds in English.

Since little is known about these issues in African languages, one is guaranteed to learn something, and since these issues are equally interesting and relevant in languages from other areas of the world, it might be a natural place to try extending the initiative beyond Africa.

Data from Afranaph. Those interested in details or commentary on these sentences should use the ID numbers to search for the examples and click on the 'details' button for each example. Caveat: Not all data has complete tone information.

Id: 2508   Strategy: RCM , Language: Kirundi

	abo	bakobwa	ba-shaka	ku-kund-an-a
(ok)	those	Girls	SM.3PL-want	SM.3PL-like-RECP-IPFV

Those girls want to like each other.

Id: 2509   Strategy: RCM , Language: Kirundi

	Solo	a-ashaka	ko	Abo	bakobwa	ba-kund-an-a
(ok)	S	SM.3SG-want	that	those	girls	SM.3PL-like-RECP-IPFV

Solo wants that those girls like each other

Id: 114   Strategy: RFM , Language: Ikalanga

(ok)
Soli u-no-shak-a   ku-zwi-khuz-a.
Sol   SM.c1a-PRS-want-FV to-RFM-praise-FV
Sol wants to praise himself.

Id: 115   Strategy: RFM , Language: Ikalanga

(ok)
Sol u-no-shak-a   kuti Alisi a-zwi-khuz-e.
Sol SM.c1a-PRS-want-FV that Alice SM.c1a-RFM-praise-FV
Sol expects Alice to praise herself.

Id: 160   Strategy: RCM , Language: KiNande

(ok)
aba-síká Bá-sóndire   eri-pip-án-a
c2-girls   SM.c2-want INF-praise-RCM-FV
The girls want to praise each other.

Id: 948   Strategy: OM , Language: KiNande

(ok)
Nyísóndire   indí   ínabugá   bolóbólo
SM.1SG.c1-want SM.1SG.c1-say that-SM.c1-speaks slowly
I want that he speak slowly.

Id: 1223 Strategy: ara-ASSOC-pronoun , Language: Yoruba

(ok)
Ṣọlá Fẹ́   yin   ara   rẹ̀
Sola Want praise body his
Sola wants to praise himself.

Id: 1224 Strategy: ara-ASSOC-pronoun , Language: Yoruba

(ok)
Ṣọlá Fẹ́   kí   Álíìsì yin   ara   rẹ̀
Sola Want COMP Aliisi praise body his
Sola expects Alice to praise herself.