On the cross-linguistic distribution of same-subject and different-subject 'want' complements: Economic vs. iconic motivation

Martin Haspelmath
Max-Planck-Institut für evolutionäre Anthropologie, Leipzig

1. Subject omission and formal simplicity in 'want' complements

In this paper, I argue that the asymmetrical formal patterns that we find in same-subject (SS) and different-subject (DS) 'want' complement clauses across languages can be explained by frequency biases (economic motivation in Haiman’s 1983 sense). ‘Want’ complements overwhelmingly show referential identity between the wanter argument and the subject (the A or S argument) of the complement clause. Sentences like (1a) are much more frequent than sentences like (1b), apparently in all languages. (Here the symbol Ø is meant to show that the notional subject of the complement clause is not expressed overtly.)

(1) a. Kim₁ wants [Ø₁ to go home].
   b. Kim₁ wants [him₂ to go home].

I claim that two aspects of same-subject complement clauses are economically motivated: (i) The fact that the notional complement subject is unexpressed in many languages (as illustrated in (1a), where the notional subject of 'to go home' cannot be expressed overtly in its clause), and (ii) the fact that different-subject complement clauses often show a formally more complex pattern, as illustrated by German:

(2) German
   a. Kim₁ will [Ø₁ nach Hause gehen].
      Kim wants to home go
      ‘Kim wants to go home.’

   b. Kim₁ will, [dass er₂ nach Hause geht].
      Kim wants that he to home goes
      ‘Kim wants him to go home.’

In the different-subject pattern (2b), there is an additional complementizer (dass), and the verb form is finite (geht ‘goes’), contrasting with the infinitival form in (2a), so the DS pattern is more complex in two different ways.

* An oral version of this paper was presented at the International Conference on Cognitive Linguistics, Stockholm, July 1999, and similar talks were given at the University of Erfurt in May 1998, at the DGfS conference in Marburg in February 2000, and at the typology symposium at the University of Frankfurt am Main in November 2002. I thank all audiences for useful comments.
I refer to these frequently encountered properties of same-subject complements as **subject omission** and **formal simplicity**. It should be noted that neither subject omission nor formal simplicity are universal properties. While both English and German exhibit obligatory subject omission, there are quite a few languages that must express the subject overtly in same-subject complements of 'want'. Two such languages are Modern Greek and Standard Arabic, as illustrated in (3)-(4). Here the subject in the SS complement in (a) is expressed in the same way as the subject in the DS complement in (b).

(3) Modern Greek

a. (SS) θέλω να δουλέω.
   want-1SG [SBJV work-1SG]
   'I want to work.' (Lit. 'I want (that) I work.]

b. (DS) θέλω να δουλέω-
   work-2SG
   'I want you to work.' (Lit. 'I want (that) you work.]

(4) Standard Arabic

a. (SS) ʔuriid-u ʔan ʔa-takallam-a.
   1SG-want-IND [that 1SG-talk-SBJV]
   'I want to talk.' (Lit. 'I want (that) I talk.]

b. (DS) ʔuriid-u ʔan ta-takallam-a.
   1SG-want-IND [that 2SG-talk-SBJV]
   'I want you to talk.' (Lit. 'I want (that) you talk.]

And while German exhibits formal simplicity (of the same-subject pattern compared to the different-subject pattern), English does not: Apart from the obligatory omission of the subject, (1a) is not simpler than (1b).

However, in a broader context the subject omission and formal simplicity can be seen as universal properties: Compared to other verbs such as ‘believe’, ‘want’ apparently always shows a greater or equal tendency to exhibit subject omission and formal simplicity. We can thus formulate the following implicational universals:

(5) a. If a language has subject omission in (same-subject) complements of ‘believe’, it also has subject omission in complements of ‘want’.

b. If a language shows formal simplicity in same-subject complements of ‘believe’, it also shows formal simplicity in complements of ‘want’.

Even though 'believe' tends to be coded differently from 'want', languages may also treat both of them in the same way. An example of a same-subject ‘believe’ construction that shows both subject omission and formal simplicity is shown in (6a) from French, contrasting with the different-subject pattern in (6b). The verb croire thus allows the same constructions in which vouloir ‘want’ occurs (cf. 7a-b).  


---

1 However, the complement clause with que allows both a same-subject and a different-subject reading in (6b), in contrast to (7b). Thus, croire has all the possibilities of vouloir and some additional ones. (But note
(6)a. *Kim croit être seul.
   Kim believes to.be alone
   ‘Kim believes that he is alone.’
b. *Kim croit qu’il est seul.
   Kim believes that he is alone
   ‘Kim believes that he is alone.’

(7)a. *Kim veut être seul.
    Kim wants to.be alone
    ‘Kim wants to be alone.’
b. *Kim veut qu’il soit seul.
    Kim wants that he be alone
    ‘Kim wants him to be alone.’

So the contrast between ‘believe’ and ‘want’ is not universal, but when there is a difference between them, then ‘want’ shows subject omission and formal simplicity, while ‘believe’ does not. 2

2. Redundancy and economic motivation

The first formal characteristic of same-subject complements, that the subject is often left unexpressed, is perhaps not very surprising, because its referent can be readily inferred from the context. However, things are not that simple, for three reasons.

First, the referential identity in (1a) is not simply pragmatically inferred, but is rigidly prescribed by the grammar. In other words, the complement subject is obligatorily controlled. It is not just an available or preferred reading, but it is the only possible reading. Such obligatory control is found in many languages, not just in English.

Second, the inference is not watertight, because complements of ‘want’ do not have to have a subject that is identical to the wanter. In this, they contrast with ability verbs, for example, where a different-subject pattern is simply nonsensical:

(8)  a. *Sasha1 is able [Ø1 to lift the suitcase].
    b. *Sasha1 is able [for him2 to lift the suitcase].

Here it is quite clear that expressing the subject of the complement clause overtly in (8a) would be redundant, because there is no possible contrast. The reference of the complement subject can be predicted with 100% certainty (so I call this deterministic

that the complement verb is in the subjunctive in (7b) but in the indicative in (6b), so (6) and (7) are not completely parallel.)

2 Unfortunately, the evidence for the claims in (5a-b) is purely impressionistic at the moment, derived from my casual observations of many descriptions of languages. But in order to formulate the general tendencies in the coding of ‘want’ complements as testable universals, they need to be contrasted with another type of complement clause. More evidence needs to be provided for these claims in future work.
redundancy). Still, some languages opt for full explicitness and express it redundantly in such constructions as well, e.g. Modern Greek:

(9) I Kopéla bor-i [na anij-i ti fiáli].

The girl can open the bottle.

‘The girl can open the bottle.’ (Lit. ‘...is able that she opens the bottle’)

So the fact that English (like most other languages) does not repeat the subject in (8a) is evidently an exploitation of the redundancy that derives from the special semantics of ability. But an analogous explanation is not immediately evident in the case of ‘want’, because the counterpart of (8b) is perfectly possible (see 1b).

Third, it is by no means always the case that referential identity with an element in the immediate context leads to omissibility of the referential expression. Consider reflexive situations such as (10).

(10) María saw herself in the mirror.

Here the object argument is referentially identical with the subject argument, so one might think that it can be inferred easily, and that there would be a tendency to omit it. However, the opposite is the case: While English does not allow object omission at all (some kind of pronominal object has to be used), many other languages allow objects to be implicit, but in such cases the object normally gets a disjoint reading (non-identity with the subject), not a reflexive reading.

Thus, it does not go without saying that subject omission in sentences like (1a) gets a same-subject interpretation. This must be a convention of the grammar that could in principle be the opposite and that calls for an explanation. There are thus actually three aspects of the SS-DS contrast in (1a–b) that need to be explained: the omission of the notional subject in (1a), the obligatory coreference with the wanter in (1a), and the greater formal complexity of the DS pattern in (1b). I claim that all these aspects can be explained in terms of the frequency bias (economic motivation).

An alternative explanation that one finds in the literature is in terms of iconic motivation (Haiman 1983, Givón 1990, Cristofaro 2003). The idea is that the “functional integration” of SS complements is iconically reflected in “formal integration”. According to Givón (1990: 560), the “the degree of finiteness is an iconic expression of the degree of integration of the main and complement events”. Givón writes:

"Given a hierarchy of degree of finiteness (or its converse, degree of nominality) of verb forms found in a language, the more integrated the two events are,
(i) the more noun-like is the complement verb likely to be, and
(ii) the less finite verbal morphology – such as tense-aspect-modality and pronominal agreement – is the verb likely to display." (Givón 1990: 561)

In much the same vein, Cristofaro (2003: 252) writes that "lack of TAM and person agreement distinctions (as well as lack of overtly expressed arguments) leads to syntactic integration between clauses, and iconically reflects semantic integration between states of affairs."
The explanation in terms of frequency biases is a universalist usage-based explanation, i.e., it derives universal tendencies of grammatical form from universal frequency biases. We thus need to consider what the cross-linguistic formal patterns are, and what the cross-linguistic usage frequencies are. In the next section, we start with the frequency biases, before looking at the cross-linguistic formal patterns in §4.

3. The frequency bias for same-subject constructions

While there is no deterministic redundancy in the case of same-subject ‘want’ constructions, I argue that subject omission is due to redundancy exploitation here, too. The redundancy is probabilistic rather than deterministic, but this is sufficient to lead languages to choose subject-omitting constructions in many cases (even the majority, as we will see in §4). Thus, the argument is that many languages (including English) have constructions like *I want Ø to go home* because ‘want’ complements overwhelmingly show subject identity, so that the subject can be easily predicted by the hearer and does not have to be expressed overtly by the speaker. Such cases of zero expression exploiting probabilistic redundancy are very widespread in language structure (cf. Haspelmath 2008a), and no ad hoc assumption need to be made here.

But of course, exploiting the redundancy that derives from the frequency bias is not necessary. Languages may instead opt for explicit coding (cf. (3)–(4) above), just as they may opt for explicit coding even in the case of deterministic redundancy (cf. example (9) above).

The strong usage preference for same-subject ‘want’ constructions can easily be seen in text counts. Table 1 shows data from English, and Table 2 shows data from written Italian. The trend is so overwhelming that it would be pointless to look at larger corpora.

Table 1. Frequency of same-subject and different-subject complements of *want* in English

<table>
<thead>
<tr>
<th></th>
<th>540</th>
<th>89%</th>
</tr>
</thead>
<tbody>
<tr>
<td>same-subject</td>
<td>76</td>
<td>11%</td>
</tr>
<tr>
<td>different-subject</td>
<td>616</td>
<td>100%</td>
</tr>
</tbody>
</table>

Total: cases of *want* with clausal complement

Source: ICE-GB (International Corpus of English), Schmidtke-Bode (2011+:§3.2)

Table 2. Frequency of same-subject and different-subject complements of *volere* ‘want’ in written Italian

<table>
<thead>
<tr>
<th></th>
<th>444</th>
<th>87%</th>
</tr>
</thead>
<tbody>
<tr>
<td>same-subject</td>
<td>65</td>
<td>13%</td>
</tr>
<tr>
<td>different-subject</td>
<td>509</td>
<td>100%</td>
</tr>
</tbody>
</table>

Total: cases of *volere* ‘want’ with clausal complement

(Letteratura Italiana Zanichelli (LIZ) on CD-ROM)
Now one might suspect that the frequency asymmetry that we see in English and Italian is not universal, and is in fact due to the formal asymmetry, so that the direction of causality is exactly opposite from what I have claimed. In other words, it could be that same-subject complements are so much more frequent than different-subject complements because they are shorter and involve less coding effort. Thus, we should also look at languages in which there is no formal asymmetry, and one such language is Modern Greek. Some data from written Modern Greek are given in Table 3 (as I had no electronic corpus available, the absolute figures are very low).

Table 3. Frequency of same-subject and different-subject complements of théló ‘want’ in written Modern Greek

<table>
<thead>
<tr>
<th>Same-subject complements</th>
<th>Different-subject complements</th>
<th>Total: cases of théló ‘want’ with clausal complement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>38</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>88%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

Source: Kóstas Tzamális, Stin Athína tu Periklí, Athen: Estia/Kollaru, pp. 22-122.

Despite the limited amount of data, I regard the corpus counts as fairly good evidence that the frequency bias is strong and cross-linguistically systematic, regardless of the formal encoding of same-subject and different-subject complements of ‘want’. And the frequency difference is of course not surprising: Humans are egocentric, and our own actions are much more important to us than other people’s actions, so it is expected that we talk more about our desires concerning our own actions than about our desires concerning other people’s actions.

Let us now look in more detail at the cross-linguistic encoding patterns.

4. ‘Want’-constructions in 80 languages world-wide

My data are from a world-wide convenience sample of 80 languages, for which I have collected data on same-subject and different-subject ‘want’ constructions from reference grammars and other kinds of grammatical descriptions. These languages are listed in Table 4, with some genealogical information, a bibliographical reference, and the type of relationship between the SS ‘want’ pattern and the DS ‘want’ pattern. Seven different types are distinguished:

Type 1: no coding asymmetry (as in Modern Greek and Standard Arabic, see (3)-(4))
Type 2: simple subject omission (see §5)
Type 3: simple complementizer omission (see §6)
Type 4: complementizer omission and different verb form (see §7)
Type 5: shorter ‘want’ verb (see §8.1)
Type 6: ‘want’ verb expressed as desiderative marker (see §8.2)
Type 7: different-subject construction does not exist (see §8.3)
<table>
<thead>
<tr>
<th>language</th>
<th>family</th>
<th>subfamily</th>
<th>type</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ju</td>
<td>’hoan</td>
<td>(“Khoisan”)</td>
<td>Ju</td>
<td>type 3</td>
</tr>
<tr>
<td>Mende</td>
<td>Niger-Congo</td>
<td>Mande</td>
<td>type 1</td>
<td>Innes 1971:122</td>
</tr>
<tr>
<td>Dagbani</td>
<td>Niger-Congo</td>
<td>Gur</td>
<td>type 1</td>
<td>Olawsky 1999:25</td>
</tr>
<tr>
<td>Yoruba</td>
<td>Niger-Congo</td>
<td>Defoid</td>
<td>type 3</td>
<td>Rowlands 1969:66, 74</td>
</tr>
<tr>
<td>Swahili</td>
<td>Niger-Congo</td>
<td>Bantu</td>
<td>type 4</td>
<td>NT</td>
</tr>
<tr>
<td>Lango</td>
<td>Nilo-Saharan</td>
<td>Nilotic</td>
<td>type 4</td>
<td>Noonan 1992:221, 223</td>
</tr>
<tr>
<td>Krongo</td>
<td>Nilo-Saharan</td>
<td>Kadugli</td>
<td>type 4</td>
<td>Reh 1985:335</td>
</tr>
<tr>
<td>Koyraboro Senni</td>
<td>Nilo-Saharan</td>
<td>Songhay</td>
<td>type 1</td>
<td>Heath 1999:326</td>
</tr>
<tr>
<td>Standard Arabic</td>
<td>Afro-Asiatic</td>
<td>Semitic</td>
<td>type 1</td>
<td>own knowledge</td>
</tr>
<tr>
<td>Maltese</td>
<td>Afro-Asiatic</td>
<td>Semitic</td>
<td>type 3</td>
<td>Sandro Caruana, p.c.</td>
</tr>
<tr>
<td>Coptic</td>
<td>Afro-Asiatic</td>
<td>Egyptian</td>
<td>type 4</td>
<td>Lambdin 1983:54, 56, 74</td>
</tr>
<tr>
<td>Somali</td>
<td>Afro-Asiatic</td>
<td>Cushitic</td>
<td>type 1</td>
<td>Berchem 1991:253</td>
</tr>
<tr>
<td>Iraqw</td>
<td>Afro-Asiatic</td>
<td>Cushitic</td>
<td>type 2</td>
<td>Mous 1993:109, 266, 291</td>
</tr>
<tr>
<td>Mupun</td>
<td>Afro-Asiatic</td>
<td>Chadic</td>
<td>type 3</td>
<td>Frajzyngier 1993:471-2</td>
</tr>
<tr>
<td>Hausa</td>
<td>Afro-Asiatic</td>
<td>Chadic</td>
<td>type 1</td>
<td>Kraft &amp; Kirk-Greene 1973:167</td>
</tr>
<tr>
<td>Georgian</td>
<td>Kartvelian</td>
<td></td>
<td>type 4</td>
<td>Vamling 1989</td>
</tr>
<tr>
<td>Lezgian</td>
<td>Nakh-Daghestani</td>
<td>Lezgic</td>
<td>type 4</td>
<td>Haspelmath 1993:369</td>
</tr>
<tr>
<td>Godoberi</td>
<td>Nakh-Daghestani</td>
<td>Andic</td>
<td>type 2</td>
<td>Hasepmlath 1996:188</td>
</tr>
<tr>
<td>Basque</td>
<td>–</td>
<td>–</td>
<td>type 4</td>
<td>Andolan Eguzkitzia, p.c.</td>
</tr>
<tr>
<td>German</td>
<td>Indo-European</td>
<td>Germanic</td>
<td>type 4</td>
<td>own knowledge</td>
</tr>
<tr>
<td>English</td>
<td>Indo-European</td>
<td>Germanic</td>
<td>type 1, 3</td>
<td>own knowledge</td>
</tr>
<tr>
<td>Modern Greek</td>
<td>Indo-European</td>
<td>Greek</td>
<td>type 1</td>
<td>own knowledge</td>
</tr>
<tr>
<td>Kashmiri</td>
<td>Indo-European</td>
<td>Indic</td>
<td>type 4</td>
<td>Wali &amp; Koul 1997:46, 50</td>
</tr>
<tr>
<td>Hindi</td>
<td>Indo-European</td>
<td>Indic</td>
<td>type 4</td>
<td>NT</td>
</tr>
<tr>
<td>Finnish</td>
<td>Uralic</td>
<td>Finnic</td>
<td>type 4</td>
<td>Karlsson 1999:182</td>
</tr>
<tr>
<td>Khanty</td>
<td>Uralic</td>
<td>Ugric</td>
<td>type 2</td>
<td>Nikolaeva 1999:25, 46</td>
</tr>
<tr>
<td>Turkish</td>
<td>Turkic</td>
<td></td>
<td>type 4</td>
<td>own knowledge</td>
</tr>
<tr>
<td>Evenki</td>
<td>Tungusic</td>
<td></td>
<td>type 6</td>
<td>Nedjalkov 1997:27-8</td>
</tr>
<tr>
<td>Mandarin Chinese</td>
<td>Sino-Tibetian</td>
<td>Sinitic</td>
<td>type 2</td>
<td>Bingfu Lu, p.c.</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>Hmong-Mien</td>
<td></td>
<td>type 3</td>
<td>Harriehausen 1990:216, 220, 230</td>
</tr>
<tr>
<td>Japanese</td>
<td>–</td>
<td>–</td>
<td>type 6</td>
<td>Kaoru Horie, p.c.</td>
</tr>
<tr>
<td>Korean</td>
<td>–</td>
<td>–</td>
<td>type 5</td>
<td>Chang 1996:65, 126</td>
</tr>
<tr>
<td>Tagalog</td>
<td>Austronesian</td>
<td>Philippine</td>
<td>type 2</td>
<td>Schachter &amp; Otanes 1972:268, NT</td>
</tr>
<tr>
<td>Acehnese</td>
<td>Austronesian</td>
<td>Sundic</td>
<td>type 7</td>
<td>Durie et al. 1994:177-178</td>
</tr>
<tr>
<td>Indonesian</td>
<td>Austronesian</td>
<td>Sundic</td>
<td>type 4</td>
<td>Sneddon 1996:271, 275, 296</td>
</tr>
<tr>
<td>Madurese</td>
<td>Austronesian</td>
<td>Sundic</td>
<td>type 4</td>
<td>Davies 1999:46</td>
</tr>
<tr>
<td>Language</td>
<td>Type</td>
<td>Subtype</td>
<td>Type number</td>
<td>Reference(s)</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>---------------</td>
<td>-------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Kiribati</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 4</td>
<td>Groves et al. 1985:53, 56, 153</td>
</tr>
<tr>
<td>Taba</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 1</td>
<td>Bowden 2001: 391–92</td>
</tr>
<tr>
<td>Erromangan</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 6</td>
<td>Crowley 1998:134</td>
</tr>
<tr>
<td>Drehu</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 5</td>
<td>Moyse–Faurie 1983:181–2</td>
</tr>
<tr>
<td>Boumaa Fijian</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 5</td>
<td>Dixon 1988:39, 91, 279, 286</td>
</tr>
<tr>
<td>Samoan</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 5</td>
<td>Mosel 1994:337–338</td>
</tr>
<tr>
<td>Maori</td>
<td>Austronesian</td>
<td>Oceanic</td>
<td>type 2</td>
<td>Bauer 1993:42</td>
</tr>
<tr>
<td>Kalam</td>
<td>Trans-New Guinea</td>
<td>Madang</td>
<td>type 7</td>
<td>Pawley 1994:400</td>
</tr>
<tr>
<td>Kombai</td>
<td>Trans-New Guinea</td>
<td>Awju-Dumut</td>
<td>type 7</td>
<td>de Vries 1993</td>
</tr>
<tr>
<td>Abun</td>
<td>West Papuan</td>
<td>North-Central Bird’s Head</td>
<td>type 1</td>
<td>Berry and Berry 1999:167, 176</td>
</tr>
<tr>
<td>Warembori</td>
<td>Lower Mamberano</td>
<td></td>
<td>type 1</td>
<td>Donohue 1999:48–49</td>
</tr>
<tr>
<td>Mparntwe</td>
<td>Pama-Nyungan</td>
<td></td>
<td>type 2</td>
<td>Wilkins 1989:451–1</td>
</tr>
<tr>
<td>Arrernte</td>
<td>Pama-Nyungan</td>
<td></td>
<td>type 4</td>
<td>Dench 1995:256</td>
</tr>
<tr>
<td>Martuthunira</td>
<td>Pama-Nyungan</td>
<td></td>
<td>type 2</td>
<td>Blevins 2001:132–3</td>
</tr>
<tr>
<td>Nhanda</td>
<td>Pama-Nyungan</td>
<td></td>
<td>type 2</td>
<td></td>
</tr>
<tr>
<td>Maranungku</td>
<td>Daly</td>
<td>Western Daly</td>
<td>type 6</td>
<td>Tryon 1970:56</td>
</tr>
<tr>
<td>Kayardild</td>
<td>Tangkic</td>
<td></td>
<td>type 7</td>
<td>Evans 1995:259</td>
</tr>
<tr>
<td>Bininj Gun-wok</td>
<td>Gunwinyguan</td>
<td></td>
<td>type 1</td>
<td>Evans 2003:640–1</td>
</tr>
<tr>
<td>Labrador Inuttut</td>
<td>Eskimo-Aleut</td>
<td></td>
<td>type 5</td>
<td>Smith 1982</td>
</tr>
<tr>
<td>Dogrib</td>
<td>Athapaskan</td>
<td></td>
<td>type 1</td>
<td>Saxon 1984</td>
</tr>
<tr>
<td>Halkomelem</td>
<td>Salishan</td>
<td></td>
<td>type 1</td>
<td>Gerdts 1988</td>
</tr>
<tr>
<td>Mohawk</td>
<td>Iroquoian</td>
<td></td>
<td>type 1</td>
<td>Baker 1996</td>
</tr>
<tr>
<td>Lakota</td>
<td>Siouan</td>
<td></td>
<td>type 2</td>
<td>Van Valin 1977</td>
</tr>
<tr>
<td>Diegueño</td>
<td>Yuman</td>
<td></td>
<td>type 1</td>
<td>Gorbet 1998:11</td>
</tr>
<tr>
<td>Maricopa</td>
<td>Yuman</td>
<td></td>
<td>type 1</td>
<td>Gordon 1986:248–50</td>
</tr>
<tr>
<td>Hopi</td>
<td>Uto-Aztecan</td>
<td></td>
<td>type 3</td>
<td>Kalectaca 1978</td>
</tr>
<tr>
<td>Tümpisa</td>
<td>Uto-Aztecan</td>
<td></td>
<td>type 7</td>
<td>Dayley 1989</td>
</tr>
<tr>
<td>Ute</td>
<td>Uto-Aztecan</td>
<td></td>
<td>type 4</td>
<td>Chamoreau 2000: 92, 121, 163</td>
</tr>
<tr>
<td>Purépecha</td>
<td></td>
<td></td>
<td>type 4</td>
<td></td>
</tr>
<tr>
<td>Tzutujil</td>
<td>Mayan</td>
<td></td>
<td>type 3</td>
<td>Dayley 1985:391–3</td>
</tr>
<tr>
<td>Mam</td>
<td>Mayan</td>
<td></td>
<td>type 4</td>
<td>England 1983:302</td>
</tr>
<tr>
<td>Chalcatongo Mixtec</td>
<td>Oto-Manguean</td>
<td></td>
<td>type 3</td>
<td>Macaulay 1996</td>
</tr>
<tr>
<td>Guarani</td>
<td>Tupi</td>
<td></td>
<td>type 6</td>
<td>Gregores and Suarez 1967:128, 177, 179</td>
</tr>
<tr>
<td>Trumai</td>
<td></td>
<td></td>
<td>type 7</td>
<td>Girardello 1999:146, 175</td>
</tr>
</tbody>
</table>
5. Simple subject omission (type 2)

The most common type of frequency effect that we find is the omission of the subject (the S or A argument) in same-subject complements. This happens in the great majority of the languages of the sample. (Only 15 of the 80 languages show the Modern Greek pattern, with no subject omission.\(^3\))

There are 14 languages in the sample in which this is the only economy effect (type 2). One such language is English, as we saw in the very first example (1a–b). A few further examples of languages in which the different-subject construction is identical to the same-subject construction except for the omission of the subject pronoun are given in (11)–(16) below. (In the examples below, the 'want' verb is printed in boldface, for ease of orientation.)

(11) Mandarin Chinese (Bingfu Lu, p.c.)
(S) \textit{Lisi xiwang Ø zao dian huijia.}
\text{Lisi want early little return}
‘Lisi wants to go back home early.’
(D) \textit{Lisi xiwang Zhangsan zao dian huijia.}
\text{Lisi want [Zhangsan early little return}
‘Lisi wants Zhangsan to go back home early.’

(12) Haitian Creole (NT)\(^4\)
(S) \textit{Eròd té vlé Ø touyé-l.} (Mt 14:5)
\text{Herodes PST want kill-him}
‘Herod wanted to kill him.’
(D) \textit{Li pa-t vlé pèsonn konnin li té la.} (Mk 7:24)
\text{he not-PST want [nobody know he PST there]
‘He didn’t want anybody to know that he was there.’

\(^3\) However, in the larger and more balanced sample of Haspelmath (2005/2008c), a quarter of the languages (72 out of a total of 283) have an overtly expressed subject, and almost all of them are like Modern Greek and Standard Arabic.

\(^4\) For Haitian Creole and a number of other languages, my evidence comes from the New Testament (NT) in these languages.
(13) Manipuri (Bhat & Ningomba 1997:309)
(SS) oy Ø cat-po pam-mi
   I   go-INF want-NFUT
   ‘I want to go.’
(DS) oy ma-na-Ø cat-po pam-mi
   I   he-NOM go-INF want-NFUT
   ‘I want him to go.’

(14) Retuarã (Strom 1992: 160)
(SS) waʔia Ø-eʔe-ri-ka ko-yapa-yu
   fish   get-NMLZ-N 3SG.F-want-PRS
   ‘She wants to get fish.’
(DS) waʔia yi-ʔe-ri-ka ko-yapa-yu
   fish   1SG-bekōm-N NMLZ-N 3SG.F-want-PRS
   ‘She wants me to get fish.’

(15) Godoberi (Haspelmath 1996:188)
(SS) ilu-Ł qʷaraʕ-an-da Ø b-al-i.
   mother-DAT want-CVB-COP N-read-INF
   ‘Mother wants to read.’
(DS) ilu-Ł qʷaraʕ-an-da waša caXawa wu-n-i.
   mother-DAT want-CVB-COP boy[ABS] away M-go-INF
   ‘Mother wants the boy to go away.’

(16) Tagalog
(SS) ibig ni Herodes na ipapatay si Juan
    want ART.ERG Herodes COMP kill ART.ABS John
    ‘Herod wanted to kill John.’ (Mt 14:5)
(DS) ibig ba ninyo-ng palayain ko ang Hari ng mga Judio?
    woll Q you.ERG-COMP release L.ERG ART.ABS King GEN PL Jew
    ‘Do you want me to release for you the King of the Jews?’ (Mk 15:9)

In addition to simple subject omission, we observe a number of further economy effects, which have in common that the SS pattern tends to be formally simpler or shorter. These are described and exemplified in §6-8.

6. Simple complementizer omission (type 3)

In a number of languages, there is an overt complementizer in different-subject complements, but this complementizer omitted in same-subject complements, and this is the only difference between the SS and the DS pattern. Languages with this contrast type are Ju|h’hoan, Maltese, Chalcatongo Mixtec, Hmong Njua, and Tzutujil.
It-tifel jrid jiġi d-dar kmieni.

‘The boy wants to come home early.’

The boy wants him to come home early.

kuní=ri kée=rí

‘I want to eat (something).’

I want you to go.

Pam as nös-ni-qe naawakna.

‘He wants to eat.’

He wants me to eat.’

n-dem n-man ar dò mo cin dì

‘I want to know how they do it.’

I want them to do it.’

In some cases, the complementizer is reduced and merges with the ‘want’ verb, but only in the same-subject construction. This is the case in English and in Yoruba:

I wanna do it. (< want to) (cf. What do you wanna do?)
I want her to do it. (cf. *Who do you wanna do it?)
Simple complementizer omission (or complementizer shortness) is rarer in my sample than complementizer omission and different verb form, a type to which we now turn.

7. Complementizer omission and different verb form (type 4)

The largest group of languages in my sample (20 languages) is those that omit the subject, the complementizer, and also have a different verb form (usually called "infinitive") in same-subject constructions. This was already illustrated for German in (2a-b) above, and some further examples are given below in (23)-(33). 5

(23) Lango (Noonan 1992: 223–4)
(SS) á-mittò 便秘
1SG-want.PROG come.INF
'I want to come.'
(DS) á-mittò 便秘 3SG-come.SBJV
'I want her to come.'

(24) Hindi (NT)
(SS) shair ne nazm parb-nii caab-ii.
poet ERG poem[F] read-INF.F want-PST.F
'The poet wanted to read a poem.'
(VS) tuu kyaa caab-taa bai ki mañ tere liye kar-û?
you what want-PRS AUX.2SG [that I you for do-SBJV.1SG]
'What do you want me to do for you?' (Mk 10:51)

(25) Finnish (NT)
(SS) me tabdo-mme nāb-da sinu-lta merkin (Mt 12:38)
we want-1PL see-INF you-ELAT sign
'We want to see a sign from you.'
(VS) mihin tabdo-t, että valmista-mme pääsiäislampaan...? (Mt 26:17)
where want-2SG [that prepare-1PL Passover.meal]
'Where do you want us to prepare the Passover meal?'

5 It appears that when the complementizer is omitted and a different verb form is used, the subject is always omitted as well. At least I am not aware of a language that shows complementizer omission and a different verb form in the same-subject construction, but does not omit the subject. I have no explanation for this.
(26) Indonesian (NT)
(SS) *Herodes ingin membunuh-nya.* (Mt 14:5)
Herodes want kill-3SG.OBJ
‘Herodes wanted to kill him.’
(VS) *aku ingin, supaya kamu hidup tanpa kekuatiran.* (1 Cor 7:32)
I want, that you
‘I want you to be without worry.’

(27) Basque (Andolin Eguzkitza, p.c.)
(SS) *Aita-k ō goiz etorr-i nabi du etxe-ra.*
father-ERG early come-PTCP want IND.3SG.ERG house-LOC
‘Father wants to return home early.’
(DS) *Aita-k ama etxe-ra goiz etordadin nabi du.*
father-ERG [mother(ABS) house-LOC early come SBJV.3SG.ERG] want IND.3SG.ERG
‘Father wants mother to return home early.’

(28) Lezgian (Haspelmath 1993: 369)
(SS) *Nabisata-z ktab k’el-iz k’an-zawa.*
Nabisat-DAT book read-INF want-IPFV
‘Nabisat wants to read a book.’
(DS) *Nabisata-z ruš-a ktab k’el-na k’an-zawa.*
‘Nabisat wants the girl to read a book.’

(29) Coptic (NT)
(SS) *ten-uōš e-naw e-u-mēni əntot-k* (Mt 12:38)
1PL-want ALL-see ALL-ART-sign from-2SG
‘We want to see a sign from you.’
(DS) *u p-ete-k-uāš-f ṅonta-ai-f na-k?* (Mk 10:51)
what ART-REL-2SG-want-3SG.OBJ 1SG.SBJV-do-3SG.OBJ for-2SG
‘What do you want us to do for you?’

(30) Swahili (NT)
(SS) *tw-a-taka ku-ona isbara kwa-ko* (Mt 12:38)
1PL-PRS-want INF-see sign from-you
‘We want to see a sign from you.’
(DS) *m-na-taka ni-wa-fingu-li-e mfalme wa Wayahudi?* (Lk 19:14)
2PL-PRS-want 1SG-2PL.OBJ-free-APPL-SBJV king GEN Jews
‘Do you want me to release the king of the Jews for you?’

(31) Krongo (Reh 1985: 335–37)
(SS) *n-átàasà ā?àŋ ākò óudà* 1-want.IPFV I INF.eat meat
‘I want to eat meat.’
8. Other types of formal simplicity

8.1. The verb ‘want’ is shorter in the same-subject construction (type 5)

In quite a few languages, it is not so much the complement construction, but the verb ‘want’ itself that is more reduced in the same-subject pattern. In the present sample, this is the case in Samoan (cf. 34), Boumaa Fijian (cf. 35), Korean, Drehu, and Labrador Inuktitut.

(34) Samoan (Oceanic) (Mosel & Hovdhaugen 1992: 710, 714)

(SS) e fia si’i e Leona Iosefa
    GENR want carry ERG Leona Iosefa
‘Leona wants to carry Josefa.’

(DS) e le mana’o le teine e fasi ia le tama
    GENR NEG want ART girl [GENR hit she ART boy]
‘The girl doesn’t want the boy to hit her.’

(35) Boumaa Fijian (Dixon 1988: 91)

(SS) au via nasu’a bulumakau yai.
    I DESID tie-TR cow this
‘I want to tie up this cow.’
8.2. ‘Want’ is expressed as a desiderative affix (type 6)

The expression of ‘want’ may be so short that it is affixed as a desiderative marker to the main verb (8 languages in the current sample; 45 out of 283 languages in Haspelmath 2005). For different-subject sentences, a different verb has to be used.

(36) Erromangan (Crowley 1998:134)
(SS)  *yacam-ampy-omonki.*
1SG.PRS-DESID-drink
'I want to drink.'

(37) Japanese (Kaoru Horie, p.c.)
(SS) *Taroo-wa orenzi-ga tabe-tai.*
Taro-TOP orange-NOM eat-DESID
‘Taro wants to eat an orange.’

(38) Tümpisa Shoshone (Dayley 1989: 384)
(SS) *hi-nna üü bi-pi-suwa-nna?*
what-OBJ you drink-want-GEN
‘What do you want to drink?’

According to Dayley (1989: 385), "there is no direct equivalent of [the different-subject construction]; the closest would be with a verb of telling instead of wanting... Thus, instead of saying 'I want someone to do something', one would say something like 'I told/will tell someone to do something'."

Similarly, in Acehnese, the verb *tém* can be used in SS patterns such as (39):

(39) Acehnese
(SS) *Haha-wa  Taroo-ni suupu-o non-de bosii.*
mother-TOP Taro-DAT soup-ACC drink-CVB want
‘Mother wants Taro to eat a soup.’

8.3. The different-subject construction does not exist (type 7)

In some languages (six in my sample), there simply is no different-subject construction, and some kind of paraphrase has to be used. For example, in Tümpisa Shoshone, *-suwa* expresses 'want', but only in a same-subject configuration:

(38) Tümpisa Shoshone (Dayley 1989: 384)
(SS) *bi-nna üü bi-pi-suwa-nna?*
what-OBJ you drink-want-GEN
‘What do you want to drink?’

According to Dayley (1989: 385), "there is no direct equivalent of [the different-subject construction]; the closest would be with a verb of telling instead of wanting... Thus, instead of saying 'I want someone to do something', one would say something like 'I told/will tell someone to do something'."

Similarly, in Acehnese, the verb *tém* can be used in SS patterns such as (39):

The use of an affix in different-subject ‘want’ constructions is very rare; an example is Labrador Inuttut (Smith 1982:173).
(39) Acehnese (Durie, Bukhari and Mawardi 1994:177)

\[ Lôn-têm \ woe. \]
\[ I \- want \ return \]
'I want to return.'

However, the closest equivalent of the DS pattern is a sentence with *lakèe* 'ask', such as (40) (Durie et al. 1994:178):

(40) \[ Lôn-lakëe \ droeneuh \ beu-neu-woe. \]
\[ I \- ask \ you \ SBJV\-you\-return \]
'I ask you to return.' (May be used for 'I want you to return.')

Non-existence of a pattern is not strictly speaking a case of “formal complexity”. However, one may take the point of view that every meaning can be expressed in some way, and if there is no grammatically paradigmatic expression, then there is some other way of expressing the relevant idea. What matters is that the more frequent pattern is the one that can be expressed compactly and economically by grammatical means.

9. Economy vs. iconicity

There are two main arguments in favour of the economy explanation, and against the iconicity explanation proposed by Givón and Cristofaro.

First, the crucial concepts of “conceptual closeness” and “morphosyntactic integration” are both rather vague, and difficult to measure objectively. The economy explanation is much easier to test empirically: Frequency of occurrence can easily be measured by examining texts in almost any language (see §2), and shortness of expression can also be determined quite easily.

Second, it is unclear why in the specific case of same-subject and different-subject complements of 'want', participant sharing (on the conceptual side) should be iconically reflected by complementizer omission (on the morphosyntactic side). One would expect it to be reflected by subject omission (as happens in many cases, of course), but not by other kinds of shortness effects. But in §5-7 we saw that these other kinds of formal simplicity effects are also widely found in languages, as predicted by the economy account.

In Haspelmath (2008a), I showed that there are quite a few other areas of language structure where economy explanations are the best explanations, even though linguists have sometimes proposed iconicity explanations.

10. Explanation by functional adaptation and diachronic change

In the foregoing, I have noted a cross-linguistically general correlation between some aspects of language form and usage patterns. But this does not amount to an explanation yet – we still need to demonstrate a possible causal link between the motivating factor of frequency and occurrence and the formal patterns. This is a challenging task, but fortunately, it is not specific to the particular case at hand. Frequency has strong effects on
all aspects of language structure, and since Zipf (1935), it has been clear that something like a principle of least effort must be at play, whatever the exact mechanisms (see also Croft 2003:§4.3, Haspelmath 2008a).

It is also clear that the implementation of the usage-form correlation must happen via diachrony. We speak the way we do because we follow other speakers' behaviour, not because we strive for a particular economical design of our language. But making this more precise is very difficult, because we know so little about language change. Only very few languages have an attested history, and the attestation of this history is very rudimentary. Nevertheless, I would like to suggest that there are two diachronic pathways by which the economical patterns that we saw above can be created: differential phonetic reduction, and differential selection of constructions (see Haspelmath 2008b).

We see differential phonological reduction in the case of English and Yoruba (cf. (21)-(22) above), in the well-known case want to > wanna. It has long been known that frequent combinations undergo greater phonetic reduction because the information is more predictable, and speakers can afford to speak with less effort. This is the classical Zipfian explanation.

However, it does not seem likely that this diachronic pathway is particularly important in explaining the patterns of 'want' complements. Subject omission and complementizer omission are not very likely to be due to phonological reduction. Instead, these patterns seem to arise by selection during the process of grammaticalization. In general, tightly constrained patterns of grammar ultimately derive from looser constructions. Verbs with the meaning 'want', which tend to have tightly constrained complement clause patterns, tend to derive from verbs such as 'wish, desire, need', which often have less tightly constrained complement clause patterns. Compare German 'wollen', which only allows a single pattern, in (41a-b), with wünschen 'wish', which allows both a finite dass-clause and an infinitival complement in the same-subject configuration (see 42a-b). This is thus a looser pattern, and if something like (42) turns into something like (41), a selection takes place: Only the more economical pattern (42b) survives.

(41) a. Ich will früh heimkehren. (*Ich will, dass ich heimkehebr.)
    'I want to return home early.'
    b. Ich will, dass du früh heimkehrst. (*Ich will du früh heimkehren.)
    'I want you to return home early.'

(42) a. Ich wünsche mir, dass ich früh heimkehre.
    'I desire that I return home early.'
    b. Ich wünsche mir, früh heimzukehren.
    'I desire that I return home early.'

Another looser construction type that may be the antecedent of a 'want' construction is a pattern with nominalization, as in (43). Here the subject may be made explicit by a possessive pronoun, as in (43b-c), but it may also remain implicit.

---

(43) a. Ich wünsche mir eine frühe Heimkehr. ‘I desire an early return.’
b. Ich wünsche mir meine frühe Heimkehr. ‘I desire my early return.’
c. Ich wünsche mir deine frühe Heimkehr. ‘I desire your early return.’

Again, if in diachronic change, a pattern such as (43a) (with no explicit subject) survives and becomes the only possibility, there is no reduction, just differential selection of economical patterns during grammaticalization.

At the earlier stage of pattern freedom, speakers must have preferred the infinitival pattern for the same-subject case because it allowed more economical utterances, whereas the finite das pattern eventually survived only in the different-subject case. Thus, speakers’ preferences in utterances can give rise to functionally adaptive patterns in grammar. This account is somewhat speculative, and I cannot prove it, but something along these lines must eventually be said for a complete economy-based explanation of the observed tendencies of ‘want’ complement clause coding.

11. Conclusion

My economy-based or frequency-based explanation is an instance of what is now generally called "usage-based" explanation of general patterns of language structure. This approach stands in a certain tension to the Saussurean view with its separation of langue and parole (cf. Newmeyer 2003). In Ferdinand de Saussure’s *Cours de linguistique générale*, we find the following metaphor:

[O]n peut comparer la langue à une symphonie, dont la réalité est indépendante de la manière dont on l’exécute; les fautes que peuvent commettre les musiciens qui la jouent ne compromettent nullement cette réalité.8
(Saussure 1972[1916]:36)

But in contrast to Saussure’s metaphor, in human language, competence (=langue) is indeed influenced by performance (=parole), through the different treatment of construction with different frequencies in language change.

However, this was not unknown in Saussure’s time, and in fact we also find another quotation in the *Cours*:

la langue est à la fois l’instrument et le produit de la parole9
(Saussure 1972[1916]:37)

Thus, a language is crucially different from a symphony, which is a blueprint for the musicians’ performance, but cannot be said in any way to be its product. The frequency with which a certain structure is used by speakers has a strong impact on its precise form,

---

8 "One can compare langue to a symphony, whose reality is independent of the manner in which it is performed; the errors committed by the musicians do not compromise this reality in any way."

9 "Langue is at the same time the instrument and the product of parole."
through the perpetual recreation of language in language leaning and language change, which has no analog in the creation of a symphony.

**Abbreviations**

Category abbreviations follow the Leipzig Glossing Rules. Additional abbreviations:

- **CONST** construct
- **DESID** desiderative
- **ELAT** elative
- **GENR** generic
- **MOD** modal
- **OPT** optative
- **POT** potential
- **SUBORD** subordinate

**References**


Scaliger, Julius Caesar (1584[1540]), *De causis linguae Latinae libri tredecim*. [Heidelberg]: Apud Petrum Santandreamum.


