

The Grammaticalization of Complementizers

Author(s): Evelyn N. Ransom

Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society (1988), pp. 364-374

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

The Grammaticalization of Complementizers
Evelyn N. Ransom
Eastern Illinois University

I. Introduction. One type of grammaticalization has been described by linguists like Givón, Traugott, Lehmann, and Bybee and Pagliuca as consisting of stages in simultaneous semantic and phonological reduction-- from full lexical items with specific and general content to more reduced forms and more generalized meanings. These stages have been viewed as occurring on a unidirectional scale, like the following:

A GRAMMATICALIZATION SCALE

FULL LEXICAL -----> REDUCED LEXICAL
MEANING AND FORM MEANING AND FORM

This type of grammaticalization can be found with complementizers from different languages. First, we will examine the way that complementizers can be viewed as representing different stages of development on the grammaticalization scale, then the way the current meanings and forms of the complementizers are related to their lexical sources and their position on the grammaticalization scale.

The questions we will address are "To what extent are the current meanings of complementizers correlated with their lexical sources and their particular stage of development?" and "What implicational universals can be observed?"

Before we look at the data, a warning is in order. The specific complementizer analyses are based on isolated examples and commentary. More in-depth study of these languages and the history of the forms is needed for reliability.

II. Stages of Complementizer Development. When we look at complementizers in different languages, we find that many of them are related to morphemes still being used for other purposes.¹ Those related to full lexical content words, such as nouns and verbs, will be viewed as representing an early stage of complementizer development, shown on the left end of the grammaticalization scale. Those related to lexical function words, such as determiners or pronouns, will be viewed as representing a medial stage of development, shown in the middle of the grammaticalization scale. Those related to bound forms with more abstract functions, such as mood-like forms, will be viewed as representing the last stage of development before the form is

completely lost, as shown on the right side of the grammaticalization scale. Examples of these stages of complementizer development are shown in Table 1.

TABLE 1
Stages of Complementizer Development

FULL LEXICAL MEANING AND FORM	PARTIALLY REDUCED LEXICAL MEANING AND FORM	REDUCED LEXICAL MEANING AND FORM
--	---	---

Korean <u>kes</u> 'thing' or 'fact'	Kanuri <u>de</u> & <u>-ro</u> def.art.&loc./dat.	Basque <u>-ela</u> , <u>-teko</u> , <u>-en</u>
Thai <u>thii</u> 'place' or 'at'	English <u>that</u> dem.pro. <u>for</u> and <u>to</u> loc./dat.	
Thai <u>hây</u> 'cause' or 'give' etc.	Russian <u>cto(by)</u> rel.pro. <u>ctò</u> interrog. pro	
Ewe <u>bé</u> 'say'	Maori <u>ki</u> loc./dat.	

Because the words in the full lexical stage retain lexical content, linguists have questioned their status as complementizers, just as they have questioned the status of the more mood-like forms in the reduced stage. What characterizes all of these forms as complementizers is that syntactically they can be used to signal the closed boundary of the clause, the one that does not expand (left boundary for SVO; right for SOV); and semantically they can be used to signal the modality of the complement proposition. This scalar approach to the grammaticalization of complementizers makes it possible to deal with some of the problems in treating complementizers as a unified category and in showing their relationship to mood forms.

First let us look at the words in the left column, which can be used either with their full lexical meaning and form or as complementizers. When one of these words is used with full lexical content, it has certain semantic and syntactic values peculiar to that word. For example, the Korean word kes in (1) has the meaning of a concrete thing and functions as the subject of the sentence:

- (1) kes as a noun = 'thing'
Ku kes un chayk iey yo
That thing is a book. (Martin and Lee 1969)

In (2), kes could not be interpreted as a concrete thing, since one cannot know concrete things, but it could be interpreted with an abstract meaning as the

nominal object 'fact' taking an appositive clause or else as a complementizer marking an object clause (the brackets are mine):

- (2) kEs as an abstract fact or a complementizer
na nin ki ka o-nin kEs lil al-nin-ta
I he come-PRS COMP/[fact] ACC know-PRS-M
I know [the fact] that he is coming. (Kim 1974)

However, in (3), neither nominal interpretation would be possible, since one cannot order things or facts but only acts, and a complementizer interpretation is used:

- (3) na nin ki eke ka-l kEs lil myEnglyEngha-ess-ta
I him go-FUT COMP ACC order PAST-M
I ordered him to go. (Kim 1974)

In both (2) and (3), kEs can be seen as syntactically marking the closed boundary of the object clause (the right boundary, since Korean is SOV) and as semantically signaling the modality of the complement proposition, namely that the proposition is expected to be the case.²

Now let us look at the examples in the middle of the scale. These forms are usually monosyllabic function words with little content, and they are usually viewed as the typical complementizer. For example, English that can be used as a demonstrative pronoun or as a complementizer. According to the OED, the complementizer form is generally thought to be derived from the neuter singular demonstrative pronoun -at in combination with an appositive clause as in "We all know that: he once lived here."³

As a demonstrative pronoun, that has a full vowel, a plural form those, and refers to something pointed out at a distance (in contrast to this and these, which refer to something pointed out nearby), as shown in (4):

- (4) That is his sportscar/Those are his sportscars.

When that is used as a complementizer, it is usually unstressed, allowing the vowel to be reduced to a schwa; it cannot take a plural form, and does not have either the general meaning of something pointed out or the more specific distal meaning, as shown in (5):

- (5) They know that this is his sportscar.

In (5), that functions as a complementizer. Syntactically it marks the closed boundary of the object clause (the left boundary, since English is SVO), and semantically it signals the modality of the complement proposition, namely that it is inclined to be true.

Now let us look at the examples on the right side of the scale where the meanings lack specific and general content and the forms are reduced to bound endings. Basque seems to have the only clear-cut examples of this stage that I have found so far. The endings -ela, -teko, and -en are attached to either the auxiliary verb or the infinitive, making them somewhat like mood forms:

- (Txomin Joseba Koldobika: native consultant)
- (6) -ela as complementizer or completive mode
 Marik atea hertsi zu-ela gogoratu zuen
 Mary door closed COMP remembered
 Mary remembered that she closed the door
- (7) -teko as complementizer or genitive gerund
 Marik Jon atea herts-teko bultzatu zuen
 Mary John door close-COMP forced
 Mary forced John to lock the door
- (8) -en as complementizer or conjunctive mode
 Marik atea hertsi ote zu-en edo ez zalantzan zen
 Mary door closed COMP or not doubt
 Mary wondered whether she had locked the door

Traditionally, the bound form -ela is called a Completive Mode but has no specific meanings; -en is called a Conjunctive Mode with no specific meanings but is said to resemble a subjunctive mood; and -teko is called a Genitive Gerund, since it has part of the form of a gerund (-te) and part of the form of a genitive of inanimate possession or location (-eko). However, since Basque is a SOV language, these verbal endings function as markers of the closed boundary of the clause and as signals of the modality of the complement proposition, just as complementizers do.

Thus we see that complementizers can occur at different stages of development on the grammaticalization scale with full or reduced meanings and forms. Now let's turn to the meanings of complementizers.

III. Complementizer Meanings. In language after language, complementizers are used to signal the modality meanings of the complement proposition. There

are two types of modality meanings: the Information Modalities, which are related to the epistemic-deontic or root contrast, and the Evaluation Modalities, related to what Palmer calls judgments (cf. Palmer 1986 and Ransom 1986).

The Information Modalities consist of four types, (Truth, Future Truth, Occurrence, and Action) each distinguished by the type of complement predicate the proposition can take. The Truth Modality treats a complement as being about the truth or falsity of its proposition and has no restrictions on the type of predicate, as in "We know, believe, hope that her opponent is a female."

The Future Truth Modality treats a complement as about the future truth of its proposition, and requires the complement predicate to be interpreted as capable of change in the future so that permanent states, like being a female are unacceptable unless we can find a changeable interpretation. It would be acceptable for Rene Richards to say he expected/was eager to be a female.

The Occurrence Modality requires the complement predicate to be interpreted as an event and thus requires nonstative verbs or states viewed as a process of coming into being. It may be strange to say "We watched Harry become tall", but we could say "We watched Alice in Wonderland become tall".

These three modalities are related to the epistemic modalities since they are about knowledge of the world, whether something is true, will come true, or did occur. The fourth type of Information Modality is more closely akin to the deontic or root modalities. It is not about knowledge of the world but about volitional, controllable actions and thus like the imperative, promises, and permission in that the predicate must be interpreted as a controllable act and its subject must be interpreted as an agent who is either motivated by others or self-motivated. Thus neither states nor events are usually acceptable. One cannot decide to be a woman, to be tall, or to fall, unless one is Rene, Alice or a fall guy.

These four Informational Modalities can be represented by complementizers, as in the contrast in English between tell that and tell to.

The Evaluation Modalities also consist of four types: Predetermined, Determined, Undetermined, and Indeterminate. They describe judgments about the alternatives available to a proposition. A Predetermined Evaluation is interpreted as definitely the case

with no alternatives, as in the complements of sentences like "She regretted that he left"; "She forced him to leave"; or "She managed to leave".

A Determined Evaluation is interpreted as probably the case but with some alternatives, as in the complements of sentences like "I am afraid that she left"; "I told her to leave"; and "I decided to leave".

An Undetermined Evaluation is interpreted as possibly the case but with many alternatives, as in the following complements: "I hope that she left"; "I permitted her to leave"; and "I was willing to leave".

An Indeterminate Evaluation is interpreted as possibly the case or not -- with equal alternatives, as in the following complements: "I wonder whether she left"; "I wonder whether to leave"; and "I told them whether to leave".

The Evaluation Modalities are typically represented by modals like must, should, may, and may or may not, but they can also be represented by complementizers like Korean kes, ko, ki, and ci or English tell that and tell whether.

The four Information Modalities and the four Evaluation Modalities combine to make sixteen possible modalities that a complement can have. Table 2 shows examples of higher predicates that typically select certain combinations of those modalities for their complements, just as they select the semantic features of other types of arguments.

TABLE 2
HIGHER PREDICATES SELECTING THE COMBINED MODALITIES

	TRUTH	FUT TRUTH	OCCUR	ACTION SELF-DIRECTED-OTHER	
PREDET	know regret true	foresee anticipate certain(sr)	watch cause occur	remember get obligatory	force compel
DETER	believe say probable	expect predict likely(sr)	wait tend	decide promise important	persuade command
UNDET	hope speculate possible	look for eager desirable	like(i) ready fun	willing able acceptable	permit allow
INDET	wonder know uncertain	foresee predict	watch	wonder remember clear	tell remind

(i=infinitive; sr=subject raised construction)

It is usually these sixteen combinations of modalities that are signalled by complementizers. English uses whether or if for Indeterminate Truth and Action; that for Predetermined, Determined and Undetermined Truth and for-to for Predetermined, Determined, and Undetermined Action.

Given that complementizers can be used to represent certain combinations of modality meanings, let us next look at the lexical sources of these complementizers, keeping in mind the questions that we asked earlier: "To what extent are the current meanings of complementizers correlated with their lexical sources and their particular stage of development on the grammaticalization scale?" and "What implicational universals can be observed?"

IV. Lexical Sources for Complementizers. We saw earlier that complementizers across languages differ according to their stages of development on a grammaticalization scale. Some are derived from nouns and verbs with full lexical meanings and forms; some come from function words which are usually lacking in lexical content, and some come from mood-like forms with no specific meaning. Now let us compare complementizers that come from similar lexical sources.

First of all, let us look at complementizers that come from nominals, like Korean kEs and Japanese koto meaning 'thing' or 'fact' and Thai thii meaning 'place' or 'at'. These nouns have the meaning of something concrete or abstract existing, and one thing their cognate complementizers have in common is that they are typically used to represent a Predetermined Truth Modality in the complement (= factives or epistemic necessity constructions). They differ in the ranges of modalities they can represent. Korean kEs is also typically used to represent Predetermined Action (=deontic necessity). Japanese koto tends to represent Predetermined and Determined ranges of Truth and Action, allowing more alternatives than the others do.

Complementizers that come from determiners and pronouns seem to have a similar range to those from nominals. Kanuri de is a definite article with the meaning 'all of a set'. When it is combined with the locative marker -ro, it can be used as a complementizer for Predetermined Truth complements.

Also definite are pronouns, a major source of complementizers. English and other West German languages, and Aramaic derived a complementizer from a demonstrative pronoun while Russian, Latin, and Greek

derived a complementizer from a relative pronoun. It would be interesting to trace the earliest histories of the complementizer uses to see if they began with Predetermined Truth complements, which I would predict, and then to check the current uses of the forms to see what ranges they could occur with. In the OED, the earliest examples of that as a complementizer appear to be with factual complements, but a study of earlier texts is needed for a more reliable sampling.

In current usage, English that and apparently Russian cto occur with Predetermined, Determined, or Undetermined Truth complements. English that and Russian cto by with a subjunctive mood can occur with certain Action complements. English that + subjunctive can occur with Predetermined and Determined Action complements functioning as subjects, and a few exceptional object complements.

English whether, which is derived from an interrogative marker meaning "which of the two" retains its interrogative and its alternative meaning by representing Indeterminate Truth and Action.

These nominal, determiner, and pronominal sources tend to have in common that their complementizers can be used with Truth complements. Now let us turn to verb sources. The most common verbal source for complementizers is the verb say, and as one might suspect, the complementizers can all occur with indirect speech complements, especially Determined Truth. In Thai waa is still used to mean 'say' and the complementizer carries that meaning also. Typically, waa occurs with Determined, Undetermined, and Indeterminate Truth complements, but not with the complements of emotive factives or certain subject complements, and not with Action complements. In Krio (Givón 1980), the complementizer se from the verb 'say' appears to occur mostly with Determined and Undetermined Truth complements. Korean ko is from the verb say, and it tends to occur with both Determined Truth and Action. Lord (1976) points out that Uto-Aztecan (Munro 1974), has a complementizer from the verb 'say' that occurs only with verbs of saying, while others like Ewe, Yoruba, Akan, Tamil, Telugu, and Kera use it more extensively.

So far we have looked at the lexical sources of complementizers that tend to be used for Truth complements. Now let us look at the lexical sources of complementizers that tend to be used for Action complements. Looking first at verb sources, Thai has the complementizer hai that is derived from a verb meaning 'give' or 'cause', which is about events or

actions and so it is not surprising to find it taking Action complements.

A very popular source of complementizers is from locative and dative markers, and these tend to occur with Action complements. English for and to typically occur with Predetermined, Determined or Undetermined Action. In Kanuri, the locative-dative marker -ro appears to occur with Predetermined, Determined, and Undetermined Action complements, but it also occurs with Determined and Undetermined Truth complements. Thus the locative and dative meanings of direction, futurity, uncertainty, location, or animate receiver have become generalized to the point where these markers can occur with an extensive range of complement modalities.

By looking at the lexical sources of complementizers, we can compare languages whose complementizers may have a similar source but a different range of meanings, as with Kanuri de and English that, or a similar range but a different source, as with English that and the Thai quotative wāa. We can find out, from the types of metaphorical extensions of concrete lexical referents to abstract grammatical markers, how certain semantic fields are associated cognitively, such as definiteness and existence with truth, or direction with futurity, possibility, or purpose.

V. Conclusion. In conclusion, we can say that there appear to be tendencies for both the lexical sources and the stage of development of the complementizer to influence the kinds of modality the complementizer can be used to signal. In order to pin down these tendencies, more in-depth study of complementizers and their development in particular languages is needed. Can we claim that nominal sources tend to become grammaticalized into complementizers marking Truth complements first before spreading to other modalities, while the more adverbial locatives tend to become grammaticalized into complementizers for Action complements first before spreading to other modalities? When verbs become grammaticalized into complementizers, is it possible to claim that they tend first to become complementizers for the kinds of complements that they as verbs could take before spreading to other kinds of modalities? Based on reliable data, both synchronic and diachronic, it might be possible to establish implicational generalizations about the directions of semantic change.

BIBLIOGRAPHY

- Bybee, Joan L. and William Pagliuca. 1985. Cross-linguistic Comparison and the Development of Grammatical Meaning. *Historical Semantics: Historical Word Formation*, ed. by Jacek Fisiak. 59-83.
- Givón, Talmy. 1980. The Binding Hierarchy and the Typology of Complements. *Studies in Language*. 4.333-377.
- Kim, N. 1974. *Studies in the Syntax of Korean Complementation*. University of Washington Dissertation.
- Lehmann, Christian. 1985. Grammaticalization: Synchronic Variation and Diachronic Change. *Lingua e Stile* 20:303-318.
- Lord, Carol. 1976. Evidence for Syntactic Reanalysis: From Verb to Complementizer in Kwa. *Papers from the Parasession on Diachronic Syntax*. Chicago Linguistic Society.
- Noonan, Michael. 1985. Complementation. *Language Typology and Syntactic Description: Complex Constructions*, ed. by Timothy Shopen.
- Martin, Samuel and Young-Sook C. Lee. 1969. *Beginning Korean*. Yale University Press.
- Palmer, F. R. 1986. *Mood and Modality*. Cambridge University Press.
- Ransom, Evelyn. 1986. *Complementation: Its Meanings and Forms*. Benjamins.
- Traugott, Elizabeth Closs. 1982. From Propositional to Textual and Expressive Meanings: Some Semantic-Pragmatic Aspects of Grammaticalization. *Perspectives on Historical Linguistics*, ed. by Winfred P. Lehmann and Yakov Malkiel. Benjamins. 245-271.