Some Asymmetries between Initial and Final Subordination Markers

Jiro INABA

1. Introduction

The possibility of recursive embedding can be regarded as one of the most fundamental properties of human language (cf. e.g. Chomsky 2008: 136). A matrix clause can contain within it another clause which is subordinate to the former, e.g. in the form of a sentential object. In such cases of clausal embedding or subordination, we often find a subordination marker or a clause linker signaling the subordination relationship:

(1)   a. I believe [that he will come]
        b. ich glaube [dass er kommt]
        I believe [Comp he comes]
(2)   boku-wa [kare-ga kuru to] omou
        I-Nom [he-Nom come Comp] think

In this paper, I will concentrate on such subordination markers that are called "complementizers" (underlined above) from various languages.[1] More specifically, I will investigate their positioning within the subordinate clauses, in relation to the positioning of the subordinate clause with respect to the matrix verb. I will also go into the syntactic properties of these subordination markers, which are usually subsumed under the functional category Comp in the current theoretical literature.

This paper is organized as follows: Section 2 demonstrates on the basis of empirical

---

[1] Huddleston & Pullum (2002: 954f) claim that the head of the clausal complement in a sentence like (i) is not *that*, which they call marker, but the following part *it's wrong* (see also Matthews 2007: Ch.3 for discussion):
i)   I know [that it's wrong].
This is reminiscent of the S-S'-schema before Chomsky (1986) (e.g. Chomsky 1981) or of the idea adopted by Pollard & Sag (1994). As evidence for the non-headedness of *that* in question, Huddleston & Pullum mention the "omissibility" of *that* and a certain licensing requirement. While I do not share their opinion concerning the English *that* or the corrensponding elements in the relevant languages, I in principle use a theory-neutral term "subordination marker" or "subordinator" in this paper for those elements that are usually called Comp(lementizer) in the generative literature. Just in glosses I stick to the traditional "Comp" for the sake of simplicity and without any theoretical implication.

data that there is a universal tendency concerning the positioning of the sentential object and the subordination marker within it; the matrix predicate and the subordination marker of the embedded object clause tend to be adjacent. In section 3, I point out that there is a systematic difference between the clause-initial and the clause-final subordinators cross-linguistically: While clause-initial subordinators can have properties of a general complementizer subordinating (exclusively) clauses, it is not the case with clause-final ones. In the ensuing section, I will provide an account for this asymmetry by appealing to a sentence-processing strategy (cf. Hawkins 1994). Section 5 concludes the discussion of this paper.

2. Positioning of sentential objects and subordination markers

Languages of the world exemplify a certain tendency as to the relative positionings of the matrix predicate and its sentential object, on the one hand, and the position of the subordination marker within that object clause, on the other. For the sake of concreteness, let us list up the possible ordering variations of the matrix predicate (V_1), the subordination marker (Comp), and the embedded sentence (S_2). We want to concentrate on cases in which the matrix verb and its sentential object form a constituent or are adjacent, which is semantically motivated at least in the underlying structure. We also abstract away from the question of whether the S_2 unexceptionally forms a constituent that excludes the subordination marker; the following schemata should be understood linearly:

(3)   a. V_1 [Comp S_2]
         b. [S_2 Comp] V_1
(4)   a. #/* [Comp S_2] V_1
         b. #/* V_1 [S_2 Comp]

As is already pointed out in the literature (cf. among others Kuno 1974, Grosu & Thompson 1977, Dryer 1980, Hawkins 1990, 1994: Ch.5.6, Bayer 1996: Ch.6), the patterns in (3) are more common among languages, while those in (4) are unacceptable, rare, marked or not canonical. This is exemplified by the data as below:[2]

---

[2] Let us ignore, for example, the so-called stylistic inversion in Japanese (or anything similar), a realization of (4b), that is indeed frequently observed in colloquial speech, is nonetheless regarded as somehow not canonical in this OV-language.

[German]

(5)   a. weil ich glaube [<u>dass</u> er kommt] (cf. (1b))

        b. #/?*weil ich [<u>dass</u> er kommt] glaube

[Japanese]

(6)   a. boku-wa [kare-ga kuru <u>to</u>] omou (= (2))

        b. # boku-wa omou [kare-ga kuru <u>to</u>]

Interestingly, this state of affairs applies also in languages, like Bengali, in which there are both clause-initial and clause-final subordination markers; subordinate clauses with the initial subordinator appears to the right of the matrix verb, conforming to the pattern in (3a), while for those with the final subordinator, the opposite positioning (= (3b)) is attested (Bayer 1996: 255):

[Bengali]

(7)   a. chele-Ta Sune-che [<u>je</u> [or baba aS-be]]

        boy-CF hear-Pst3 [<u>Comp</u> [his father come-Fut3]]

        'The boy has heard that his father will come.'

        b. *chele-Ta [<u>je</u> [or baba aS-be]] Sune-che

(8)   a. chele-Ta [[or baba aS-be] <u>bole</u>] Sune-che

        boy-CF [[his father come Fut3] <u>Comp</u>] hear-Pst3

        'The boy has heard that his father will come.'

        b. *?chele-Ta Sune-che [[or baba aS-be] <u>bole</u>]

The generalization one can derive from these data essentially corresponds to what Bayer (1996: 193) proposes as "C-visibility": "Where CP is selected by V, its head tends to be linearly adjacent to V." We can reformulate it in a theory-neutral way as follows:

(9)   When a matrix predicate selects a sentential object, the subordination marker of the latter tends to be adjacent to the former.

This seems to be a typologically valid generalization, as we will ascertain in the following shortly.

        What is interesting from a cross-linguistic perspective is that in some languages the nominal and the sentential objects behave differently with respect to their positioning to the matrix verb. To be more concrete, let us look at the following table:

(10)

	V > nom.obj.	nom.obj. > V
V > sent.obj.	(a) English	(b) German
sent.obj. > V	(c) ---	(d) Japanese

English is a uniformly head-initial language, at least as far as a verb and its complement are concerned, while Japanese is a strictly head-final language. In each of these languages, the canonical position of the object relative to the verb remains the same, independent of whichever syntactic category the object belongs to. This is not the case in German: In this OV-language, complements (and also adjuncts) are in principle realized to the left of the verb. When a clause appears as a complement, however, its positioning to the right of the verb is preferred to the OV-order (see (5)) which otherwise is the canonical pattern in this language.

This state of affairs is apparently connected with the fact that the subordination marker *dass* in German is clause-initial: Only by way of positioning the clausal object with *dass* to the right of the matrix verb can the requirement (9) be fulfilled. Exactly the same is the case with the sentential object with *je* in Bengali, another OV-language (see (7)). From this observation, one could say that the cross-linguistic generalization in (9) overrides the object-verb-ordering canonical in a given language.[3]

As an exemplification of the schemata (3) and (4), resulting in the generalization (9), we will look over some more data. Let us first recapitulate some of the relevant examples with preverbal sentential objects:

[Bengali]
(11)　　a. chele-Ta [[or baba aS-be] bole] Sune-che (= (8a))
　　　　　　boy-CF [[his father come Fut3] Comp] hear-Pst3
　　　　　　'The boy has heard that his father will come.'
　　　　b. *chele-Ta [je [or baba aS-be]] Sune-che (= (7b))
　　　　c. *chele-Ta [bole [or baba aS-be]] Sune-che (Bayer 1999: 246)

---

[3] In case of conflict between the generalization (9) and the VO- vs. OV-setting of head parameter, what happens is the opposite positioning of the sentential object as a whole (see (5) and the discussion in text) and not the reordering of the initial complementizer into the final position within the embedded clause; i.e. we do not find sentences like the following, although the "C-visibility" is maintained here:
i)　　　　*weil ich [er kommt dass] glaube
ii)　　　　*chele-Ta [[or baba aS-be] je] Sune-che (Bayer 2001: 15)
So far as I see, functional categories, if present at all, (or their counterparts in other languages) seem to be universally fixed with respect to the relative ordering to their complement (or to the element they are attached to). See Inaba (2007: 165f) for discussion.

[Japanese]

(12)   a. boku-wa [kare-ga kuru <u>to</u>] omou (= (6a))

         I-Top [he-Nom come <u>Comp</u>] think

         'I think that he will come.'

         b. *boku-wa [<u>to</u> kare-ga kuru] omou

As already noted, the subordinators of the preverbal object clauses appear in the right-peripheral position. Further data that support (9) are given below from OV-languages with postverbal clausal objects:

[Persian] (Dryer 1980: 130)

(13)   Ān zan mi-dānad [<u>ke</u> ān mard sangi partāb kard].

         that woman Cont-knows [<u>Comp</u> that man rock throwing did]

         'The woman knows that the man threw a rock.'

[Hindi] (Bains 1989: 25)

(14)   Ram kehtA hE [<u>ki</u> anu yeh afsAnA pare gi].

         Ram says is [<u>Comp</u> Anu this story read will]

         'Ram says that Anu will read this story.'

[Megrelian]

(15)   Lela pikrens [<u>namda</u> malas disertacia-s ipcvenki].

         Lela-Nom think-3sg-Prs [<u>Comp</u> soon dissertation-Dat defend-1sg-Prs]

         'Lela thinks that she will soon defend her dissertation.'

In these examples, the preverbal positioning of the clausal object renders each sentence ungrammatical (see Inaba 2007: Ch.6). I should also mention that the languages cited here are in principle all OV-languages, belonging to the German-type (10b), as can be detected by the word order within the embedded clause: In these languages, too, the position of the subordination marker within the embedded clause is so crucial that the canonical OV-ordering is suspended in the case of sentential complementation.

We have thus observed a couple of languages from different language families. Before closing this section, I would like to deliver some remarks concerning generalizations and counter-examples. Although the generalization (9) has a wide range of validity, as we have just seen, there are data that appear to be counter-examples to it:

[Basque] (Lehmann 1984: 59, Hawkins 1994: 389f)

(16)   Esa-ida-zu [zer ari z-era-<u>n</u>]

       tell-Imp-Dat-Erg [what do Abs-Aux-<u>Comp</u>]

       'Tell me what you are (in the process of) doing.'

[Harar Oromo, spoken in Ethiopia] (Dryer 2007: 100)[4]

(17)   [<u>akká-n</u> d'ufé-n] beexa

       [<u>Comp</u>-1Sg came-1Sg] know

       'I know that I came.'

Independent of the exact syntactic character of the subordination markers here, data of this kind cannot be subsumed under (9), so long as the "Comp" plays the role of signalizing the subordination of the embedded clause.

      In the face of this state of affairs, I still would like to stay with the generalization (9). In almost any linguistic research, and especially in typologically oriented ones, any non-trivial generalization is likely to be met by counter-examples, which is true also for the present study. What is asserted as a cross-linguistic generalization can be a mere tendency, and one would just need to bring forward one out of many thousands of languages in order to falsify it. I nonetheless regard the generalizations that hold for most cases as meaningful. I share Dryer's (1980: 188) opinion when he says: "I assume that the existence of real counterexamples weakens a generalization, but does so minimally if they are rare. I assume that generalizations that are true of most languages are no less interesting and just as much in need of explanation as generalizations that are true of all languages." Thus, the existence of a small number of counter-examples to (9) should not invalidate it.

      In this section, I reviewed data from a couple of languages that support the generalization in (9). Some more data for this purpose will be presented also in the next section, where I will mainly look into the properties of the initial and the final subordination markers, respectively.

3. Properties of subordination markers

I demonstrated in the previous section that the patterns in (3), here repeated as (18), are found

---

[4] Dryer (2007: 100) considers this language "atypical among OV&CompClause languages in that the complement clause occurs in normal object position before the verb".

in the case of sentential complementation across languages:

(18)   a. V₁ [Comp S₂]

       b. [S₂ Comp] V₁

At first glance, each of these two variants seems to be a mirror image of the other, the syntactic properties or relationship among the relevant elements being identical. In this section I argue, however, that there is a substantial difference between the initial and the final subordination markers in their syntactic properties. I claim that only initial subordination markers can be regarded as a general complementizer that belongs to the functional category Comp and serves the function of linking a clause, while it is not the case with the final subordination markers.

     In order to make the point clear, let us first take up a "typical" example of initial subordination markers; the English *that* as in (1a). *That* can well be regarded as a general complementizer in that (i) it can introduce embedded finite clauses in general so long as they are declarative, and (ii) *that* can introduce not only object clauses selected by the verb but also clauses associated with or dependent on the noun (cf. (19)).[5] Givón (1991: 25) maintains in this regard that "the sharing of the same subordinating morphemes by REL-clauses and V-complements is widely attested in many languages."[6] As a further support for the complementizer-status of *that*, one could add that (iii) it can function as a linker also in some adverbial clauses (cf. (20)):

(19)   a. the people [that I have never met]

       b. the rumor [that she killed her husband]

(20)   He must be crazy [that he should go out now].

To be noted is the observation that what is introduced by *that* is a (finite) clause and nothing else. The German *dass* also fulfills these criteria. I do not go into the question of how many of these criteria must be met in order for an element to be regarded as a clause-linking complementizer; I just list them as some of the typical properties of it.

---

[5] I regard the criterium (ii) to be met if at least one of the constructions in (19) are attested. I am not sure to what extent it makes sense from a cross-linguistic perspective to distinguish between the so-called complement clause of a nominal and the restrictive relative clause (cf. e.g. Kayne 2005: 239f and An 2007: 53ff for discussion).

[6] Kayne (2005: 237) remarks, however, that the "English *that*-relative does not occur in any Germanic OV language".

For the three other languages with a clause-initial subordinator cited shortly above (see (13)-(15)), I just briefly mention that they pattern more or less with the English *that*. For Persian, for example, Klaiman (1976: 7) maintains: "Persian has only an invariant complementizer *ke* which is positioned clause initially and which doubles as a relativizer. It might be best described as a general subordinizer." See also Windfuhr (1987: 124). Samushia (2007) claims that *namuda/namda* in Megrelian, one of the South Caucasian or Kartvelian languages spoken in Georgia, combines with still other types of subordinate clauses than subject and object clauses, such as consecutive, causal, or purpose clauses. In Hindi, the *ki*-clause can appear as a complement to a series of such verbs and nouns as are presented so far for various languages (Montaut 2004: 243ff).[7]

Let me further explicate the point from just another (less well-known) language: Twi is one of the dialects of the Akan language belonging to the Kwa language family and is spoken in Ghana by about 7 million people. It is an SVO-language and has a clause-initial complementizer *sɛ*.[8] Lord (1993: 159) gives the following examples:

[Twi]

(21)   a. ama pɛ [sɛ kofi bɛyɛ adwuma no]

Ama wanted [Comp Kofi Fut-do work the]

'Ama wanted Kofi to do the work.'

b. na ama nim [sɛ kofi yɛɛ adwuma no]

Pst Ama know [Comp Kofi did work the]

'Ama knew that Kofi had done the work.'

To be noted is that *sɛ* can introduce various other clauses than a declarative complement clause; embedded questions (cf. (22)), purpose clauses (cf. (23)), conditionals (cf. (24)), reason clauses (cf. (25)) (Lord 1993: 159ff, cf. also Boadi 1972: 145):

---

[7] Something similar applies also for Kashmiri, another Indo-Aryan language, which exhibits verb-second (cf. Wali & Koul 1997: Ch.1.1).

The "main conjunction" *ki* in Hindi, which "seems to have been introduced in the system through Persian influence", can furthermore function as a "resulting marker" (i.e. *(so) that* in English), a temporal conjunction (in the sense of *when*), and can also constitute a subpart of various other clausal conjunctions (Montaut 2004: 243ff). As for relative clauses, although *ki* cannot play the role of a general relativizer in Hindi, the relative pronoun *jo*, if appositive, "is frequently reinforced by an 'expletive' *ki*," which signalizes, according to Montaut (2004: 248), the affinity of this kind of relatives with "the general pattern of the *ki* subordination."

[8] As a full verb, *sɛ* means 'be like', 'deserve', 'be necessary', etc. Its funciton is thus not limited to subordinating a clause. See Lord (1993: Ch.7.1).

[Twi]

(22)   ɔhwe [sɛ kofi beba anaa]

        he-look [Comp Kofi Fut-come or-not]

        'He ascertained whether Kofi would come.'

(23)   kofi yɛɛ adwuma no [sɛ yaw bɛpɛ n'asɛm]

        Kofi did work the [so-that Yaw Fut-like his-manner]

        'Kofi did the work so that Yaw would like him.'

(24)   [sɛ kofi yɛ adwuma no â] metua no ka

        [Comp Kofi do work the Cond] I-Fut-pay him salary

        'If/When Kofi does the work, I will pay him.'

(25)   oguanee [sɛ/efisɛ osuro]

        he-ran-away [Comp he-was-afraid]

        'He ran away because he was afraid.'

In (25), the verb *fi*, meaning 'come from', is combined with the impersonal pronoun *e-* and the whole subordinator *efisɛ* literally corresponds to 'it comes-from that'. Although *sɛ* alone can introduce reason clauses, the use of *efisɛ* makes this reason-reading more explicit, ruling out possible confusion with other readings of *sɛ*. In addition to *efisɛ*, there are some other "more-or-less frozen combinations" (Lord 1993: 172): "The complementizer *sɛ* has merged with other morphemes to form adverbial subordinating conjunctions."[9] In view of the observation that *sɛ* thus has a potential of subordinating a clause of various types, we can regard this clause-initial subordinator as a general clause linker in this language. Essentially the same thing applies, according to Lord (1973, 1993: Ch.7.4), to the clause-initial complementizer *bé* in Ewe, another Kwa language spoken in Ghana, Togo and Benin.

       There are still other languages, not necessarily related among each other, that exhibit essentially the same characteristics. Langacker (1975: 47) reports something to this effect for the clause-initial subordinator *in* in Classic Nahuatl, an SOV-language belonging to the Uto-Aztecan language family and spoken at the time of the 16th-century Spanish conquest of Mexico: "[...] a wide variety of subordinate clause types occur with *in*, including complement clauses, adverbial clauses, embedded questions, cleft sentences, and (according to [his] analysis) relative clauses, both restrictive and appositive." Givón (1991) presents ample

---

[9] The examples listed include *besi sɛ* 'until', *kãnsɛ* 'although', *gye sɛ* 'unless', *anasɛ* 'or', and *sɛnea* 'how', all of which are followed by a clause (Lord 1993: 172ff).

examples with the clause-initial subordinator *'asher* in Biblical Hebrew (later *she* as a result of phonological contraction); *'asher* (or *she*) can introduce complement clauses to a verb (p.263f, 277), relative clauses (p.263f), and several types of adverbial clauses (p.264). Essentially the same is true for "the general complementizer" *hogy* ('that') that appears clause-initially in Hungarian (Kenesei 1994).

I have thus listed a couple of languages each with a clause-initial subordination marker that serves the purpose of subordinating a clause in general. This observation motivates us to postulate a functional category Comp as a clause linker in these languages. Traditionally, the category Comp has been adopted for elements that can serve the purpose of linking a subordinate clause to the superordinate one or to some element within it.[10] It is, however, not self-evident whether all sorts of subordination markers, specifically for the clause-final ones, have the property of a general clause linker in the sense mentioned above. In the remaining part of this section, immediately below, we are confronted with this problem.

As already observed, Bengali has both an initial and a final subordinator, *je* and *bole*, respectively. Relevant examples are repeated below:

[Bengali]
(26)   chele-Ta Sune-che [je [or baba aS-be]]   (= (7a))
         boy-CF hear-Pst3 [Comp [his father come-Fut3]]
         'The boy has heard that his father will come.'
(27)   chele-Ta [[or baba aS-be] bole] Sune-che   (= (8a))
         boy-CF [[his father come Fut3] Comp] hear-Pst3
         'The boy has heard that his father will come.'

Bayer (2001: sec.2) claims that these two subordinators are to be differently characterized not only lexically or distributionally, but also syntactically. The following data from Bayer (2001: 15) show that the *bole*-clause can be used only in a subset of the cases in which the *je*-clause appears (cf. also Singh 1980: 192):

---

[10] Originally, Rosenbaum (1967: 24) called "complementizers" such "complementizing morphemes" as *to* in infinitives, the possessive marker *'s* and *-ing* in gerunds as well as *that* and *for*. For Bresnan (1979: 6), the complementizers are "those S-initial morphemes which distinguish clause types", including "*that, for, than, as*, and *WH*, or 'Q'." It seems that the term "complementizer" has in the meantime taken on a more theoretical character, in the sense that it "heads" a subordinate clause (cf. Chomsky 1986). See also fn.1.

[Bengali]

(28)   {janlam / bhablam / Sunlam / dekhlam / OSombhob} [je ram kolkata-y jacche]

        {knew-I / thought-I / heard-I / saw-I / unlikely} [Comp Ram Calcutta-Loc goes]

        '{I knew / I thought / I heard / I saw / It is unlikely} that Ram is going to Culcutta.'

(29)   [ram kolkata-y jacche <u>bole</u>] {janlam / bhablam / Sunlam / *dekhlam / *OSombhob}

        [Ram Calcutta-Loc goes <u>Comp</u>] {knew-I / thought-I / heard-I / *saw-I / *unlikely}

        '{I knew / I thought / I heard / *I saw / *It is unlikely} that Ram is going to Culcutta.'

As its traditional name "quotative" suggests, *bole* functions primarily "to set the preceding discourse in quotes" (Bayer 1999: 236, 2001: 13) or to signal it as "direct report" (Wurff 2002: 131). Let us look at further examples from Bayer (2001: 16):

[Bengali]

(30)   [colo colo <u>bole</u>] Se taRataRi hɔTte Suru korlo

        [go go <u>Comp</u>] (s)he quickly walk begin made

        'Saying "let's go", (s)he quickly began to walk.'

(31)   a. Se bollo: 'colo colo'

        (s)he said: 'go go'

        '(S)he said: "go, go!"'

        b. *Se bollo je colo colo

According to Bayer (2001: 16), *bole* in the relevant sense is always possible with direct speech, while *je* cannot select anything but a finite clause. The function of *bole* is, in other words, not restricted to subordinating a sentence, as opposed to *je*; the latter can also introduce a clause associated with a noun (Bayer 2001: 21):

[Bengali]

(32)   chele-Ta <u>e kOtha</u> jane na [<u>je</u> baba aS-be]

        boy-CL <u>this story</u> knows not [<u>Comp</u> father come-will]

        'The boy does not know it that his father will come.'

By way of comparison with the clause-initial subordinator *je*, which is, like *that* in English, to be considered a general complementizer in Bengali,[11] the dubious status of *bole* as a

---

[11] Bayer (1999: 238) further presents the following example from Assamese, another Indo-Aryan

complementizer thus becomes more conspicuous.[12]

        Let us next turn to another kind of preverbal sentential objects as represented below:

[Turkish] (Dryer 1980: 131)[13]

(33)   Adam ban-a [Ayşe-nin kitab-ı oku-duğ-u-nu] söyle-di.

        man me [Ayşe-Gen book-def-Acc read-Nom-3sg-Poss-Acc] told

        'The man told me that Ayşe read the book.'

[Persian] (Öhl 2003: 182)

(34)   Man [[in [ke gorbe-hā shir dust dārand]] rā] mi-dānam

        I [[Det [Comp cats milk friend have]] Acc] know

        'I know that cats like milk.'

[Japanese]

(35)   Gakusei-wa [[[sensei-ga yasumu] koto] wo] nozonde-ru.

        student-Top [[[teacher-Nom be-absent] thing] Acc] hope

        'The students are hoping that the professor will be absent.'

It should be noted here, however, that each embedded clause in the above examples is nominalized, which is also exemplified especially by Case-marking. The same is true also for Kanuri, an East Saharan language, where a Case-marker[14] plays the role of a subordinator in (36):

---

language closely related to Bengali:
i)   moi ne-janu [kak$_i$ je bill-e juwal kali t_i deshisil]
          I Neg-know [who$_i$ Comp Bill-Nom yesterday t_i seen-has]
          'I don't know who Bill saw yesterday.'
Bayer claims that *je* is not an operator but rather a complementizer while the operator *kak* moves into the specifier position, SpecCP, headed by the *je*. The argumentation here provides indirect support for the complementizer-status of *je* in question.

[12] Bayer (1999: 248) points to the possibility that *bole* is actually a postposition. This view does not seem to be quite out of the mark but rather is compatible with the observation that (i) PPs in Bangali are postpositional, (ii) there are not rarely cases in which complementizers and adpositions have something in common (cf. Emonds 1985, Corver & Riemsdijk 2001, Baker 2003, etc.), (iii) there are languages in which the same morpheme used as a postposition or a case marker functions as a subordinator (cf. e.g. Genetti 1991).

[13] See also Noonan (2007: 96) for further similar and relevant data from Uzbek.

[14] One might find cases across languages where a subordinator is actually a Case-marker rather than a complementizer in the sense discussed here. Also for English, Haumann (1997: Ch.6) claims that *for* is not a complementizer, as is taken for granted in the current generative literature, but a Case-assigner whose occurrence is required only by the presence of the DP in need of Case.

[Kanuri] (Noonan 2007: 57)

(36)   a. Àvá-nzə́-yè shí-<u>rò</u> kúŋə̀nà cîn.

        father-his-Nom him-<u>Dat</u> money give-3.sg.

        'His father gives him money.'

        b. [Sávà-nyí íshìn-<u>rò</u>] təmăŋə̀nà.

        [friend-my comes-<u>Dat</u>] thought-1.sg.-Perf

        'I thought my friend would come.'

From the standpoint of complementation by the matrix verb, what is selected as the complement in these examples is a nominal (NP/DP or KP (Kase phrase; cf. Bayer et al. 2001)) which contains a sentence within it. The data here thus represent cases not of CP-complementation but of NP-complementation by the matrix predicate. That is, the constituent selected by the matrix verb is not headed by a complementizer that subordinates a clause. Thus, the clause-final elements marking subordination are, here again, not complementizers in the strict sense.

In (35) above, we have just observed a case of preverbal sentential objects in Japanese whereby the embedded clause is nominalized with the help of *koto*. We would now like to review another strategy of subordinating a clause in Japanese, namely the use of *to* as appears in (12a), here repeated as (37):

[Japanese]

(37)   boku-wa [kare-ga kuru <u>to</u>] omou (= (12a))

        I-Top [he-Nom come <u>Comp</u>] think

        'I think that he will come.'

The *to* here is usually considered a complementizer in most of the literature without discussion. It is true that the function of *to* in sentences like (37) corresponds to that of the complementizer *that* in English. A closer look at it, however, renders this naïve view untenable or at least not self-evident. While I leave a profound investigation of this subordination marker in Japanese for a separate research, I just point out here that the remarks to the same effect as for *bole* above apply also for *to*: (i) The distribution of *to* as a clause linker is quite limited in Japanese:[15]

---

[15] In (29) above, we observed that the *bole*-clause in Bengali can be embedded by the verb 'hear' but not by 'see'. It might be interesting to note at this point that virtually the same applies to *to*-clauses in

[Japanese]

(38)   a. Okaasan-wa [[Mariko-ga keeki-o tabeta] {koto wo / (?)to}] sitte-iru.

        mother-Top [[Mariko-Nom cake-Acc ate] {thing Acc / (?)Comp}] know

        'The mother knows that Mariko ate the cake.'

    b. Okaasan-wa [[Mariko-ga keeki-o tabeta] {koto wo / ?*to}] sir-anai.

        mother-Top [[Mariko-Nom cake-Acc ate] {thing Acc / ?*Comp}] know-Neg

        'The mother does not know that Mariko ate the cake.'

(39)   Gakusei-wa [[sensei-ga yasumu] {koto wo / ?*to}] nozonde-iru. (cf. (36))

        student-Top [[teacher-Nom be-absent] {thing Acc / ?*Comp}] hope

        'The students are hoping that the professor will be absent.'

In some of the contexts in which *that* can appear as a subordination marker in English, *to* is excluded in Japanese; the object clauses here must be nominalized with the help of *koto*. The data reveal that *to* behaves rather idiosyncratically as a clause linker. Consequently, we can say that *to* can no more be regarded as a general complementizer than the Bengali *bole* (cf. (28) and (29)).

    (ii) Unlike *that* in English as in (19), *to* cannot function as a subordinator for the purpose of associating a clause with a nominal head:

[Japanese]

(40)   a. [kanojo-ga kekkon-shita {*to/to-iu}] uwasa

        [she-Nom marriage-did {*to/to-iu}] rumor

        'the rumor that she got married'

    b. [heiwa-ga otozureru {*to/to-iu}] kiboo

        [peace-Nom come-over {*to/to-iu}] hope

        'the hope that there will be peace'

Japanese:
i)   Mariko-wa [okaasan-ga modotte-kuru to] kiita.
        Mariko-Top [mother-Nom come-back Comp] heard
        'Mariko heard that her mother will come back.'
ii)   *Mariko-wa [okaasan-ga modotte-kuru to] mita.
        Mariko-Top [mother-Nom come-back Comp] saw
        'Mariko saw that her mother was coming back.'
These data lend support to the view that *bole* and *to* are actually quote markers. The different behaviour between 'hear' and 'see' is supposedly attributed to the unproblematic assumption that what is heard is suitable to be directly quoted in words, while what is seen is not necessarily so.

What relates a subordinate clause with a nominal is *to-iu*, which consists of the subordination marker *to* in question and the verb 'say' in its attributive form (for discussion on these two "complementizers" in Japanese, see Ogawa 2001: Ch.3.6.4, where *to-iu* is claimed to be a "verbal complementizer"). (40) thus has the structure [NP [VP [S *to*] *iu*]] N] (literally 'the N that says S') rather than [NP [CP S [C *to-iu*]] N]. The *to* (or any other element in Japanese) corresponding functionally to the English *that* does not thus seem to possess the complementizer-status comparable to the latter.

Another example of a clause-final subordination marker comes from Hopi (Voegelin & Voegelin 1975: 385f):

[Hopi]

(41)   pam [nɨy maːmac-<u>qa</u>-y] ʔinɨmi paŋquawɨ

she [me recognize-<u>Comp</u>-Obl] to-me tell

'She told me she recognized me.'

As Voegelin & Voegelin (1975) name it, *-qa* in Hopi is actually rather a nominalizer.[16] This is also evident by the fact that the embedded clause in (41) is Case-marked. Furthermore, *-qa* can also be attached even to a single verb (p.381):

[Hopi]

(42)   yaw [moːtiti-<u>qa</u>] yatkina-t maqiwni

Quot [be-first-<u>Comp</u>] saddle-Obl will-be-given

'The winner will be given a saddle.'

The sentential object selected by the matrix verb as in (41) is thus not a CP but a nominal, which is also detected by Case marking. This is just reminiscent of cases in some other languages observed above.

Before concluding this section, let me briefly touch on one aspect of relative clause formation, another kind of clausal subordination along with sentential complementation, in languages with prenominal relative clauses. As is expected (cf. e.g. Hawkins 1990, 1994), prenominal relative clauses have a final subordination marker, if at all.[17] Examples are from

---

[16] The nominalizer *-qa* appears also as a final "relativizer" in the sequence [NP N Case$_i$ S-qa Case$_i$].

[17] Data contradicting this generalization are found in Masica (1991: 412ff) from various New

Lahu and Chinese (Lehmann 1984: 62ff, Hawkins 1994: 388f):

[Lahu]

(43)   [và^ʔ-ó+qō thà^ʔ cɔ̂ tā <u>ve</u>] yâ+mî+ma

      [pig's-head Acc cook Perf <u>Comp</u>] woman

      'the woman who has cooked the pig's head'

[Chinese]

(44)   [nǐ gěi wǒ <u>de</u>] shū

      [you give I <u>Comp</u>] book

      'the book that you gave me'

The use of these relativization markers, *ve* and *de*, is not, however, limited to subordinating a clause in each language. They can appear in the final position also of a non-clausal constituent that modifies a noun:

[Lahu]

(45)   a. [dà^ʔ jâ <u>ve</u>] ɔ+lî

      [very nice <u>Comp</u>] custom

      'a very nice custom'

      b. [qhâ^ʔ+šɛ <u>ve</u>] ɔ́+qā

      [chieftain <u>Comp</u>] buffalo

      'the chieftain's buffalo'

[Chinese]

(46)   a. [bù hǎo <u>de</u>] lái-wǎng

      [Neg good <u>Comp</u>] come-go

      'non-good contact'

      b. [gè rén <u>de</u>] xīn

      [Cl man <u>Comp</u>] heart

      'each man's heart'

Relativization markers that appear in the final position of a prenominal relative clause should thus be regarded as an "attributor" (cf. Lehmann 1984) that serves the purpose of modifying a

Indo-Aryan languages; prenominal relative clauses with initial relative markers. Masica (1991: 413f) calls this type of constructions "a typological anomaly".

noun rather than subordinating a clause.[18] This exhibits a clear contrast to sentence-initial relative markers such as *that* in English whose function is limited to subordination of a finite clause.[19] Here we find the same kind of difference between clause-initial and clause-final subordinators as we observed in the case of sentential complementation.

We have in this section established on the basis of empirical data that there is a substantial difference between the clause-initial and the clause-final subordination markers.[20] The former can have properties of a general complementizer subordinating a wider range of clauses, while the latter do not;[21] clause-final subordinators are in one case nominalizers, whereby postpositions and case-markers are included, and in another quote-markers, whose function is not limited to attaching to a clause. In either case, final subordinators cannot be regarded as a clause-linking complementizer, at least in the sense comparable to the clause-initial ones. When we take the complementizer in our strict sense, we might then as well say that there exist no clause-final complementizers universally.

---

[18] Lehmann (1984: 61) calls the Lahu *ve* "nominalizer" ('Nominalisator') and characterizes it as a "universal particle" ('Universalpartikel') due to its "outstanding polyfunctionality" ('außerordentliche Polyfunktionalität'). Just as one of its functions, *ve* can appear also at the end of the sentential object, followed by case marking:

i)   [ɔ̀+šɨ̄ tɔ̂ʔ la <u>ve</u>] thàʔ nɔ̀ mâ ga mɔ̀ lâ

        [blood out come <u>Comp</u>] Acc you Neg Part see Q

        'Didn't you see that blood came out?'

[19] In English, there are marginally relative clauses that are infinite:

i)   She found a good place [from which to watch the procession]. (Huddleston & Pullum 2002: 1036)

In such a case, *that* is excluded as a relativization marker. The generalization that the general complementizer *that* is restricted to finite clauses thus still obtains.

[20] Actually, remarks to this effect can be found in the literature. Kayne (2005: 240) maintains, for example, "that (with few exceptions) there can be no 'final' counterpart of *that*, and that in languages that have only final complementizers there is likely to be no counterpart of English *that* at all." Davison (2007: 179) also asserts for a variety of Indic languages that "[...] the initial and final markers are lexically distinct, both in content (semantic features) and also in category." (Cf. also Newmeyer 2005: 56f.) As far as I see, however, these suggestions are neither thoroughly worked-out nor further investigated in the current research.

[21] Basque, again, will be a counter-example (cf. Lehmann 1984: 59, Hawkins 1994: 389f):

i)   [ama-k erra d-u-<u>en</u>] libura-a

        [mother(-Def)-Erg burnt Abs-Pres(-Erg)-<u>Comp</u>] book-Def

        'the book that mother has burned'

ii)   Esa-ida-zu [zer ari z-era-<u>n</u>] (= (16))

        tell-Imp-Dat-Erg [what do Abs-Aux-<u>Comp</u>]

        'Tell me what you are (in the process of) doing.'

The clause-final subordinator here seems to behave as a general complementizer.

4. An account

We have thus far established a certain asymmetry between initial and final subordinators that is found systematically across languages: Among the clause-initial subordinators there are genuine complementizers that serve the function of subordinating or "complementizing" exclusively clauses and can head the complementizer phrase, CP, which is directly selected by the matrix predicate. One of the most familiar examples of such a general complementizer is *that* in English. The clause-final subordinators, however, lack this property; they are either nominalizers or quote particles whose function is not limited to linking a clause. In this section, an attempt will be made to give an account for this asymmetry observed as a cross-linguistic tendency. Specifically, I will draw on a strategy based on parsing à la Hawkins (1994). The crucial point is that the sentence-processing takes place uniformly from left to right, irrespective of whether the language at issue is head-initial or head-final.

Let us remember the schemata (18), which result from the generalization (9), for the languages with initial subordinators and those with final subordinators, respectively (repeated below as (47)):

(47)   a. V_1 [Comp S_2]
        b. [S_2 Comp] V_1

They can now be represented more concretely as follows:

(48) $Subj_1$ V_1 (IO_1/Adv_1) <u>Comp</u> $Subj_2$ V_2 Obj_2
(49) $Subj_1$ (IO_1/Adv_1) $Subj_2$ Obj_2 V_2 <u>Comp</u> V_1

Consider first the case of subordinator-initial languages, (48). It should be noticed that in the course of sentence-processing taking place from left to right, the Comp in (48), at the point when it is encountered, serves the function of signalizing the existence of the subordinate clause: By coming across the Comp, the speaker immediately recognizes that the constituents that follow belong not to the matrix but to the lower clause (cf. also Hawkins 1994: Ch.5.6). In this sense, the Comp here has its raison d'être as an element subordinating a *clause*. Put differently, the configuration given as (48) renders it beneficial for effective parsing that the Comp plays the role of a clausal subordinator.

Let us next turn to (49), the embedding configuration in the subordinator-final languages. Here, the left-most boundary of the embedded clause lies immediately to the left of

Subj$_2$. Now, when the parser reaches the Subj$_2$, after going through Subj$_1$ (and IO$_1$/Adv$_1$), it is at this point per se not self-evident whether the Subj$_2$ should be interpreted as a matrix or an embedded element. This is also the case, or rather applies more explicitly, for Obj$_2$; the recognition of Obj$_2$ as embedded will be furthermore harder in case either Subj$_1$ or Subj$_2$ is phonologically not realized, as is not seldom the case in pro-drop languages. The presence of the subordination marker at the right periphery of the subordinate clause, as in (49), thus does not contribute to the on-line identification of the preceding elements as embedded. When the subordinator is unable to signalize the existence of sentential embedding "in due time", namely at the point when the relevant elements are processed, its function as a sentential subordinator is not necessarily justified. Simply said, the Comp in (49) does not have to be a sentential subordinator, but is rather, if at all, supposed to link preceding elements in general, and not exclusively a clause, to the following matrix element, here V$_1$.

The reasoning here regarding the construction of type (49) is just what our empirical research in the previous section has revealed. In one case, the Comp in (49) is a nominalizer or a Case marker that serves the function of converting the preceding constituent, of whichever syntactic category, into an element of a nominal character (cf. e.g. (33)-(36)). Now a nominal can be regarded as the canonical syntactic category of an argument that the matrix predicate at issue here selects. In this sense, the nominalizer or the Comp in (49) surely links the preceding elements to the following matrix element, the former not necessarily being a clause.

In another case, the clause-final subordinator or the Comp in (49) was identified as a quote marker (cf. e.g. (12a), (27), (30)). Here, too, the preceding element is not restricted to be a clause, but can be just anything, so long as it can be uttered. Be that as it may, the subordinator here at least links the quoted part to the matrix predicate.

In this section, I provided an account for the certain asymmetry observed between the clause-initial and the clause-final subordination markers across languages. The empirical observation, established in section 3, that the final subordinators (cf. (47b)) do not have the properties of a general complementizer subordinating a clause can now be attributed to their very positioning in the right periphery of the relevant subordinate clause: Final subordinators do not contribute to the on-line processing of the embedded elements as such, but merely serve the function of connecting the preceding elements in general, and not exclusively clauses, to the matrix predicate (see (49)). Simply put, it need not be a *clause* linker.

5. Concluding remarks

In this paper, I have discussed clause-initial and clause-final subordination markers, respectively, from a typological perspective. The initial and the final subordinators differ from each other not only in their positioning within the embedded clause, but also in their syntactic properties: The former typically possess the function of subordinating primarily clauses and can thus be regarded as belonging to the functional category Comp in the sense assumed in the current theoretical literature, while it is not the case with the latter ones. I have then tried to deduce this typologically observed tendency, a right-left-asymmetry, from the mechanism of sentence processing which proceeds from left to right.

I admit that the present research leaves much still to be desired. For example, a more thorough examination of data, both quantatively and qualitatively, would be desirable in order to reinforce the cross-linguistic generalizations. Crucially, the parsing-based account for the observed data carried out in section 4 is at present not much more than a working hypothesis, whose validity should be reassessed on the basis of a wider range of phenomena from a variety of languages. These and other problems I want to pursue in my future research.

References

An, D.-H. (2007) "Clauses in the Noncanonical Positions at the Syntax-Phonology Interface." *Syntax* 10: 38-79.

Bains, G. (1989) *Complex Structures in Hindi-Urdu: Explorations in Government and Binding Theory*. Ph.D. Diss. New York University.

Baker, M. (2003) *Lexical Categories. Verbs, Nouns, and Adjectives*. Cambridge: Cambridge UP.

Bayer, J. 1996. *Directionality and Logical Form. On the Scope of Focusing Particles and Wh-in-situ*. Dordrecht: Kluwer.

Bayer, J. (1999) "Final Complementizers in Hybrid Languages." *Journal of Linguistics* 35: 233-271.

Bayer, J. (2001) "Two Grammars in One: Sentential Complements and Complementizers in Bengali and Other South-Asian Languages." In: P. Bhaskararao & K. Subbarao (eds.) *The Yearbook of South Asian Languages 2001*. New Delhi: Sage Publications: 11-36.

Bayer, J., M. Bader & M. Meng (2001) "Morphological Underspecification Meets Oblique Case: Syntactic and Processing Effects in German." *Lingua* 111: 465-514.

Boadi, L. (1972) "Sentential Complements in Akan." *Lingua* 29: 128-172.

Bresnan, J. (1979) *Theory of Complementation in English Syntax*. New York & London: Garland.

Chomsky, N. (1981) *Lectures on Government and Binding. The Pisa Lectures*. Berlin & New York: de Gruyter.

Chomsky, N. (1986) *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, N. (2008) "On Phases." in: R. Freidin, C. Otero & M.-L. Zubizarreta (eds.) *Foundational Issues in Linguistic Theory*. Cambridge, Mass.: MIT Press: 133-166.

Corver, N. & H. v. Riemsdijk (eds.) (2001) *Semi-lexical Categories. The Funciton of Content Words and the Content of Function Words*. Berlin & New York: de Gruyter.

Davison, A. (2007) "Word Order, Parameters and the Extended COMP Projection." In: J. Bayer, T. Bhattacharya & H. Babu (eds.) *Linguistic Theory and South Asian Languages*. Amsterdam: John Benjamins: 175-198.

Dayal, V. & A. Mahajan (eds.) (2004) *Clause Structure in South Asian Languages*. Boston. Kluwer.

Dryer, M. (1980) "The Positional Tendencies of Sentential Noun Phrases in Universal Grammar." *The Canadian Journal of Linguistics* 25-2: 123-195.

Dryer, M. (2007) "Word Order." in: T. Shopen (ed.) *Language Typology and Syntactic Description. Vol.I: Clause Structure*. New York: Cambridge UP: 61-131.

Emonds, J. (1985) *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Gair, J. & J. Paolillo (1997) *Sinhala*. München: Lincom Europa.

Gelderen, E. (2004) *Grammaticalization as Economy*. Amsterdam & Philadelphia: John Benjamins.

Genetti, C. (1991) "From Postposition to Subordinator in Newari." in: Traugott & Heine (eds.): 227-255.

Givón, T. (1991) "The Evolution of Dependent Clause. Morpho-syntax in Biblical Hebrew." in: Traugott & Heine (eds.): 257-310.

Haumann, D. (1997) *The Syntax of Subordination*. Tübingen: Niemeyer.

Hawkins, J. (1990) "A Parsing Theory of Word Order Universals." *Linguistic Inquiry* 21: 223-261.

Hawkins, J. (1994) *A Performance Theory of Order and Constituency*. Cambridge: Cambridge UP.

Huddleston, R. & G. Pullum (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge UP.

Inaba, J. (2007) *Die Syntax der Satzkomplementierung: Zur Struktur des Nachfeldes im*

Deutschen. Berlin: Akademie Verlag.

Kayne, R. (2005) "Antisymmetry and Japanese." in: Kayne *Movement and Silence*. New York: Oxford UP: 215-240.

Kenesei, I. (1994) "Subordinate Clauses." in: F. Kiefer & K. Kiss (eds.) *The Syntactic Structure of Hungarian*. San Diego & Tokyo: Academic Press: 275-354.

Kuno, S. (1974) "The Position of Relative Clauses and Conjunctions." *Linguistic Inquiry* 5: 117-136.

Langacker, R. (1975) "Relative Clauses in Classic Nahuatl." *International Journal of American Linguistics* 41: 46-68.

Lehmann, C. (1984) *Der Relativsatz: Typologie seiner Strukturen; Theorie seiner Funktionen; Kompendium seiner Grammatik*. Tübingen: Narr.

Masica, C. (1991) *The Indo-Aryan Languages*. Cambridge & New York: Cambridge UP.

Matthews, P. (2007) *Syntactic Relations. A Critical Survey*. Cambridge. Cambridge UP.

Montaut, A. (2004) *A Grammar of Hindi*. München: Lincom Europa.

Newmeyer, F. (2005) *Possible and Probable Languages. A Generative Perspective on Linguistic Typology*. Oxford: Oxford UP.

Noonan, M. (2007) "Complementation." In: T. Shopen (ed.) *Language Typology and Syntactic Description. Vol. II. Complex Constructions*. Cambridge: Cambridge UP: 52-150.

Ogawa, Y. (2001) *A Unified Theory of Verbal and Nominal Projections*. Oxford & New York. Oxford UP.

Öhl, P. (2003) *Economical Computation of Structural Descriptions in Natural Language*. Ph.D. Diss. University of Stuttgart.

Pollard, C. & I. Sag (1994) *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Roberts, I. & A. Roussou (2003) *Syntactic Change. A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge UP.

Rosenbaum, P. (1967) *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: MIT Press.

Samushia, L. (2007) "Durch verba dicendi geschaffene hypotaktische Konstruktionen." Ms. University of Frankfurt am Main.

Singh, U. (1980) "*Bole*: An Unresolved Problem in Bengali Syntax." *Indian Linguistics* 41: 188-195.

Traugott, E. & B. Heine (eds.) (1991) *Approaches to Grammaticalization. Vol. II. Focus on Types of Grammatical Markers*. Amsterdam & Philadelphia: John Benjamins.

Voegelin, C. & F. Voegelin (1975) "Hopi /-qa/." *International Journal of American Linguistics* 41: 381-398.

Wali, K. & O. Koul (1997) *Kashmiri. A Cognitive-Descriptive Grammar*. London & New York: Routledge.

Wurff, W. (2002) "Direct, Indirect and Other Discourses in Bengali Newspapers." In: T. Güldemann & M. v. Roncador (eds.) *Reported Discourse. A Meeting Ground for Different Linguistic Domains*. Amsterdam & Philadelphia: John Benjamins: 121-139.