

Ambiguity-Avoidance: A Universal Constraint Extraction from NP Sequences

Author(s): Eve E. Sweetser

Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society (1977), pp. 440-453

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.

has reversed the usual S O IO to IO O S.

However, informational-weight ordering cannot reverse the SO order of a double accusative, even when reversal would not cause ambiguity:

- (4a) Ich habe den Freund den Wein holen lassen.
 (4b) *Ich habe den Wein den/einen Freund holen lassen.

b. Preposing can't apply to deep objects in DA sequences.

There exists in German a preposing rule which can move virtually any constituent to preverbal position for emphasis. (5 b-c) are examples of the application of this rule to (5a):

- (5a) Ich habe gestern meine Schwester gesehen.
 I saw my sister yesterday.
 (b) Meine Schwester habe ich gestern gesehen.
 (c) Gestern habe ich meine Schwester gesehen.

However, from double accusatives this rule can only prepose subjects:

- (6a) Den Freund habe ich den Wein holen lassen.
 (b) *Den Wein habe ich den Freund holen lassen.
 (c) Den Wein habe ich von dem/einem Freund holen lassen.

I let the wine be brought by a/the friend.

The deep object in a transitive clause embedded under lassen, hören, or sehen may only be preposed by passivizing the lower clause first, thus marking the deep subject with a von-phrase and making the deep object the subject of the lower clause.

2. Application of rules into ADA's.

There are no further constraints on movement from DA's unless extraction is a potential source of ambiguity—that is, there are no more constraints except on extraction from ambiguous double accusatives (ADA's). From ADA's only subjects can be relativized, and only objects can be questioned:

- (7) Der Freund, den ich meinen Bruder anrufen liess,..
 The friend that I had \emptyset call my brother...
 (*The friend that I had my brother call \emptyset ...)
 (8) Wen liessdest du deinen Bruder anrufen?
 Who(m) did you have your brother call \emptyset ?
 (*Who(m) did you have \emptyset call your brother?)

That is, in (7), the relativized noun must be interpreted as deep subject (not deep object) of the lower clause, while in (8), the questioned noun must be interpreted as deep object (not deep subject).

I examined two other German extraction rules, es-relativization and comparative deletion. Es-relativization is similar to English it-relativization,⁴ and produces sentences like (9):

- (9) Es war ein Polizist, der den Soldaten geschlagen hat.
It was a policeman who hit the soldier.

Comparative NP-deletion (not comparative ellipsis⁵) produces sentences like (10):

- (10) Er sah mehr Soldaten als sein Bruder \emptyset sah.
He saw more soldiers than his brother saw.

Es-relativization extracts only deep subjects from ADA's, and comparative deletion (CD) only deep objects: proofs of these two claims are here omitted for lack of space, but appear in Sweetser 1976. I shall now discuss tests of the definiteness of the targets of German extraction rules, so that these tests can then be applied to the four extraction rules in question.

3. Tests for definiteness of German NP's.

a. es gibt: a test of indefiniteness.

The German es gibt construction (used much like English "there is") can only be followed by an indefinite NP, not by a definite one.

- (11) Es gibt (*die) schöne Blumen in dieser Stadt.
There are (*the) beautiful flowers in this city.

b. Tests of definiteness.

There are three German constructions which only allow definite or generic NP's; non-generic NP's cannot occur in these constructions. First, as in English, only a definite or a generic NP may be subject of an adjectival predicate:

- (12)(a) Der Professor ist intelligent.
The professor is intelligent.
(b) Ein Professor ist intelligent. (*specific)
A professor is intelligent.

(12b) is fine under a generic reading where it means that professors are intelligent, but bad under a reading where it refers to a specific professor. "Tough" sentences in German (as in English) have this same restriction: only a definite or a generic NP can occur as the object of a clause embedded under a "tough" adjective, and hence only such NP's can be tough-moved.

- (13a) Mein Freund ist leicht zu erreichen.
My friend is easy to get hold of.
- (b) Ein Freund ist leicht zu erreichen. (*spec./
A friend is easy to get hold of. OK gener.)

Thirdly, selbst/selber intensive reflexives can only occur adjacent to definite NP's, never indefinites:

- (14) Der } Student selber { hat den Professor gefragt.
*Ein } { kann Homer lesen.
The/*A student_i himself_i { asked the professor.
{ can read Homer.

Selbst or selber can also occur separated from the NP to which it refers: when it so occurs, it can refer to either a definite or a generic NP, but not to a specific indefinite NP:

- (15a) Der Professor kann selber Homer lesen.
The professor_i can himself_i read Homer.
- (b) Ein Professor_i kann selber_i Homer lesen.
A professor_i can himself_i read Homer.
(*specific, ⁱOK generic)
- (c) Viele Studenten konnten selber Homer lesen.
Many students_i could themselves_i read Homer.
- (d) Ein Student hat den Professor_i selber gefragt?
*A student_i asked the professor himself_i.
(OK A student asked the professor_i himself_i.)

In (15d), a finite past-tense verb has made a generic interpretation of the subject impossible: the result is that the subject cannot be interpreted as the referent of selber.

4. Targets of German extraction rules.

The constructions discussed in the previous section constitute tests for the definiteness or indefiniteness of NP targets of extraction rules in German: e.g., if we find that the target of a given rule may not occur in the environments which bar definite NP's, and may occur in environments which require indefinite NP's, we know that that rule has an indefinite target. When applied, these tests showed that relativization and es-relativization have definite targets, while question-movement and CD have indefinite targets. There is space for only one of these proofs in the present paper: because it is fairly well-accepted that relativization and question-movement universally have definite and indefinite targets respectively, I shall present the proof for es-relativization (henceforth esrel). Esrel is similar to the English rule of it-relativization, which turns structures like "It_i (I ate X_i) was a banana" into structures like "It was a banana

that I ate." (The target of extraction is the relativized lower coreferent NP, which becomes the surface "that".) Similarly in German:

- (16) *Es waren (die) schönen Blumen, die es \emptyset in dieser Stadt gab.
It was (the) beautiful flowers that there were in this city.
- (17) Es war {ein Polizist, der {so klein war.
{mein Bruder} {schwer zu finden war.
It was {a policeman} who was {so small.
{my brother} {hard to find.
- (18) Es war ein/der Professor, der \emptyset selber Homer gelesen hatte.
It was a/the professor_i who himself_i had read Homer.

In (16), it can be seen that the es gibt construction, which permits only indefinite NP's, is not a possible environment for the target of esrel. (17)-(18), on the other hand, show that the target of esrel may be the subject of an adjectival predicate or of a tough-moved sentence, and may also be adjacent to a selber intensive reflexive: all of which are constructions which bar non-generic indefinites, while selber occurs adjacent only to definite NP's. We can thus conclude that only a definite NP may be the target of esrel.

Below is a summary of the results of the application of the definiteness tests to the targets of our four German extraction rules, and also of which rules extract subjects and which extract objects from ADA's.

Rule	Target	moves from ADA's
Relzn	def.	subj's only
<u>esrel</u>	def.	subj's only
Q-mvt.	indef.	obj's only
CD	indef.	obj's only

We have seen that extraction from German double accusative sequences obeys the proposed constraint. The question now is whether the constraint is of wider validity: does it also apply to extraction from fixed-order sequences of morphologically undifferentiated NP's (ADNP's) in other languages? The facts for extraction from Dutch double accusatives exactly parallel the German facts, and I shall now give other examples of the crosslinguistic applicability of the extraction-constraint.

5. Icelandic nominative-accusative sequences.

The usual Icelandic word-order is SVO, but preposing of the direct or indirect object may occur, causing the subject to be placed immediately following the verb.

Hence one possible Icelandic word order is IO V S DO:

- (19) Ég framseldi lögreglumni Jón. (basic order)
 I turned-over police-dat. John-acc.
 "I turned John over to the police."
 (20) Lögreglumni framseldi ég Jón.
 ((19) with preposed IO)

Some Icelandic nouns have nominative-accusative case syncretism (for indefinite NP's only). Hence, if verb morphology does not disambiguate (that is, if the two NP's are both of the same number and person), Icelandic postverbal S-DO sequences can be ADNP's, as in (21):

- (21) Kónginum framseldu lögreglumenn kennara.
 king-dat. turned-over policemen teachers
 3 pl. (nom/acc.) (nom/acc.)
 "Policemen turned over teachers to the king."

From ADNP's such as that in (21), relativization may extract only subjects and comparative deletion may extract only objects. (Question-movement is unrestricted; since there is no case-syncretism in the question-words, no ambiguity can be created by questioning from such a sequence.)

Icelandic intensive reflexives and adjectival predicates place exactly the same restrictions on the definiteness of NP's as do the corresponding constructions in German. Using these constructions as tests, it can be shown that Icelandic relativization has a definite target and CD an indefinite target.

Icelandic results:

Rule	Target	moves from ADNP's
Relzn	def.	subj's only
CD	indef.	obj's only

It can thus be seen that the extraction constraint holds in the case of these Icelandic ADNP's.

6. Tzotzil.

Tzotzil word order is fixed VOS, and nouns are not case-marked. Thus almost any Tzotzil sentence with animate subject and object (identical in person and number) contains an ADNP sequence, since only word order differentiates subject from object.

- (22) lah-s-mah zeb li vinik-e.¹¹
 past-3sgsubj-strike (the)-girl the man
 "The man struck the girl."
 (23) lah-s-melzan mesa li Petul-e.
 past-3sgs-make (the)-table Petul.
 "Petul made the table."

The question is how extraction rules will apply to NP's in Tzotzil ADNP sequences. And in fact, only subjects are relativizable in Tzotzil (not only from ADNP's, but under any circumstances- objects are simply unrelativizable), and only objects can be questioned (under any circumstances). And other extraction rules which are not normally restricted to subjects or to objects as targets, are so restricted in applying to NP's in ADNP sequences.

Tzotzil has a rule of clitic-attachment (or perhaps of pro-drop) which allows the non-appearance in the surface sentence of a definite third-person singular subject or object pronoun if no ambiguity arises from the omission:

- (24) lah-s-ve li kerem-e.
 past-3sgs-eat the boy
 "The boy ate it."

However, from ADNP's, only subjects can be cliticized:

- (25) lah-s-hip li kerem-e.
 past-3sgs-throw the boy
 "He threw the boy." (*"The boy threw it/him.")

The missing NP in (25) can only be interpreted as being the subject of the sentence, not the object.

Tzotzil also has a transformation of ?a-topicalization, which moves only definite NP's marked with the demonstrative li...e: these NP's are moved to sentence-initial position and marked with a topicalization-marker ?a. Either a subject or an object may be topicalized, if semantics or verb-morphology differentiates subject from object: but from an ADNP, only a subject may be topicalized:

- (26) lah-s-mah zeb li vinik-e.
 struck (the)girl the man
 "The man struck the girl."
 (27) ?a li vinik-e, lahsmah li zeb-e.
 "As for the man, he struck the girl."
 (28) ?a li zeb-e, lahsmah li vinik-e.
 *"As for the girl, the man struck her."
 "As for the girl, she struck the man."

Clitic-attachment and ?a-topicalization both have definite targets, since only definite pronouns can be cliticized and only definite demonstrative-marked NP's can be topicalized. The extraction-constraint holds for these rules, therefore, since they extract only subjects from ADNP's. I do not have the necessary information to construct definiteness-tests for Tzotzil, so I could not determine the nature of the targets of relativization and question-movement: however, it seems reasonable to

propose that these two transformations universally have a definite and an indefinite target respectively, and if that is so, then they also behave in accordance with the constraint on extraction.

Tzotzil results:		
Rule	Target	moves from ADNP's
clitic att.	def.	subj's. only
?a-top'zn.	def.	subj's. only
relzn.	def.	subj's. only
q-mvt.	indef.	obj's. only

An interesting corollary of the Tzotzil evidence just presented is that the constraint must be formulated to refer to extraction of subjects and objects, rather than of leftmost or nonleftmost NP's in a sequence: since Tzotzil is VOS, a constraint based on word order could not account for both the Germanic facts and the Tzotzil facts.

7. Navajo subject-object sequences.

The basic word order of Navajo is SOV, and in SOV sentences the verb is marked with a prefix yi-; but there is a variant OSV word order, and in OSV sentences the verb is marked with a prefix bi-. Navajo does not casemark nouns: as in Tzotzil, the order of subject and object NP's is crucial to distinguishing their syntactic functions. Hence in any sentence where semantics and verb morphology do not disambiguate, subject and object constitute an ADNP. Relativization (presumably definite-targeted) can extract only subjects from yi-marked sentences, and only objects from bi-marked sentences; and pronominalization (definite pronominalization, done by deletion) is restricted in the same way. This might a priori seem to argue against the formulation of the constraint which is necessary to account for the Tzotzil facts: it seems as if the Navajo constraint should be stated in terms of leftmost vs. nonleftmost NP rather than in terms of subject and object. However, it has been argued that the bi-marked sentences are passivized, and that the leftmost NP's in them are thus derived subjects.¹² If this is so, then Navajo constitutes further evidence for the constraint as stated, since definite-targeted extraction rules only extract subjects from ADNP sequences.

8. Conclusion: the universal basis for the constraint on extraction from ADNP's.

Examination of three Indo-European languages and two other totally unrelated ones indicates that the constraint on extraction from ADNP's may well be a universal one, and that crosslinguistically subjects must be definite and

objects indefinite to be extractable from such sequences.

It is known that there is a universal tendency to place old information (including definite NP's) leftwards in a sentence, and new information (including indefinite NP's) rightwards. Recent work by Edward Keenan has also established a universal correlation between subjecthood and definiteness. These two linguistic universals may well explain why subjects universally tend to precede objects, and why subject-initial word-order is common (while subject-final order is rare). Since relativization has a definite target, it seems logical that subjects, which tend to be definite, are also universally more relativizable than objects. Keenan has argued that this is the result of an NP-accessibility hierarchy, and that subjects are simply more extractable than objects - but the German facts show that subjects are only more extractable by definite-targeted transformations, rather than generally more accessible.

What I would like to argue is that when a language needs to develop a constraint to avoid ambiguity, it can freeze universal tendencies into requirements. In German, for example, extraction from an ADNP is impossible unless both of the word-order tendencies (definite before indefinite, and subject before object) are fulfilled. The NP target of an extraction from an ADNP must be a "classical" subject or object to be extractable. The hearer knows, when a definite NP has been extracted from such a sequence, that it was the deep subject and that it was the first of the two NP's in the sequence.

The universal constraint is independent of language-particular word order and is dependent on the subjecthood-definiteness correlation. Tzotzil is VOS and has the constraint in the same form as the four SO languages discussed. Since Tzotzil is an exception to the universal tendency for subjects to precede objects, it cannot require convergence of that tendency with the informational-weight ordering for extraction from ADNP's to be allowable. However, it seems natural that in German and other languages where a convergence of the two tendencies is possible, the order of ADNP's should follow both principles. Thus far, I have discovered no ADNP sequence whose fixed word order is not also the most natural order for those two constituents when they do not constitute an ADNP sequence. The two ordering principles taken together give the speaker an empirical reason to expect subjects and definites early in the sentence, and objects and indefinites late; they can thus hardly be unrelated to the universal constraint. Yet the Tzotzil use of the definiteness/subjecthood correlation shows that it in itself is a sufficient basis for the universal extraction-constraint.

The constraint which I have proposed would be hard to

state in most of the frameworks now in use for analyzing the structure of sentences. If stated in transformational terms, it is not merely a transderivational constraint, but a semantically based transderivational constraint, rather than a syntactically based one, such as that proposed in Hankamer 1973. For even though my constraint itself can be stated in syntactic terms, it applies only when semantic ambiguity would result from unconstrained extraction. That is, it is a functional constraint: it uses universal linguistic tendencies to construct a perceptual strategy for the univocal comprehension of what would otherwise be ambiguous sentences. Any meaningful formulation of this constraint must thus be able (1) to state the constraint in terms of grammatical relations rather than in terms of word order, and (2) to deal explicitly with the teleological nature of the constraint.

One final question is how my constraint fits into the dispute over a possible universal definition of subjecthood. Relational grammar has argued cogently for such a definition, yet there have been arguments against it, largely on the basis of so-called "subjectless" languages. It would be interesting to test my proposed constraint in such a language.

Footnotes

0. This paper is a highly condensed form of my undergraduate honors thesis, written in 1975-76 for the Harvard University Linguistics Dept. The German section of the thesis appeared in full in Harvard Studies in Syntax and Semantics, vol.2 (290r), spring 1976. The German work was made possible by six volunteer native informants, whom I thank. Two linguists, Annie Zaenen and Höskuldur Thráinsson were my native informants for Dutch and Icelandic respectively. The Navajo data comes from Ken Hale's handouts for an MIT course, explained to me by Phil LeSourd. The Tzotzil data is from Judith Aissen's field notes; I thank her for advice and criticism even more than for the data. And the most thanks of all to Jorge Hankamer, my adviser, sine qua non. In the thesis can be found the full proofs that belong in the gaps where this paper just asserts that proofs exist.

1. So far as I know, lassen, hören, and sehen are the only verbs in the German language which can have double accusatives in which both NP's are animate. Hence only with these three verbs do ADA's occur.

2. This generalization was first suggested to me by Jorge Hankamer.

3. The definiteness tests I use in this paper were suggested largely by Morgan's (1972) tests for definiteness of English NP's, which he used to prove that English relativization has

a definite target.

4. German es-relativization is a relativization rule, not a clefting rule; see note (9) below.

5. Comparative NP-deletion is separate from comparative ellipsis. Ellipsis, which gives rise to German sentences like "Er sah mehr Soldaten als sein Bruder \emptyset " (He saw more soldiers than his brother), is not an NP-extraction rule, and hence is not at issue in this paper.

6. English adjectival predicates and "tough" sentences show the same pattern of definite and generic NP's behaving alike as opposed to specific indefinite NP's: it is not at all peculiar to German.

7. The existence of the second, good reading of (15d) is irrelevant to the badness of the first reading. There exist German sentences where the reference of selber is ambiguous, e. g. Der Student hat den Professor selber gefragt, which means either "The student asked the professor, himself_i" or "The student_i asked the professor himself_i."

8. Proofs of the definiteness/indefiniteness of the targets of German relativization, question-movement, and CD appear in Sweetser 1976.

9. (16)-(18) also constitute a proof that es-relativization is in fact relativization, and not clefting. A clefting rule would derive a sentence such as "It was a banana that I ate" from a base-form such as "I ate a banana." The focal NP in a clefted sentence thus originated in the lower clause. However, the focal NP in an it-relativized sentence did not originate in the lower clause and has not been moved. Thus, if esrel were clefting, we would expect indefinite focal NP's to be good exactly when the hole in the lower clause could have contained an indefinite NP, and bad exactly when indefinites are barred in the target-position in the lower clause. However, we do not find this: for example, in (16), we see that neither die schönen Blumen (barred with the lower es gab) nor schönen Blumen is good; while in (17)-(18) we see that indefinite focal NP's are good even when the target-position bars indefinite NP's. We must then assume that what has been moved from below is not the focal NP, but the relative pronoun, and that we are not looking at clefting but at esrel. If esrel has a definite target, we can explain the grammaticality of (17)-(18) and the ungrammaticality of (16).

10. The Dutch and Icelandic facts exist in full in my B.A. thesis.

11. In Tzotzil, li...e around a noun serves to mark that NP as demonstrative. Li...e can only occur on definite NP's, and cannot occur on two adjacent NP's: in an OS sequence, it will thus appear only on the subject and be blocked from occurring on the object unless the subject is moved away.

12. The only argument I have heard against a passive analysis of bi-marked Navajo sentences is that this marking is not reserved for sentences where subject and DO have been reversed in order, but also occurs when other NP's (such as the objects of postpositional phrases) have switched places with subject NP's. However, if English passive can produce "This bed was slept in by George Washington," I don't see why Navajo passive should be rigidly held to applying to DO's only.

Bibliography

Aissen, Judith, unpublished Tzotzil field notes from San Cristobal de las Casas, Chiapas, Mexico (1974)

Breckenridge, Janet, Rules which nothing undergoes: an investigation of impersonal passives and object-raising constructions in German, Harvard honors thesis 1975.

Cowan, Marion, Tzotzil Grammar, Summer Institute of Linguistics 1969.

Hankamer, Jorge, "Unacceptable Ambiguity," Linguistic Inquiry vol. IV, no. I, winter 1973.

Hankamer, Jorge, and Laura Knecht, "The Role of the Subject/Nonsubject Distinction in Determining the Choice of Relative Clause Participle in Turkish," Harvard Studies in Syntax and Semantics, vol. II, 1976.

Hankamer, Jorge, and Jessie Pinkham, "Deep and Shallow Clefts," CLS XI, 1975.

Keenan, Edward, "NP Accessibility and Universal Grammar," 1972.

Mathesius, V., "On Information-Bearing Structure of the Sentence," trans. Tsuneko Olga Yokoyama, in Harvard Studies in Syntax and Semantics, vol. I, 1975.

Morgan, Jerry, "Some Aspects of Relative Clauses in English and Albanian," CLS VIII, 1972.

Milsark, Gary, On the Existential Predicate Construction in English, MIT PhD. thesis, 1974.

Reis, Marga, "Is there a rule of subject-to-object raising in German?" CLS IX, 1973.

Sgall, P., E. Hajičová, and E. Benešová, Topic, Focus, and Generative Semantics, Kronberg Taunus 1973.

Sweetser, Eve, "Constraints on Extraction from Ambiguous Double Noun Phrase Sequences," Harvard honors thesis 1976.

Sweetser, Eve, "Avoiding Ambiguity in German Double Accusatives," Harvard Studies in Syntax and Semantics, vol. II, 1976.

Thráinsson, Hóskuldur, "Gapping in Icelandic: a functional explanation of the No-Ambiguity Constraint," CLS XI, 1975.

Weathers, Nadine, "Morphological Analysis of a Tzotzil Text," IJAL, vol. XVI, pp. 91-8.