

Towards Understanding Raising-to-Subject

Scott Grimm

Dept. of Linguistics, Stanford University

sgrimm@stanford.edu

April 22, 2011

Raising to Subject: Background

Topicality

Evidentiality

Cross-Linguistic Correlates

Controlling for Other Factors

Implications: Scope

Raising to Subject: Background

A class of verbs and adjectives, including *seem*, *appear*, *likely*, take both sentential and infinitival complements

- (1) It seems that Barnett understands the formula.
(SentComp)
- (2) Barnett seems to understand the formula. (InfComp)

The standard view is that these two expressions are in some sense equivalent, either truth-conditionally (Davies and Dubinsky, 2004, p. 4) or via selectional restrictions

This talk will contribute new empirical evidence and isolate some of the factors that distinguish the two forms of expression (topicality and evidentiality).

Raising to Subject: Background

3 Common Assumptions:

- ▶ At the level of thematic selection, both InfComp and SentComp constructions select for a proposition
- ▶ The raising predicate does not select for its subject
(*Barnett isn't engaged in some act or state of 'seeming'*)

The subject is selected with respect to the proposition:

(3) Barnett_i [seems [t_i to understand]]

- ▶ *seem* can have wide-scope in both:

seem [Barnett to understand the formula]_{prop}

seem [Barnett understands the formula]_{prop}

Main Claim: The InfComp is strongly associated with a topic-comment structure, while the SentComp construction permits embedded subjects which are not topics.

This can be established by examining both constructions w.r.t. characteristics of topic-comment structures

- ▶ Topic tests
- ▶ Distribution of Information Status for subjects
- ▶ Acceptability of non-topics as subjects

Topicality: Topic Tests

Both the InfComp and SentComp pass topic tests, which evaluate for changes of meaning under topicalization (Reinhart 1981).

- (4) a. Felix seems to be back in town.
b. As for Felix, he seems to be back in town.
- (5) a. It seems that Felix is back in town.
b. As for Felix, it seems he is back in town.

⇒ Subjects of both the InfComp and SentComp are potential topics.

Do subjects of the InfComp and SentComp differ in information structure properties?

- ▶ If so, this should be reflected in the information status of the subjects found across a corpus.

British National Corpus: 200 tokens each of *seem* or *appear* with the InfComp or SentComp (800 tokens total)

Nissim et al. (2004) measures information status in terms of hearer identifiability, primarily following Prince (1992)

- ▶ *Old*: previously mentioned, pronouns, generics

6 Subtypes:

identity (co-reference with previously mentioned entity), event, general (dialogue participants), generic, ident-generic (co-referential with a generic entity), relative (relative pronouns).

- ▶ *New*: Not previously mentioned or accessible to the hearer.
No Subtypes

Mediated: “Mediated entities have not yet been directly introduced in the dialogue, but are inferrable from previously mentioned ones, or generally known to the hearer.”

9 Subtypes:

general (culturally known entities, “the moon”), bound (bound pronouns), part, situation, event, set, poss, function value, aggregation

Topicality: Information Status Distribution

The two construction differ in their distribution:

- ▶ InfComp subjects show reliable preference for old material
- ▶ SentComp subjects show greater acceptance of new material

	old	mediated	new
InfComp	234	124	13
SentComp	174	144	47

$$\chi^2 : p < .0001$$

Discourse-new subjects of the InfComp, while hearer new, were clearly speaker-identified—still qualify as topics, e.g. contrastive topics, specific indefinites

⇒ Subjects of InfComp are topical

Spoken Modality

Both to control for modality of production and as a means of independent verification, I examined the occurrences of *seem* in a version of the Switchboard Corpus (Godfrey et al. 1992)

- ▶ A portion was annotated for information status as part of the LINK project (based on Bresnan et al. (2002) and Zaenen et al. (2004)).
- ▶ Identical annotation scheme, so the comparison was straightforward

The SentComp construction had too few instances to be informative, but the InfComp instances displayed the same pattern and proportions as in the BNC

	old	mediated	new
InfComp	30	15	-

Topicality: Acceptability of Non-topics

Singular indefinite generics fail the topic test (Reinhart 1981).

- (6) A shark will never attack unless it is very hungry.
?He said about a shark that it will never attack unless it is very hungry.

If verbs such as *seem* only permit subjects which are topics, then indefinite generics should not be permitted

Generics: Background

Different generics have different interpretations:

- ▶ An inductivist use: true when “*sufficiently many relevant* individuals in the domain of the generic satisfy the predicated property.” (Cohen 2001, p. 194)

(7) Kings are generous.

- ▶ Normative generic sentences “do not get their truth or falsity as a consequence of properties of individual instances . . . instead, [they] are evaluated with regard to rules and regulations” (Cohen 2001, p. 194).

(8) Bishops move diagonally.

(9) A bishop moves diagonally.

Normative statements can come with a deontic force:

(10) A Christian is forgiving. (Papafragou 1996)

(11) Christians are forgiving.

(10) does not indicate that any Christian actually achieves this ideal, but that they *should*, in contrast to (11), which on the preferred reading does say something about the real world.

Generics: Background

- ▶ Bare plurals allow for both *inductivist* and *normative* (or *definitional*) readings
- ▶ Indefinite singular generics only permit a normative reading

(12) Kings are generous. (✓ induct. / ✓ norm.)

(13) A king is generous. (# induct. / ✓ norm.)

Topicality: Acceptability of Non-topics

The SentComp permits normative/definitional indefinite generic statements, and hence non-topic subjects:

(14) It seems that a bishop moves diagonally.

(15) It seems that a king is generous.

Embedded indefinite subjects are found in naturally-occurring SentComp constructions:

(16) It seems that a fetus is a precious life worthy of protection, but after birth, it's a harlot's kid and another welfare recipient.

(www.slate.com/id/2182590/)

Topicality: Acceptability of Non-topics

The InfComp, however, does not permit normative/definitional indefinite generic statements, and therefore does not permit non-topic subjects:

- (17) ?A bishop seems to move diagonally.
- (18) ?A king seems to be generous.

Topicality: Acceptability of Non-topics

A similar infelicity for InfComp constructions occurs for indefinite generics with a deontic reading

- (19) a. A Christian is forgiving.
b. ?A Christian seems to be forgiving.
- (20) a. A gentleman opens doors. (Burton-Roberts 1977)
b. ?A gentleman seems to open doors.

Contrast with:

while reading *The Complete Book of Etiquette*, Jones reports:

- (21) It seems that a gentleman opens doors.

Topicality: Acceptability of Non-topics

Singular indefinites can be made acceptable when they are either specific or contrastive—i.e., when they are made topical

- (22) What always strikes me about health food shops are the rows and rows of bottles and tablets. **A greengrocer** seems to be a much better source of natural products than such collections of distilled essences and the like.
(<http://www.fallacyfiles.org/adnature.html>)

Comparison: Passives and Unaccusatives

Other constructions analyzed as A-Movement do not show such constraints:

- ▶ Passive:

(23) A president is voted in by members of the country, company, or other entity. (Google)

- ▶ Unaccusative:

(24) An iceberg melts during the summer. (Google)

Nor are infinitive constructions problematic:

(25) The world wants a man to be financially effective. (Google)

seem appears to constrain possible subjects

In some sense, that the InfComp is associated with topicality is surprising, since idiom chunks, which are supposedly non-referential, are used as a diagnostic for raising structures

Idioms do however show traits of referential expressions, such as serving as antecedents for pronouns (Nunberg et al. 1994).

(26) (Nunberg et al. 1994, p. 502)

Once someone lets the cat out of the bag, it stays out of the bag for good.

If the claim that subjects of the InfComp are associated with topicality holds, it is expected that for idioms or fixed expression with truly non-referential subjects, infelicity would arise

- (27) a. A fool and his money are soon parted.
b. ?A fool and his money seem to be soon parted
- (28) a. A still tongue keeps a wise head.
b. ?A still tongue seems to keep a wise head.

The above infelicity is parallel to that of the indefinite generics

Interim Summary

On the above evidence:

- ▶ Subjects of the SentComp may be topics or not
- ▶ Subjects of the InfComp may only be topics

This evidence also implicates that the InfComp and SentComp cannot be truth conditionally equivalent in all cases

The InfComp aligns with a topic-comment structure (so-called “categorical” statements)

The SentComp is not so restricted

Evidentiality provides a second clear instance where the InfComp and SentComp differ

A perceptual experience constraint has been noted for the InfComp Postal (1974); Asudeh and Toivonen (2007).

(29) Julius Caesar appeared to be honest.

(30) It appears that Julius Caesar was honest.

This can be generalized to connect with the domain of evidentiality

This connection is increasingly recognized (Aijmer 2008; de Haan 2007)

Evidentiality: Background

- ▶ Direct evidential marking is used when “the speaker has some sort of sensory evidence for the action or event he/she is describing” (de Haan 2004).
 - ▶ Visual, auditory or other means of direct experience
- ▶ Indirect evidentials “are used when the speaker was not a witness to the event but when he/she learned of it after the fact” (ibid.).
 - ▶ This includes *inference* and *hearsay*.
- ▶ *Hypothesis*: The direct evidential function is associated with the InfComp; the indirect evidential function is associated with the SentComp

Back to the Corpus!

- ▶ Return to BNC corpus: 200 tokens each of *seem* or *appear* with the InfComp or SentComp (800 tokens total)
- ▶ As the direct evidential is the unmarked form, nearly impossible to code for it based on objective criteria (at this point)
- ▶ Indirect evidence (inference; hearsay) was often explicitly marked:

(31) It also seemed, from the feathers on the kitchen floor, that one of the pigeons had come down for a warm and had got too close . (BNC HTL)

Examples of Indirect Evidential Markers

evidential source: Provided that Ali's analysis of Hacıhasanzade's motives is correct-and one must remember that Ali is writing nearly a century after the event it would appear that already at the beginning of the sixteenth century the career of a kasabat kadi was regarded as a dead end .

for phrase: It would seem that in Capadocea they were even converted to Christianity for, in that region, St Christopher is often depicted with a dog's head.

- ▶ Markers: *from*-phrases, *therefore*, *then*, *thus* and explicit scientific conclusions

Distribution of Data

- ▶ 147/400 SentComp tokens were explicitly marked for inference or hearsay functions (37%)
- ▶ \Rightarrow SentComp is associated with an indirect evidential function

Evidentiality and Topicality

- ▶ The same acceptability patterns observed with indefinite generics by modulating construction type (InfComp vs. SentComp) can be obtained by modulating evidential type:

- (32) a. I saw that a king is generous. (inductive)
b. I heard that a king is generous.
(inductive/normative)

Implications: Lexicalization Patterns

- ▶ The association between evidential type and construction type makes a clear prediction as to lexicalization patterns of predicates which accept only one construction.
 - ▶ Predicates with direct evidential function should appear in the InfComp construction
 - ▶ Predicates with an indirect evidential function should appear in the SentComp construction
- ▶ Predicates such as *looks to* and *inferable* support this prediction

(33) a. It is inferable that Ed left.
b. *Ed is inferable to leave

(34) a. Ed looks to be tired.
b. *It looks that Ed is tired.

Implications: Lexicalization Patterns

The pattern extends beyond visual evidence:

- (35) a. I have changed the fuser hoping that was where the loud noise was coming from. No luck, it sounds to be coming from the back of the machine.
<http://forums13.itrc.hp.com/service/forums/questionanswer>
- b. *It sounds that it is coming from the back of the machine.

Cross-Linguistic Correlates: Italian

Italian shows the same empirical patterns w.r.t. evidentiality and topicality

- ▶ The sentential complement version has a strong implication of indirect evidence or inference

(36) Gianni sembra essere malato
John seems be.INF sick
John seems to be sick

(37) Sembra che Gianni sia malato
seems that John be.SUBJ sick
It seems that John is sick (Inference/Indirect Evidential)

Cross-Linguistic Correlates: Italian

- ▶ Similar patterns of infelicity also arise for generics:

(38) Sembra che un gentleman ceda il
It seems that a gentleman gives.SUBJ
passo a una signora
the way to a lady
It seems that a gentleman gives way to a lady

(39) Un gentleman sembra cedere il passo a una
A gentleman seems give.INF the way to a
signora
lady
A gentleman seems to give way to a lady
(✓induct. / # norm.)

Cross-Linguistic Correlates: Japanese

The above argued that construction choice is associated with evidential contrasts.

Japanese provides a clear example where evidential meaning depends on construction choice.

Cross-Linguistic Correlates: Japanese

The particle *soo* licenses two different interpretations: when modifying nouns or adjectives, *soo* denotes an apparent quality based on visual experience; when modifying clauses *soo* denotes reported speech

- (40) a. kono ringo wa oishi soo desu
this apple TOP delicious PART COP
This apple looks delicious.
- b. sono hi wa shugyo ga nai soo
this day TOP class NOM NEG.COP PART
desu
COP
I heard that today there is no class.

Cross-Linguistic Correlates: Japanese

Minimal pairs can be constructed showing that whether *soo* combines with an adjective or clause results in different interpretations.

- (41) a. takeshi no ringo wa oishi soo desu
takeshi POSS apple TOP delicious PART COP
Takeshi's apple looks delicious.
- b. takeshi no ringo wa oishi-i soo
takeshi POSS apple TOP delicious-INFL PART
desu
COP
I heard that Takeshi's apple is delicious.

Cross-Linguistic Correlates: Japanese

The identical situation holds of the particle *mitai*—when a nominal-adjective modifier, it indicates resemblance, while when modifying a clause, it marks inference based on visual evidence.

- (42) a. ano hito wa gorira mitai desu
that person TOP gorilla PART COP
That person looks like (resembles) a gorilla.
- b. sono hi wa shugyo ga nai mitai
this day TOP class NOM NEG.COP PART
desu
COP
It appears that today there is no class.

Controlling for other factors

- ▶ Multivariable regression models allow to control for other factors
- ▶ Subject Length (raw and log)
- ▶ Nominal Expression Type :
 - ▶ *pronoun* (including definite, personal and reflexive pronouns, as well as demonstratives)
 - ▶ *definite*
 - ▶ *indefinite* (including phrases with the indefinite article as well as bare plurals)
 - ▶ *proper name*
 - ▶ *quantifier* (such as *most*, *few*, *any*)
 - ▶ *relative pronoun*
 - ▶ *there*

Controlling for Other factors

- ▶ The more fine-grained information status subtypes yielded three significant factors which were associated with the SentComp construction:
 - ▶ *new* ($p \approx 0.002$)
 - ▶ *old-ident-generic* ($p \approx 0.003$)
 - ▶ *old-generic* ($p \approx 0.02$)
- ▶ The remaining significant factor was *indirect evidential* ($p < 2e-16$)

Implications: Scope

Topics are backgrounded and linked to the prior discourse

One would expect then that the subject of InfComp has wide scope

While there has been a long history of asserting “reconstructed readings” and quantifying-in, where the subject of the InfComp has narrow scope w.r.t. *seem*, this stance has come under scrutiny (Chomsky 1993, 1995; Lasnik 1998, 2003)

Once various confounds are controlled for, there are clear asymmetries which are consistent with the subjects of InfComp being topics

Scope Confounds

- ▶ What motivates the raising account or alternately the quantifier lowering account?
- ▶ As far as I can tell, two main sentence types.

Type 1:

(43) Some politician is likely to address John's constituency. (May 1977)

“may be taken as asserting either (i) that there is a politician, that is, Rockefeller, who is likely to address John's constituency, or (ii) that it is likely that there is some politician (or other) who will address John's constituency.”

Scope Confounds

- ▶ Is this a genuine scope phenomena?
- ▶ May's discussion only makes recourse to epistemic specificity (whether the referent is known to the speaker) not scopal specificity
- ▶ Have various confounds been controlled for?

Scope Confound: Entailments

- ▶ For scope inversion cases with two quantifiers, one reading entails the other (Ruys and Winter, Reinhart)
- ▶ Every man loves some woman:

$$(44) \quad \begin{array}{l} \text{a. } \forall x [\text{MAN}(x) \rightarrow \exists y [\text{WOMAN}(y) \wedge \text{LOVE}(y)(x)]] \\ \text{b. } \exists y [\text{WOMAN}(y) \wedge \forall x [\text{MAN}(x) \rightarrow \text{LOVE}(y)(x)]] \end{array}$$

- ▶ Whenever the wide-scope reading is true, so is the narrow scope reading.

Scope Confound: Entailments

- ▶ A similar relation holds between canonical examples of reconstruction, e.g. the sentences "A Canadian is likely to win the race" and "It is likely that a Canadian will win the race."
- ▶ Intuition: If a particular canadian is likely to win the race, then the situation where some Canadian or other wins is likely.
- ▶ (45) a. $\text{Likely}[\exists x (\text{Canadian}(x) \wedge \text{win}(x))]$
b. $\exists x (\text{Canadian}(x) \wedge \text{Likely}(\text{win}(x)))$

Scope Confounds: Intonation

- ▶ Hinterwimmer 2006 notes that focus plays a role:
 - (46) SOMEONE FROM NEW YORK is likely to win the lottery.
 - (47) Someone from New York is likely TO WIN THE LOTTERY.
- ▶ “Only focus marked constituents reconstruct.”
- ▶ However, this effect seems quite independent from reconstruction situations:
 - (48) SOMEONE knows the answer.
 - (49) Someone knows THE ANSWER.
- ▶ This is an area that needs to be explored systematically

Scope Confounds: Lexical Semantics

- ▶ Example Type 2:
- ▶ Montague (1970) and others have used examples as below to argue that raising constructions permit scopally non-specific (i.e. narrow-scope) readings of the subject.

(50) A cat seems to be in the garden. [existence]

(51) A train seems to be approaching. [appearance]

- ▶ Yet the most successful examples always use verbs of existence and appearance
- ▶ Such verbs are peculiar in both their lexical semantics and information structure properties (e.g., do not permit topic marker *wa* in Japanese)

Scope Confounds: Lexical Semantics

- ▶ When the phrases are minimally altered in the choice of the verb, the scopally non-specific reading becomes far less accessible, if not impossible, under normal intonational patterns

(52) A cat seems to be sleeping in the garden.

(53) A train seems to be leaving.

- ▶ The most natural readings involve a specific cat and a specific train.

- ▶ These examples can be made to have something resembling a narrow scope reading when focussed (mirative reading)

(54) A *CAT* seems to be in the garden.

Scope Confounds: Lexical Semantics

- ▶ A third type of verb that is popular is verbs of “negative existence”:

(55) A student seems to be absent.

- ▶ Yet these verbs result in intensional object readings in the absence of any raising verb (E. Bach reported in Dowty 1985):

(56) A screw is missing from this TV set.

- ▶ Again, a peculiar set of verbs (take the intensional genitive in Russian, etc.)
- ▶ Once all these confounds are controlled for, then the different forms show distinct interpretational differences.

Scope: Scopal Non-Specifics

- ▶ Scopally non-specific subjects are attested with the SentComp construction:

(57) It appeared that a German S.P. gun had joined the snipers and was lobbing the occasional shell into the vicinity of the orchard. (BNC A61)
⇒ there existed a German S.P. gun

- ▶ The subject of the InfComp version preferentially takes wide-scope

(58) A German S.P. gun appeared to have joined the snipers.

Scope: Quantifiers

- ▶ Partee (1971) noted that the InfComp and SentComp manifest interpretational differences when quantifiers are present

- (59)
- a. Few students are certain to pass the test.
(strong/existential)
 - b. It is certain that few students will pass the test.
(weak/proportional)

- ▶ (22a) has the strong reading, designates a set of students which the speaker has information about who did not pass : *few students* has scope over *certain*.
- ▶ (22b) has the proportional reading, i.e. few out of the students passed: *certain* has scope over *few students*.
- ▶ This is expected if the subjects of InfComp are more topical: refer to a specified set

Scope: Summary

- ▶ General scoping preferences:
 - ▶ InfComp: Subj > Predicate
 - ▶ SentComp: Predicate > Subj
- ▶ The preferential reading patterns are consistent with the subjects of the InfComp being topics, as topics are backgrounded, referential, and would refer outside of the context induced by the predicate

General Summary

- ▶ The InfComp and SentComp systematically diverge in the set of readings that they permit
- ▶ The readings which are not simple subject-predicate (topic-comment) relations are associated with the SentComp

	InfComp	SentComp
Sing. Generic Evidential Indefinite Subject	inductive direct evidential scopally specific	inductive, normative indirect evidential scopally non-specific/specific

- ▶ These divergent readings align well with the thesis that the subjects of InfComp are topics (e.g. subjects are specific/referential)

General implications for the raising analysis:

- ▶ It is improbable that *seem*, etc. select for propositions in a simple or unified manner—for not just any proposition is felicitous in the InfComp formulation
- ▶ Raising verbs in the InfComp do select for their subject, although they do not discriminate in terms of thematic content: such verbs, at minimum, select for topics, and for *seems* and *appear*, the subjects of the InfComp are constrained to be direct evidential sources

Thank you

Thanks to the following for discussion (and challenges): Eve Clark, Cleo Condoravi, Hans-Martin Gärtner, Beth Levin, Asya Pereltsvaig, Masha Polinsky, Ivan Sag, Barbara Stiebels, Tom Wasow
and to Fabio Del Prete and Chigusa Kurumada for help with the Italian and Japanese data