Copy raising as a lexical rule

Paul Kay

University of California, Berkeley and Stanford University

Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar

Online (Frankfurt/Main)
Stefan Müller, Nurit Melnik (Editors)
2021
Frankfurt/Main: University Library

pages 68–82

Keywords: copy raising, lexical rule, SBCG, weather it, SELECT feature

Abstract

The paper argues that there is compelling evidence for analyzing copy raising in English as a lexical rule that converts a subtype of perception verb with a stimulus subject (a so-called “flip-perception verb”) into a semantically bleached verb of mild evidentiary force, roughly equivalent to seem in some uses, which identifies the index of its external argument with the index of the pronominally expressed external argument of its complement.

1. Introduction

Copy raising in English is best viewed as a lexical rule that converts a subtype of perception verb with a stimulus subject (a so-called “flip-perception verb”) into a semantically bleached verb of mild evidentiary force, roughly equivalent semantically to seem in some uses. The derived verb does not subcategorize for a source of perception. It provides for a generic interpretation of the unexpressed witness of the evidence, and it identifies the index of its external argument with the index of the pronominal external argument of its clausal complement. Sign-Based Construction Grammar (SBGC) provides a formal framework for expressing this analysis.

Following Kim 2014, Landau 2011, and Potsdam & Runner 2001, among others, we observe that certain English perception verbs have two distinct senses: (i) a perception sense, in which the verb’s external argument denotes an entity that plays a semantic role which might be called the ‘perceptual stimulus’ in a report of an experience in the perceptual mode denoted by the verb, and (ii) a different sense (or a homophonous verb) that does not assign a semantic role to its external argument and which figures in sentences like (1-4).

(1) ... it’s not difficult to work out why Trump looked like he was going to win in January: the stock market was booming, unemployment was low, crime low, there were no new wars…it’s not a mystery.

(2) i have gone ahead and paid y our parts because the host of the giveaway looks like he disappeared …

1 The verbs with these two senses are not necessarily limited to those senses. For example, seem and appear may have additional senses associated with their raising and it-extraposition valences and sound has a hearsay sense distinct from both its perception and copy raising senses (i).

(i) This 74-year-old pasta sauce recipe sounds incredibly delicious.
(3) ... the bill seemed like it would easily pass...

(4) I'm so excited to get my copy of Sara's book ... this one sounds like it's going to win an immediate place in my heart.

In (1) Trump’s visual appearance is patently not at issue. Similarly in (2), as regards the visual appearance of the giveaway host. In (3) any perceptual properties the bill might have had are almost certainly not related to its likelihood of passing. In (4) the information prompting the speaker’s enthusiasm for a new cookbook is almost certainly not auditory. (All positive, numbered examples in this paper were attested on the web in July 2021, unless otherwise indicated.)

Sentences of this kind contain a subordinate clause complement, introduced by like, as if, or as though, the subject of which is a pronoun whose index is identified with the index of the matrix verb’s external argument. The latter bundle of facts has given rise to the name Copy Raising (CR). A key fact about CR sentences is that they have a paraphrase with an expletive subject and only one mention of the nominal expression in question. For example, a sentence such as (5a) Marion looks like she will be elected has two distinct rough paraphrases (5b) ‘It appears likely that Marion will be elected’ (CR) and (5c) ‘Marion’s visual appearance suggests that she will be elected’ (perception report). (Perhaps Marion is looking at the latest polls and smiling.)

(5) a. Marion looks like she will be elected. (invented example.)

   b. ‘It appears likely that Marion will be elected’ (CR)

   c. ‘Marion’s visual appearance suggests that she will be elected’ (perception report)

With regard to CR, the only structures that need special attention beyond the rest of the grammar are those in which (i) the subject pronoun of the like-phrase shares the index of the matrix subject and (ii) the meaning is that of the paraphrase just described. A sentence such as (6a) Pat looks like Marion is angry at him or even (6b) Pat looks like Marion is angry are widely understood to employ only the perception sense of the verb and be thus irrelevant to the analysis of CR, although there are certainly dissenters from this view (e.g., Asudeh and Toivonen 2012 and related papers, Lappin 1983, 1984.

[2] The phenomenon was originally discussed for English by Andy Rogers (1971, 1973, 1974a, 1974b), who gave the transformation he proposed to model the phenomenon the name “Richard”, perhaps to convey with a somewhat whimsical flourish how different this pattern seemed from raising or any other familiar grammatical pattern of English.
Heycock 1994, and Kim 2014). I will not discuss those dissents here; I find persuasive the arguments of Potsdam & Runner (2001), who conclude their examination of the relevant facts, “… true Copy Raising exists only where the pronominal copy is in subject position.”

(6) a. Pat looks like Marion is angry at him. (Invented example)
   b. Pat looks like Marion is angry. (Invented example)

An essential observation relevant to the CR phenomenon is that an expression of the form like/as if/as though + S[fin] is an ordinary kind of subordinate clause, not limited to copy raising locutions. Such clauses can appear as complements to verbs like act, behave and acquit (oneself) as in examples (7-9)

(7) … everyone high or low acquitted himself *(as if the fortune of the field depended on his own individual prowess.)

(8) Although it continued to float in midair, it acted *(like someone had cemented it to the ground).

(9) Americium, which is a pseudolanthanide, behaves *(as though it were roughly atomic No. 60.)

These clauses can also serve as complements to perception verbs, as in examples (10) through (12).

(10) Doug Collins looked *(like someone had just slapped him). (Kevin Sullivan & Mary Jordan, Trump on Trial: The Investigation, Impeachment, Acquittal, and Aftermath)

(11) The words sounded *(as if they were floating like flowers on water). (Virginia Woolf, To the Lighthouse)

(12) As for Ramsey…he had told me all along that his head was in danger, and he seemed *(as though the order was out for its removal).

Clauses of this form can also serve as adverbial modifiers of both transitive and intransitive verbal expressions.

(13) The man called her as though he was calling a little cat.

(14) This girl swims like it was something she was meant to do.

(15) She looked at him like he had lost it completely.
Kim (2014, among others) considers examples like (10-12) to be a kind of CR. The full range of examples like (7-15) suggests, on the other hand, that doing so complicates the full picture unnecessarily. There is no necessity to suppose that a sentence like (10), for example, has a grammatical form different from that of, say, (16).

(16) She looked like someone had died.

Let’s call such clauses as-if clauses; as-if clauses possess the following properties: like, as if, and as though SELECT a finite clause via the SELECT feature proposed by Allegranza (1989) and Van Eyn (2006), which groups determiners and modifiers together as functors. The finite $S$ selectee is the head of the as-if clause and consequently its external argument (XARG) is visible externally. As-if clauses are accorded a dedicated MARKING value, asif, and apply an appropriate frame to the INDEX of the head finite clause. (The term frame is the SBCG name for an MRS elementary predication. See Copestake et al. 2005.) Huddleston & Pullum (2002: 1151) argue for the lexical unity of as if, and as though, which is also usual in the CR literature, and recognize the syntactic reality of as-if clauses. They assume that like, as if, and as though are prepositions and their relation to their clausal accompaniments that of complementation rather than modification. Unsurprisingly, given the relevant publication dates, they do not appear to have considered the possibility of a functor-type analysis.

2. Expletives

Before presenting an analysis of CR, it is desirable to look briefly at how expletives behave in CR contexts, since expletives and related forms have furnished key data for many analyses of CR. For example, Sag (2012) presents the examples in (17) and (18). His implied analysis of CR consists in proposing an illustrative lexeme type, as exemplified by CR look in (19).

(17) There looks like \[
\begin{cases}
\text{there's going to be a storm} \\
\text{it's going to rain} \\
\text{*Kim's going to win}
\end{cases}
\] (Sag 2012: 150, (110a))

The point intended is that the CR version of look requires identity between the matrix subject and that of the phrase following like.

(18) ?Kim looks like \[
\begin{cases}
\text{there's going to be a storm} \\
\text{it's going to rain} \\
\text{Pat's going to win}
\end{cases}
\] (Sag 2012: 151, (110b))
Sag states that “the acceptability [of the examples in (18)] is slightly degraded.” The alleged degradation, and the attendant question mark, need not be considered grammatically relevant, as these examples appear to exemplify fully grammatical sentences that seem odd when presented in isolation only because the contexts in which they might be felicitously uttered do not spring to mind unbidden. Imagine in the first two sentences of (18) that Kim is a farmer who has hurriedly thrown on his rain gear and is dashing out the door. The examples in (18) are unremarkable in that each must be a perception report and not CR. Sag’s SBCG lexical analysis of CR is illustrated in (19).

(19) \[
\begin{array}{c}
\text{FORM} \\
\text{ARG-ST} \\
\text{NP,} \\
\text{S} \\
\text{XARG} \\
\end{array}
\] (look) \[\text{NP,} \quad \text{like} \quad \text{NP} \] (pron) \[
\text{Sag 2012:151, (111)}
\]

This analysis is not persuasive. The particle like serves only to make the XARG of the complement visible to the matrix XARG in the ARG-ST. It is otherwise unmotivated, performing no further function either in a CR environment or elsewhere in the grammar. On the other hand, we have seen that as-if clauses such as like there’s going to be a storm in (17) and like Pat’s going to win in (18) function as constituents in non-CR examples like (7) – (15) with the same apparent semantic effects as in (17) and (18). Ideally, the as-if clause type should be recognized as such in an analysis of CR clauses. As an initial step in investigating the role of expletives in CR, we note that English so-called ‘weather it’ is not properly considered expletive. In a mostly forgotten paper of nearly a half-century ago, Dwight Bolinger (1973) makes a compelling case that ‘weather it’ is not only not restricted to weather and time (It’s freezing; it’s nearly 4:00 a.m.) but rather denotes ambience generally. Some of Bolinger’s initial examples imagine a phone conversation between two forest rangers in different stations.

(20) a. ‘How’s it down there?’ – ‘It’s fairly calm’
   b. ‘How’s it up there?’ – ‘It’s practically ripping the trees out.’

Apparently, in (20) it includes the local wind conditions in its reference, but in a different location for each interlocutor. Bolinger continues:

(21) a. 'Isn't it nice out this afternoon?' – 'You must be crazy. It's so hot that it's giving me a headache.'
   b. 'It's cold enough to freeze the balls on a brass monkey.'

---

3 We will see, however, that the analysis of CR presented below is similar to that implied by (19) in being strictly lexical and in making strategic use of the XARG feature.
“Presumably in the last two sentences the *it* that refers to the weather is the same *it* that gives the headache and freezes the balls on the brass monkey” (Bolinger 1973: 262). Ambient *it*, as Bolinger shows, is protean. The examples in (22) are also from Bolinger (1973).

(22) a. It’s scary in the dark.
   b. It’s inspiring here at MIT.
   c. I’m climbing down. It’s too exposed up here.
   d. It’s her graduation next week.
   e. I like it in California.
   f. The noise makes it hard to study.
   g. It’s all finished between us.

One might disagree with Bolinger that the examples in (23), exemplify the same *it* as those in (20-22) but in any case, these tokens of *it* are also clearly not expletive. “I would maintain that the same *it* turns up in the following:”

(23) a. Stop it! (what you are obviously doing).
   b. Don’t do it! (what you are obviously about to do).
   c. Come off it! (what you are obviously insisting on)” Bolinger (1973: 263).

We proceed on the hypothesis that the *it* subject of weather sentences is not expletive. This hypothesis is consistent with the tradition in HPSG that postulates three types of index for English: expletive *it*, expletive *there*, and referential. Ambient *it* is a subtype of referential-index.

That “weather” *it* denotes ambiance entails that a weather predicate, unlike a CR verb, imposes a semantic role on its external argument. This in turn entails that a sentence like (24a) is three ways ambiguous. The lower *it* must be ambient because rain requires ambient *it* of its subject. The matrix *it* can be either ambient or expletive; if it is ambient the sentence can be CR, but it can also be a perception report with the two ambient *it* tokens referring differently: the matrix *it* to the current ambience and the lower *it* to an imagined future ambience, somewhat as we saw in (20) regarding contrasting wind conditions. If the matrix *it* is expletive, (24a) is not a CR sentence; *seems* is the familiar raising verb in its alternate, *it*-extraposition, valence, and the sentence is interpreted like a sentence such (24e). Example (24b) is bad because rain requires ambient *it*. Example (24c) is bad for the same reason; that reason also explains why CR is not possible, despite identity of the matrix subject’s index and the complement’s subject’s index. Example (24d) is bad because it lacks the identity of indices required by CR and neither perception *seem* nor *it*-extraposition *seem* allows a there subject. Sentence (24e) is a perception report, with perception *seem* and ambient *it* subject. Example
(24f) is boringly bad because there is no bound pronoun as required for CR, and both perception seem and it-extraposition seem reject there subjects. Finally, (24g), which contrasts minimally with (24c), exemplifies CR, with matching expletives there; be a storm is not a weather predicate, which would require an ambience-referring it, as subject.

(24) a. It seems like it’s going to rain soon.
   b. *It seems like there is going to rain soon.
   c. *There seems like there is going to rain soon.
   d. *There seems like it is going to rain soon.
   e. It seems like rain is coming soon.
   f. *There seems like rain is coming soon.
   g. There seems like there is going to be a storm soon.

3. Analysis of Copy Raising

Noting that the external argument of a CR verb does not denote a source of perception, we can observe that there are four English perception verbs that undergo the copy raising lexical rule. These are the four most general perception verbs: seem, appear, look, and sound. These and only these verbs can yield a hearsay reading in a sentence of the form illustrated in (25) (invented). Examples a-d have a common reading, roughly, ‘Apparently, Nero didn’t really burn Rome’; examples e and f do not have such a reading.

(25) a. It seems like Nero didn’t really burn Rome.
   b. It appears as if Nero didn’t really burn Rome.
   c. It looks as though Nero didn’t really burn Rome.
   d. It sounds like Nero didn’t really burn Rome.
   e. # It smells like/as if/as though Nero didn’t really burn Rome.
   f. # It tastes like/as if/as though Nero didn’t really burn Rome.

Assuming a multiple-inheritance hierarchy of frames (Davis & Koenig 2000) we posit a perception frame percep-fr, whose immediate subtypes include gen(eral)percep-fr, smell-percep-fr, and taste-percep-fr. The immediate subtypes of gen-percep-fr are seem-percep-fr, appear-percep-fr, look-percep-fr, and sound-percep-fr, as represented in the type hierarchy fragment (26).

(26) percep-fr
    |   gen-percep-fr  smell-percep-fr  taste-percep-fr ...
So far, nothing has been formally proposed regarding CR per se; as-if clauses are vanilla English grammar. SBCG expresses lexical rules as unary-branching, derivational constructions. The Copy Raising Construction pumps a general perception verb lexeme to a verb lexeme with a meaning that might be characterized as imparting a weak evidentiary force, perhaps similar to the meaning of seem in a sentence like (27), and whose output (MTR) lexeme identifies the index of its NP XARG with the index of the pronominally specified XARG of the as-if finite clausal complement, as shown in (28).

(27) … the bill seems like a positive step for our state.

(28) MTR [ARG-ST <NPᵢ, S[MRKG as-if, SYN[XARG NP[pron],>]]

This analysis, despite departing from that of Sag (2012), shares with it both a strictly lexical approach and critical dependence upon the XARG feature: although the selectee finite clause in an as-if clause has empty VAL(ENCE), its XARG remains available to be coindexed with the matrix subject. The key points of this approach are (i) the empirical observation that all the verbs that participate in CR have a perception-verb double (although the converse does not hold, since only the four general perception verbs have a CR double), (ii) that expressions such as like he was going to win in (1) occur freely in several non-CR contexts, as illustrated in examples (7) – (15), (iii) the fact that the it subjects of weather sentences are best conceived as referential, (iv) the All-egranza-Van Eynde innovation of the Head-Functor Construction, based on the SELECT feature, which in the present context enables the subject of the as-if complement clause of an erstwhile perception verb to be visible in the latter’s ARG-ST, and (v) that aspect of the architecture of SBCG (and of related forms of HPSG) that makes it possible for a realized external argument to nonetheless be visible in the ARG-ST of a governing predicador.

For illustration we consider the aspirational CR sentence (29). First, we take up the ordinary listeme like, which appears in examples (1-4), (10), (14), (15) and (29), recalling that examples (7-15) are not CR. (Since scope constraints play no direct role in the present analysis, we adopt a kind of MRS Lite notation, in which constraints on relative scope, such as qeq constraints, are unexpressed, in effect ignoring the distinction between LTOP and GTOP.)

(29) Trump looks like he disappeared.
The functor *like* is an adverb that selects a finite clause ([VFORM *fin*, [INV −]); the selected clause specifies a nominal external argument and an empty valence list. The semantics of *like* applies the as-if-frame to the LTOP of its selectee clause. Also, *like* is marked asif. As a functor, it will transmit [MARKING asif] to its MOTHER in a head-functor-construct.

The only addition we make to the grammar of English to account for the CR facts is the lexical rule in (31), which takes as input a general perception verb lexeme, that is, one whose FRAMES value is compatible with <gen-percep-fr>, and yields a CR verb lexeme as output. A CR verb (as specified by the mother in (31)) subcategorizes for an NP subject Z and an as-if clause complement. Semantically, it specifies a seeming-frame with three arguments: the Davidsonian event variable e, a human experiencer argument j, and the semantic information, labeled l of the as-if complement, which is the state of affairs that seems to j to be the case. Also – and essential to the CR phenomenon – the CR verb, identifies the index of its external argument Z with the index of the pronominal external argument of the as-if complement.

---

4 Comment on notation: For constraint descriptions (AVMs) [A], [B], the paired tags ‘X ! [A]’ and ‘X : [B]’ indicate that [A] and [B] are identical in all respects in which they are not shown to differ. This abbreviatory notation is commonly used in SBCG in the statement of lexical rules, though not limited to that use. (See Sag 2012: 125, including footnote, for further discussion.) AVMs in boxes represent model objects (feature structures, as against descriptions of feature structures, such as types and constructions).

77
(31) Copy Raising Construction \(\uparrow\)derivational-cxt [a lexical rule]

\[
\begin{align*}
\text{copy-raising-v-cxt} & \quad \Rightarrow \\
MTR & \quad X ! \\
\text{SYN} & \quad \begin{cases}
\text{CAT} & \left[ \text{XARG Z:NP} \right] \\
\text{VAL} & \left[ Z, \right]
\end{cases} \\
\text{SEM} & \quad \begin{cases}
\text{FRAMES} & \left\{ \text{seeming-fr(e),human-fr(i),l1} \right\} \\
\text{INDEX} & \{ e \}
\end{cases} \\
\text{DTRS} & \quad x : \\
\text{SEM} & \quad \begin{cases}
\text{CAT} & \left[ \text{VFORM fin XARG NP[pron]} \right] \\
\text{VAL} & \left\{ \right\} \\
\text{MRKG} & \{ asif \} \\
\text{LTOP} & \{ l \}
\end{cases} \\
\text{FRAMES} & \left\{ \text{gen-perception-fr(e)} \right\}
\end{align*}
\]

By way of illustration, Figure 1 shows the derivation tree for the aspirational sentence (29) *Trump looks like he disappeared.*

We take sentence (29) to have the approximate gloss ‘It seems as if Trump disappeared’. Starting at the bottom of the tree, consider the right, head-daughter sign. In the SEMANTICS value, the FRAMES value specifies the *disappear-frame* applied to the event variable \(e\), which is the INDEX of the sign, and the individual variable \(i\). The FRAMES list also contains the *past-fame* applied to \(e\). According to the tag \(\square\) the head daughter is the value of the SELECT feature of the left, functor daughter, *like*, in a head-functor-construct. The functor daughter’s FRAMES feature applies the *as-if-frame* to the local INDEX \(e\)’ and to the *disappear* and *past* predications jointly labelled \(l_1\), the value of the LTOP of the head daughter. The *as-if-frame* is assumed to be interpreted by an epistemic operator. The *like* sign also introduces the MARKING value *asif*, which is passed up to the MOTHER sign, *like he disappeared*, in a head-functor-construct. The mother of that construct, *like he disappeared*, gathers up the frames of its two daughters and inherits the rest of its information from the head daughter.

To build the VP *looks like he disappeared* with the Predicational Head-Complement Construction, we note that we have employed the Copy Raising lexical rule (31) to build the CR form of *look* from general perception *look*. Inflected CR *looks* introduces in its FRAMES list the *present-frame* predicated...
of the local index $e''$. In this example the semantic equivalent of so-called arbitrary PRO is introduced as the experiencer of the *seeming* predication by application of the *generic-fr* to the bound variable $j$, restricted by the *human-frame*. The *seeming-frame* specifies two arguments in addition to the event variable $e''$: $j$ and the *asif-fr* predication $l_z$. The *seeming* predication thus

---

Figure 1. Derivation Tree for *Trump looks like he disappeared*
specifies that the as-if predication \( l_2 \) is the state of affairs that seems to the generic human \( j \) to be the case. \textit{Looks} also specifies in its VALENCE list the CR-characteristic coindexation of its NP XARG and the XARG of its \textit{as-if} clause complement. The latter valent, bearing the tag \( \texttt{[4]} \), is discharged in the \textit{predicational-head-complement construct} by the right sister of \textit{looks}. We have already discussed this sign as the mother of the \textit{like-he-disappeared} construct. As a head, the \textit{looks} constituent bears the MARKING value \textit{unmarked}, which it passes up to its mother, \textit{looks like he disappeared}, in the \textit{predicational-head-complement-construct}.

Finally, The Head-Subject Construction (AKA Subject-Predicate Construction) realizes the remaining valent \( \texttt{[2]} \) as the subject \textit{Trump}. The mother constituent of the head-subject-construct, \textit{Trump looks like he disappeared}, gathers up the four frames of its daughters, inherits the marking value \textit{unmarked} from the head VP, and expresses with an empty VALENCE list the fact that the NP \textit{trump}, whose index \( i \) is identified with that of the pronoun \textit{he}, has satisfied the last remaining valent.

**Conclusion**

The lexical rule analysis of CR presented here has benefited particularly from three empirical observations and two formal aspects of SBCG and related versions of HPSG. The empirical observations are not necessarily new, but they have not to my knowledge been marshalled in this combination before. They are that (i) to each verb participating in CR there corresponds a homophonous general perception verb, (ii) the pronoun \textit{it} that serves as the subject of weather verbs is referential, and (iii) the word string initiated by \textit{like}, \textit{as if}, or \textit{as though} in a CR sentence forms a single constituent, a type of subordinate clause (christened here “as-if” clause) that occurs as both complement and modifier elsewhere in the grammar. The formal aspects of SBCG/HPKG that are strategically employed in the present analysis are as follows: (i) The \texttt{SELECT} feature makes it possible for the finite clause requirement of the CR-marking expression (\textit{like}, \textit{as if}, or \textit{as though}) to be a selectee rather than a complement of that expression. This circumstance identifies the XARG of the finite selectee with the XARG of the as-if clause, and so renders it potentially visible in the ARG-ST of the CR verb. (ii) Since the \texttt{XARG} feature, unlike the ARG-ST feature, percolates up the line of heads, the XARG of the as-if complement can be addressed in the ARG-ST of the CR verb and specified as a pronoun that shares its index with that of the matrix XARG. When the three empirical observations are considered, the SBCG/HPKG formalism enables an account of the CR facts with a lexical rule that inputs a general perception verb and outputs a verb that identifies the index of its external argument with the index of the external argument of its as-if comple-
ment and whose semantics specifies a mildly evidentiary *seeming* meaning of which the experiencer argument may be covert and interpreted generically.

This paper has considered copy raising only for English. Future research will have to determine whether this approach is helpful in understanding copy raising in other languages.

**Acknowledgments**

The story of CR presented here has arisen from many conversations on this subject with Laura Michaelis. I owe Rui Chaves an unpayable debt for his generous counsel on formalizing the analysis. I thank Jong-Bok Kim and Eric Potsdam for helpful email correspondence. I would also like to express my appreciation to the organizing committee of the 2021 HPSG International Conference for inviting me, as well as to the conference participants, several of whom, especially Adam Przepiórkowski, provided helpful suggestions. Remaining flaws are, of course, mine.

**References**


