

On the Notion 'Related Lexical Entry'

Georgia M. Green
University of Chicago

"When I use a word," Humpty Dumpty said, in a rather scornful tone, "it means just what I choose it to mean--neither more nor less."

"The question is," said Alice, "whether you can make words mean so many different things."

"The question is," said Humpty Dumpty, "which is to be master--that's all."

At least since the development, in the past one hundred years, of formal theories of morphology, students of linguistics have agonized over the difficulties of determining whether two phonologically indistinguishable forms were "the same word" or "the same morpheme", and frequently they suffered no small amount of intellectual discomfort at the intuitively unsatisfactory solutions their theories did lead them to. I do not pretend that what follows contains definitive answers to all the questions I am going to raise, but I hope that the raising of these questions will permit some insight into the nature and depth of the nagging doubts referred above, so that we can get down to the serious business of looking for the answers which will lead to a definitive treatment.

The agonizing difficulties I mentioned usually take the form, for example: Are all (or any) of the tokens of the form [æsk] in the following sentences tokens of the same abstract semantic entity?

- (1a) I should ask why the sky is blue.
- (1b) I should ask to go.
- (1c) I should ask them to go.
- (1d) I should ask when to go.
- (1e) I should ask them when to go.
- (1f) I should ask them for \$10.
- (1g) I should ask them to dinner.
- (1h) I should ask them the way to the Administration Building.
- (1i) I should ask \$5 for that lamp.
- (1j) I should ask descriptive adequacy of every theory I have to consider.

Ask occurs in these sentences in a variety of syntactic constructions. It has, in traditional terms, nominal adjuncts, infinitival complements, clausal complements, and various combinations and permutations of these.

Paraphrases of these sentences:

- (2a) I should EXPRESS A DESIRE TO BE TOLD why the sky is blue.
- (2b) I should EXPRESS A DESIRE TO BE PERMITTED to go.
- (2c) I should EXPRESS A DESIRE to them that they go.
- (2d) I should EXPRESS A DESIRE TO BE TOLD when to go.
- (2e) I should EXPRESS A DESIRE to them that they TELL me when to go.
- (2f) I should EXPRESS A DESIRE to them that they GIVE me \$10.
- (2g) I should EXPRESS A DESIRE to them that they COME to dinner.
- (2h) I should EXPRESS A DESIRE to them that they TELL me the way to the Administration Building.
- (2i) I should EXPRESS A DESIRE TO RECEIVE \$10 for that lamp.
- (2j) I should DESIRE AND EXPECT of every theory I consider that it BE descriptively adequate.¹

indicate that there are several different meanings distributed among the asks in (1), but that there is in all a common element of wanting, and in all but one, an element of communication. We feel cheated somehow, if our theory treats these asks as as similar as the asks in (3):

- (3a) I asked my mother for an apple.
- (3b) I asked my mother for a cookie.

or as as different as the bores in (4):

- (4a) John bored me when I last saw him.
- (4b) John bored holes in my book when I last saw him.

or the [feynt]s in (5):

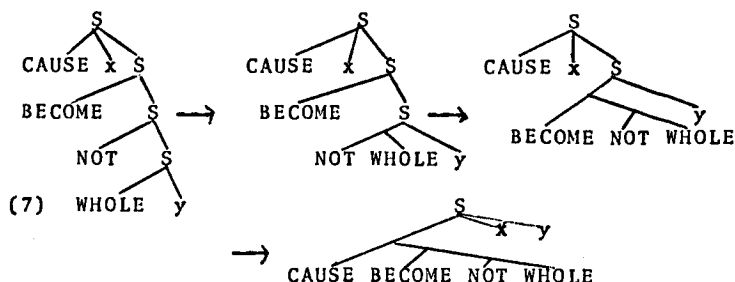
- (5a) Hamlet fainted on his sword.
- (5b) Hamlet feinted with his sword.

We do not want to, and in my opinion, should not settle for a theory which does not distinguish these in accord with our intuitions, and which does not in some satisfactory way explain how all the different meanings listed in (2) can be expressed by the same phonological form.

A theory of grammar in which the deepest level of underlying structure is that of semantic representation, and in which lexical items encode transformationally derived semantic complexes (as sketched in McCawley (1968)) provides a formal framework for discussion and investigation of the relatedness of lexical entries. That is, with a syntactic theory in which a lexical item like break encodes not only structures like (6):



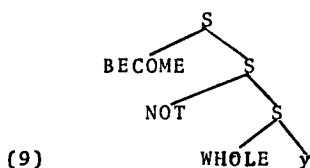
derived by successive operations of a predicate-raising transformation from a structure like (7):



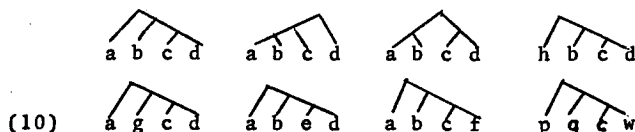
but also encodes a structure like (8):



derived in a similar manner, from (9):



we can ask, "What formal relations between derived structures constitute relatedness between lexical entries?" For example, which, if any, of the tree types in (10) could be considered related lexical entries if they should all have the same phonological form and morphological irregularities?



This question is a rather oblique way of conveying the distinction in my mind between the terms "lexical entry" and "lexical item". I conceive of the lexicon as consisting of a large set of distinct derived semantic trees, a much smaller set of underlying phonological

forms with specification of their morphological idiosyncracies (suppletions, truly irregular inflections), and some mechanism(s) relating them. The trees I will refer to as lexical entries, the phonological forms with indication of morphological properties, as lexical items.

Lexical items so defined can be considered morphemes without meaning: for each lexical item there must be specification of its underlying phonological shape(s), idiosyncratic morphological properties, and syntactic exceptionality features, if any. It follows from this conception, that I view all relations between meanings (related or not) and forms as entirely arbitrary. I am forced into this position by two considerations:

(1) Not only in languages like English with its thoroughly bastardized, irregular, and unsystematic derivational morphology, but also in languages with much more regular and systematic derivational morphology like Japanese, there are exceptions to broad and obvious generalizations about relations between form and meaning. Thus, while causative verbs in English often have forms identical to intransitive verbs with related meanings (break, cool, enlarge, roll, etc.), and Japanese causatives and intransitives are derivationally related, in neither language is the word for 'kill' (which is the semantic causative of 'die') morphologically related to the word for 'die', although the Aztec word for 'kill' (mictia) is the morphological causative of the word for 'die' (miqui). On the other hand, in Hindi, the causative of 'die' means to 'beat'.

(2) I have come to believe that the differences among languages are primarily lexical,² and that to the extent that lexical peculiarities such as the above³ exist, they are lexical and idiosyncratic, and it is inappropriate and unfeasible to treat them as semantic, syntactic, or systematic.

It follows from these two considerations that I view the establishment of a level of representation distinct from that of "word" whose units are meanings and associated forms, or forms and associated meanings (i.e. classical morphemes) as superfluous and insufficiently motivated. The denial that "morphemes" or lexical items have meaning makes it all the more important that there be a way to characterize relatedness among lexical entries. Languages differ greatly in the systematicity with which they correlate forms and meanings, and it is this I hope to be able eventually to discuss with the aid of the notion relatedness among lexical entries.

I cannot emphasize strongly enough the difference between a lexicon which conforms to what I have proposed, and one organized according to the principle "one phonological form, one lexical entry", exemplified in various otherwise quite different forms by Katz and Fodor and Fillmore. Implicit in such lexicons as the latter is

the claim that syntactic differences among the various meanings of a single lexical item are arbitrary and unsystematic, and generally, unexpected deviations from regularity. But as has been pointed out by Zellig Harris and George Lakoff (1965:VII-8), there are, and I suspect, to a greater extent than has been commonly recognized, usually, one-to-one correlations of meaning and differential syntactic properties, such as (for verbs) number and type of associated noun phrases, momentary versus continuative adjuncts, and other "selectional restrictions".

Before we can begin to answer the question I posed earlier, (and I shall not attempt to answer it here; it deserves at least ten dissertations of study), we have to have a way of detecting whether two uses of a given phonological form can be considered to be tokens of the same lexical entry (with or without accompanying transformational deletion of surrounding--this is another matter), or whether they must be considered tokens of mere homophonous lexical items. Behavior as anaphora offers one such test.⁴

For instance, the intransitive (inchoative) and transitive (causative-inchoative) breaks do not behave anaphorically with each other--the second to occur seems to be de-stressed only in contrastive situations:

- (11) She saw the window break, but she didn't know what bróke it.

although transitive break causes transitive break to become de-stressed, and intransitive break causes intransitive break to become de-stressed.

- (12) I saw the glass Floyd broke, but I couldn't tell what { it broke.
 {*it bróke.

- (13) A rock hit the window, and it bróke, causing a vase on the sill to { fáll and break.
 {*fáll and bréak.

Contrastive stress on the anaphoric it in (12) seems to come from the stress being displaced one major constituent to the left of its normal position on the last major constituent in the verb phrase; anaphoric pronouns are ordinarily totally lacking in stress.⁵

Possibilities for conjunction reduction offers another test for same lexical entry, when the lexical item in question takes the same number of complements consistently. Thus conjunction reduction is possible for (1a) and (1d):

- (14) I should ask both why the sky is blue, and when to go.

but not for (1b) and (1d):

- (15) *I should ask (both) $\left\{ \begin{array}{l} \text{to go and when to go.} \\ \text{when to go and to go.} \end{array} \right.$

or for (1f) and (1g):

- (16) *I should ask them $\left\{ \begin{array}{l} \text{for \$10 and to dinner.} \\ \text{to dinner and for \$10.} \end{array} \right.$

or for (1i) and (1j):

- (17) *I should ask \$5 for the lamp I'm selling and descriptive adequacy of the theory I'm buying.

Assuming then, that we can factor out cases of identical lexical entries,⁶ what are we really talking about when we talk about the senses of lexical items such as the asks in (1) or the refuses in (18) as different but related?

(18a) John refused to go.

(18b) John refused Mary, but he accepted Sue's invitation.

(18c) John refused the cake, but he ate the ice cream.

As I tried to argue above, the apparently idiosyncratic nature of the semantic relationships among "derived" verbs, and between verbs and adjectives and related nominalizations as well, means that the minimal meaningful units in language must be dealt with below the level of formal morphology, but that the connection between form and meaning must be a static relationship, not a process. Unpleasant implications follow from the claim that there is, for example, only one lexical entry for the lexical item refuse, and that the other meanings are derived by syntactic processes of incorporation or deletion of semantic constituents. The transformational incorporation of semantic constituents into lexical items involves lexical insertion being both cyclic and obligatory, and there are substantial arguments against both of these hypotheses (Green 1969:9-11, 58, McCawley 1968) based on the undesirability of extending the location and distribution of exceptionality in a grammar. Lexical deletion rules would likewise have to be very specific, deleting specific lexical items, and highly sensitive to context. For example, to account for the lack of ambiguity of (19) and (20):

(19) Harry refused Bill's request.

(20) Harry refused Bill's offer.

since:

- (19) = Harry refused to grant Bill's request.
 ≠ Harry refused to accept Bill's request.
- (20) = Harry refused to accept Bill's offer.
 ≠ *Harry refused to grant Bill's offer.

one would need the post-lexical deletion transformations:

- (21) accept → ∅ in env. refuse _____ $\left\{ \begin{array}{l} \text{offer} \\ \text{help} \\ \text{aid...} \\ \text{[class of all} \\ \text{concrete ob-} \\ \text{jects]} \end{array} \right\}$
- (22) grant → ∅ in env. refuse _____ $\left(\begin{array}{l} \text{NP} \\ +\text{animate} \end{array} \right) \left\{ \begin{array}{l} \text{request} \\ \text{aid} \\ \text{help...} \end{array} \right\}$
- (23) grant → ∅ in env. refuse _____ $\left[\begin{array}{l} \text{NP} \\ +\text{animate} \end{array} \right] \left[\begin{array}{l} \text{class of} \\ \text{all con-} \\ \text{crete} \\ \text{objects} \end{array} \right]$

There is reason to believe that a certain amount of context-sensitivity is inevitable in a description of the lexicalization of meaning (Green 1969), but it is not necessary that the context-sensitivity be lexical and located in highly specific post-lexical transformations. If the requisite context-sensitive rules are pre-lexical, they will be in terms of semantic units, and will express the generalizations that such statements as (21) through (23) obscure.

I still have not gotten to the point: which of the configurations in (10) should be considered different but related lexical entries. I continue to have very little in the way of an answer, but some unrelated research of mine (Green 1969) suggests that there may be reason to believe that lexical entries which differ merely in the order or configuration of their semantic constituents can be realized by the same lexical item. For example, unlikely might be a realization of either (24) or (25):

- (24) NOT LIKELY (25) LIKELY NOT

as in (26) and (27) respectively:

- (26) It's unlikely that John will come too.
 (27) It's unlikely that John will come either.

since (27) presupposes that someone else won't come,

while (26) presupposes that it is not likely that someone else will come, or that some other event is not likely.⁷

But for the present, I am incapable of providing a clearly defined theory of relatedness among lexical entries. My main purpose here, however, is to make it clear that it is not merely desirable, but necessary, that we have one. Without it, we cannot describe with appropriate generality cross-linguistic similarities such as the fact that languages as diverse as English and Hopi use the same lexical items for expressing possession and willful causation:

- | | |
|----------------|--------------------------|
| (28) siwa-?yta | He has a younger sister. |
| (29) somi-?yta | He <u>has</u> it tied. |

or the fact that the morphologically unrelated words make in English, and faire in French cover essentially the same wide range of uses, for example

- | | |
|-----------------------------------|------------------------------------|
| (30) God made man in his image. | Dieu a fait l'homme à son image. |
| (31) God made all men equal. | Dieu a fait tous les hommes égaux. |
| (32) The sun makes the snow melt. | Le soleil fait fondre la neige. |
| (33) Make him read this letter. | Faites-lui lire cette lettre. |
| (34) He made a mistake. | Il fit une faute. |
| (35) He made a gasp. | Il fit un petit "oh" de surprise. |
| (36) He made the beds. | Il a fait les lits. |
| (37) 2 and 2 make 4. | 2 et 2 font 4. |
| (38) That makes twice. | Ça fait deux fois. |

or the fact that languages all over the world use the same expression (word or case) for 'means' and 'accompaniment'. Nor can we describe appropriately without a clear notion of relatedness the cross-linguistic distinctions in the use of a single term for related concepts. For example, the English verb refuse has several senses. By and large, but with exceptions everywhere, Italian, French, and German also use a single lexical item for these senses. So we have the situation charted in (39).

- Meaning 1: "WILL" NOT DO
 Meaning 2: "WILL" NOT ?DO ACCEPT
 Meaning 3: "WILL" NOT ?DO "GRANT"
 Meaning 4: "WILL" NOT ?DO "GIVE"
 Meaning 5: "WILL" NOT ?DO ACKNOWLEDGE

	English	Italian	French	German
1	<u>refuse</u> to	<u>rifiutarsi</u>	<u>refuser</u>	<u>sich weigern</u> <u>verweigern</u> [nmz.]
2	<u>refuse</u> [sthg. offered]	<u>rifiutare</u>	<u>refuser</u> <u>rejeter</u> <u>repousser</u>	<u>ausschlagen</u>
3	<u>refuse</u> [s.o. a requested state]	<u>rifiutare</u>	<u>rejeter</u> <u>repousser</u>	<u>abschlagen</u> <u>verweigern</u>
4	<u>refuse</u> [s.o. a requested act]	<u>rifiutare</u>	<u>rejeter</u> <u>repousser</u>	<u>abschlagen</u>
5	<u>disown</u>	<u>rifiutare</u>	<u>désavouer</u>	<u>ableugnen</u>

(39)

This "description" of the data is only observationally adequate. Nonetheless, in merely bringing together the data it contains, it raises questions for which an explanatory theory of grammar must provide answers, or at least the means to find answers. For instance, refuse is an equi-subject verb. Sentences such as (40) are ungrammatical:

(40) *John refused for Mary to wash her hands.

Why is the Italian equivalent reflexive, the German equivalent suppletively reflexive, and the French not reflexive at all, even though French has many reflexive verbs? I suspect these are syntactic facts about the various languages, rather than lexical facts, but I haven't the theoretical apparatus yet to demonstrate it. Why is German the only one of these languages to permit either a sentential or a nominalized complement for the least complex meaning of refuse, with concomitant suppletion? Although again I cannot demonstrate it, I imagine this to be a lexical fact about German, of the same status as the idiosyncratic fact about English that the terms of correlative conjunctions (both...and, either...or, neither...nor) are not identical--but compare French et...et, Albanian edhe...edhe, etc.

Even if making the proper cross-linguistic generalizations (and separating them from language history) in an appropriate fashion is to be dismissed as a goal reserved for the distant future, there remains a more immediate problem which a clear notion of relatedness will facilitate coming to grips with, namely, the investigation of the semantic structure and associated pre-lexical syntax of terms whose meanings are not obvious. For instance, what does either mean in the following sentences:

(41) Bill won't go, and I won't go either.

(42) You will vote for Nixon in 1972. I will not either!

(43) Either Bill will go, or I will.

Does it mean the same thing in each sentence? If not, do the meanings have anything in common?

So we are still doubting, with Alice in Wonderland, that there is no limit to what words can be used to mean. The vital question is, in our terms, what are the constraints on how two or more meanings realized by the same lexical item can differ, or more simply, just what constitutes relatedness among lexical entries?

FOOTNOTES

For the stimulating discussions which led to my interest in this topic, I am indebted to Jerry L Morgan, George Lakoff, Robert I. Binnick, and James D. McCawley.

¹I am unable to find an adequate paraphrase of (1j) using a noun phrase for the first adjunct of ask, which seems to have to be a nominalization. Cf.:

- (i) *I should ask a phonemic level of every theory
I have to consider
- (ii) She asks a lot of her sons.
- = (iii) She desires and expects of her sons that they do a lot.

To what extent the notions of expectancy and desire are presupposed and not asserted here is not clear.

²This is by no means a new idea. Compare for example the Baconian quotation in Salus' article in this volume.

³Such gaps in lexical relations are widespread in English. On the one hand, there are productive, morphologically unmarked systems of word "extension" such as the causative-transitive (motion) verbs and their intransitive counterparts as in Tables I and II:

Table I: Motion Verbs

move	move
walk	walk
sail	sail
go	send
come	bring

Table II: Other Change of State

break	break
cool	cool
harden	harden
die	kill
believe	persuade

and the de-nominal conveyance verbs (Binnick 1968a,b) and instrumental communication verbs in Tables III and IV:

Table III: Conveyance

punt	punt
jet	jet

Table IV: Communication

radio	radio
signal	signal

But these are not 100% productive; in addition to the go/send, die/kill, letter/write kind of suppletion, there are unexplainable holes in the pattern. The relationship of transitive skip to intransitive skip is not the same semantic relationship as transitive walk to intransitive walk. To sneak someone into a room is not to cause him to sneak into a room. There are no transportation verbs corresponding to whole fleets of nouns: chariot, wagon, train, stagecoach, car, truck, bus. One can jet to Miami, but in English, one can't plane there. You can radio for help, you can sneeze for help, if a sneeze is a signal previously agreed upon, but I can't conceive of loudspeaking for help, or turning that some enemy had been sighted.

⁴This test has been proposed independently, in various guises by several linguists, including Lakoff and Postal.

⁵Let me add a caveat, however, in offering this test, that presuppositions seem to affect stress exactly as do previous occurrences of lexical items (Green 1969). Thus:

(i) How does it feel to be a beautiful girl?

is a felicitous question only when its utterer believes his addressee is a beautiful girl, whereas

(ii) How does it feel to be a beautiful girl?

may be asked felicitously under just about any circumstances. (ii) has normal sentence intonation; in (i) the stress is displaced one major constituent to the left, as if the final major constituent were a second occurrence.

⁶Alice Davison has pointed out to me that volitional and non-volitional transitive-causatives are treated as identical to each other for the purpose of destressing, as in:

(i) This is the glass Floyd broke with his fist all right, but we can't tell what $\left\{ \begin{array}{l} \text{it broke,} \\ \text{*it bróke.} \end{array} \right.$

(ii) Splinters from the glass which broke covered the light bulbs Floyd used his fist to break.

Note however, that conjunction reduction is not permitted in a conjunction of volitional and non-volitional causatives:

(iii) John and { *the wind } broke a window and a
 { Harry
 shutter, respectively.

(iv) *Tuberculosis and Othello killed Camille and
 Desdemona, respectively.

so it appears that identity is defined differently for
 the purposes of these two syntactic phenomena. How,
 and why, is not clear.

⁷On the other hand, it may merely be the case that
unlikely in (27) results from pre-lexical negative-
 raising (Green 1969)--cf. parallel to (26) and (27):

- (i) It is not likely that John will come too.
- (ii) It is not likely that John will come either.

BIBLIOGRAPHY

- Binnick, Robert I. 1968a. On the nature of the "lexical item". Papers from the Fourth Regional Meeting, Chicago Linguistic Society, ed. By Bill J. Darden, Charles James N. Bailey, Alice L. Davison, 1-13. Chicago, Department of Linguistics, University of Chicago.
- _____. 1968b. The characterization of abstract lexical entries. Paper read before the Association for Computational Linguistics, 6th annual meeting.
- Fillmore, Charles. 1968. Lexical entries for verbs. Foundations of Language. Vol. 4, No. 4, 373-393.
- Green, Georgia M. 1969. Some theoretical implications of the lexical expression of emphatic conjunction. University of Chicago, duplicated.
- Katz, Jerrold, and Jerry Fodor. 1964. The structure of a semantic theory. The structure of language: readings in the philosophy of language, ed. by Jerry A. Fodor and Jerrold J. Katz, 479-518. Englewood Cliffs, N.J.
- Lakoff, George. 1965. On the nature of syntactic irregularity. Indiana University Dissertation.
- McCawley, James D. 1968. Lexical insertion in a transformational grammar without deep structure. Papers from the Fourth Regional Meeting, Chicago Linguistic Society, 71-80.