WHY THE ONLY INTERESTING SYNTACTIC DIALECTS ARE THE UNINTERESTING ONES

Linda Coleman
University of Michigan

Any discussion of some part of linguistic theory can be brought to ruin by a naive speaker (or another linguist) pointing out that he finds some of the starred sentences acceptable, or vice versa. Such discussions, furthermore, are extremely difficult--for the neophyte, at least--to follow when, as not infrequently happens, he disagrees with the author on the acceptability of many of the sentences on which the entire framework of the argument rests, with no idea why he disagrees or whether his disagreement affects the discussion at all. One can, of course, chalk it all up to "dialect variation," for the concept of syntactic dialects was originally fudged into existence to account for precisely such intuitional variations between speakers when more urgent matters were under discussion, but one is still left with no idea of whether this lack of concord will be apt to affect his ability to accept the author's argument, or where else reader and author are likely to disagree.

It is necessary, then, to determine for purposes of reference, if for no other reason, whether there is any real sense in the sorts of syntactic variation which occur between speakers. The purpose of the study presented in this paper was simply to try to find out whether such variation is at all logical and predictable or whether it is merely the result of individual idiosyncracies. If the former is true, it will be necessary to determine what sort of variation is occurring, whether the variation behavior of an individual speaker can be predicted, whether correlations can be made between various individual speakers, and finally, whether syntactic variation has any effect on any significant part of linguistic theory.

The study was conducted with a final total of 32 informants. The test itself was in three parts, two in written form and one oral:

(a) Preference test: the informant was given 15 questions, with two sentences each, and asked to indicate which of the two sentences he preferred. He was also asked to mark any sentence which seemed very odd or unacceptable.

(b) Rating test: the informant was given 52 sentences, in random order, and asked to rank them on a 4-point scale, as follows:[1]

   1 = you would say it naturally, with no problem.
   2 = it's a bit awkward, but you still might say it.
   3 = it's very awkward. You would not say it, but you could imagine someone else saying it, perhaps.
   4 = it's rotten. Nobody would say it.

(c) Elicitation session: during the course of what was ostensibly a casual conversation, I attempted to elicit the informant's reactions to sentences tested in (b) by trying to get him to say sentences he had rejected, as well as by using test sentences myself and eliciting his reaction by suggesting that I regarded a given sentence as good or bad, depending upon his rejection or acceptance of it, respectively, under test (b).[2]

Among the items tested for were there-insertion, relative clause formation, relative clause placement (extraposition from NP), passivization of verbs of thought and communication, use of complementizers with the predicate _likely_, and passivization of sentences of pattern (6), below. The test was admittedly limited and used only a few sentences of each type, for my main purpose was to determine whether or not syntactic variation follows any logical pattern; the question of what sort of rules operate to produce such a pattern must be left until more research has been done.

Due to lack of space, only one of the charts is reproduced. The chart on the following page shows the results for sentences (1) through (5), dealing with complementizer restrictions. It is representative of the results for most of the other items.

I. (1) It is likely that John will leave.
    (2) That John will leave is likely.
    (3) It is likely for John to leave.
    (4) John's leaving is likely.
    (5) For John to leave is likely.

(1) is not included in the chart, since all responses were positive. The brackets on (3) and (4) indicate that the number of informants who found an acceptability difference between these two sentences is probably not significant. Thus, another continuum is possible: (1)(2)(4)(3)(5), which would include 25 out of the 32 informants—still a decent number for this test.

Other tested items which fit into continua are: II – extraposition from NP, III – relative clause formation, IV – passives of verbs of belief, communication, etc., and V – passivizations of sentences of the pattern:

(6) NP_1 + Verb + Preposition + NP_2

The sentences are ranked in order from most to least acceptable, in accordance with informant responses:

II. (7) A performer will appear shortly who is dressing now.
    (8) I let the cats out which were meowing.
    (9) Who(m) do you know who(m) you can trust implicitly?
    (10) Who was there who I am acquainted with?

CHART

Informant Responses on Set I[3]

Informants	Sentences			
	2	3	4	5
1	–	–	–	–
2	–	–	–	–
3	–	–	–	–
4	–	–	–	–
5		–	–	–
6		–	–	–
7		–	–	–
8		–	–	–
9		–	–	–
10		–	–	–
11		–	–	–
12		–	–	–
13		–	–	–
14		–	–	–
15		–	–	–
16		–	–	–
17			–	–
18			–	–
19			–	–
20				–
21				–
22				–
23				–
24				–
25				
26				
27				
28				
Deviants				
29		–		–
30	–	–		–
31		–	–	
32			–	

III. (11) Diomedes skipped town, which was stupid.
    (12) Diomedes skipped town, which I told him not to do.
    (13) Diomedes skipped town, which I wish he hadn't done.
    (14) Diomedes skipped town, which not too many people know.

IV. (15) It is said that Mussolini is planning on being reincarnated sometime next week.
    (16) It is believed by everyone that Ambrose is ill.
    (17) It was regarded as stupid that he should leave.
    (18) It is considered that Spiro likes Baklava.

V. (19) The baby was looked after by the nurse.
    (20) This matter will be seen to by our agents.
    (21) The booze was begged for by the team.

Apparently, informants who reject (2) through (5) have trouble with it-deletion under certain circumstances, as well as with certain complementizers. I am not at all sure what is going on here, but it appears that three rules are operating--one dealing with the simple, binary choice of whether or not it-deletion should apply; another concerned with complementizer choice, and the third involving choice of predicate, with all its semantic intricacies. For instance, most informants who rejected (3)-(5) accepted (22), and fewer, although still more than 50%, accepted (23):

    (22) It is necessary for John to leave.
    (23) John's leaving is necessary.

I have no idea why three rules operating in conflict should produce such a neat continuum of sentences.

Again, in the case of Set V, the limitations on passivization seem to be closely related to the tightness of binding between the verb and the preposition. The reason for the greater acceptability of (19) over (21) is obvious: for most speakers, look after is an idiom and its sequence of elements may not be altered or interrupted, e.g.,

    (24) *The nurse looked carefully after the baby.
    (25) The team begged earnestly for the booze.

Unfortunately, this tells us nothing of why (20) should be worse than (19), for see to is also an idiom, with the same restrictions as look after, so far as I know. Apparently, there is a ranking of tightness of binding for this sort of idiom. A very tentative hypothesis might be formed: if an informant will passivize a certain sentence of pattern (6), he will also accept passivizations of such sentences involving more tightly bound verb+preposition units; if, on the other hand, he rejects a given sentence, we may assume that he will also reject any more loosely bound units.

Set IV seems to be a part of a ranking of verbs of this particular type. Unfortunately, I have no data on informant behavior on restrictions other than passive which act on these verbs. It would be interesting to see if these restrictions are correlative with the passive restriction, as well as how they relate to other restrictions on passivization. Set III is a total mystery to me. I suspect that (11) through (14) may represent pieces of a logical continuum, most of which is missing here. As with all of the other items in these tests, a good deal more research needs to be done.

The defining characteristic of all this variation is the difference of intuition as to where to draw the line of acceptability. Given these sets of sentences, one finds very little disagreement on the order of the sentences. A few differences may arise on the question of whether sentence (f) is better than sentence (g), but there will be essentially no disagreement as to whether sentence (e) is better than (g). On the whole, such disagreements are not significant. The variations occur when different groups of informants disagree on the first sentence of the series which is to be classed as unacceptable. From that point, however, informants are predictable: if an informant accepts (f), for example, he must accept also (a) through (e); if he rejects (f), he must also reject (g) through (n).

Given this sort of limited data, we may assume either a continuum of sorts, with informants ranked from most conservative to most liberal, and with sentences ranked from most acceptable to least acceptable, or else we may posit dialects. Thus, Set I can be described as the overlapping of two dialects--one which restricts it-deletion, and a second which restricts the use of for-to and poss-ing complementizers. Alternatively, Set I may be described as a continuum formed by the interaction of the three rules of it-deletion, complementizer choice, and predicate choice.

Unfortunately, there are disadvantages to both these models. The test was limited, and I was able to include only a few sentences which are minimally different, e.g. (7) and (8). There are significant differences either of structure or semantics between most of the sentences tested. It would be interesting to test whether, given two sets of sentences, A and B, where there are significant syntactic or semantic differences between A and B themselves, but not between any sentences within the set, naive informants would group the sentences within one set into a continuum, or whether it would simply be the case that all the sentences in A would have approximately equal acceptability, but be decidedly better (or worse) than all the sentences in B. If the former occurred, the dialect model would clearly be too cumbersome, for we would have to set up dialects for nearly every possible sentence in English. Based on the admittedly limited evidence of (7) and (8), however, it may be tentatively hypothesized that informants will fall into groups based on some sort of minimal allowable difference between sentences.[4] If this is found to be

the case, the dialect model may be as acceptable as the continuum model to describe variation. The entire question, however, must be set aside until further research is done.

One of the problems with the continuum model is that it is not adequate to describe my results on sentences involving there-insertion or a definite pronoun referring to an indefinite antecedent. Thus, given (26) through (30), which are ranked in order of decreasing acceptability in my own speech, informants showed no evidence of producing any sort of continuum, except that between (29) and (30); (30) was never accepted by any informant who rejected (29), and (29) was never rejected by any informant who accepted (30). This, however, may be attributed to a hierarchy of verbs with which raising is permitted after there-insertion.

VI. (26) There is a man with a gun in this crowd.
     (27) There was a house torn down by Jones' Wreckers.
     (28) There was a fly ball hit by Yastremski.
     (29) Elaine expects there to be a chariot available
          for her.
     (30) Max believes there to be a cat on his bed.

There was no trace of a continuum with (27) and (28) and either (29) or (30). (26) was, of course, universally acceptable.

Three hypotheses are possible here: (a) There-insertion variation does not fit the continuum model, and must be described in another way; (b) There-insertion may fit the continuum model, but does not appear to do so in VI because of various other factors which are not immediately apparent. Thus, we may have some sort of "intersection of continua," which is throwing the results of VI off. Of course, we have indications of intersecting continua above, e.g., Set I, and the result is still a continuum. Furthermore, most sentence sequences are going to involve more than merely minimal differences, resulting in the intersection of two or more continua, and the result should be a continuum--if not, there must be another way to describe what is happening. (c) Informants are not used to dealing with there-insertion + passive or raising. As a result, they may not have been able to give true answers to Set VI, finding it difficult to decide whether they would ever say a sentence it has probably never occurred to them to use in casual speech. (c) appears to be the most logical of the three, for in my elicitation sessions, I noticed that there-insertion is almost exclusively used with the copula in the casual speech of most of my informants.

The case is somewhat different with sentences involving a definite pronoun with indefinite antecedent:

VII. (31) A dog is pawing at the door, isn't it?
     (32) A man with a brown plaid sports coat just ran by
          here, didn't he?
     (33) One can't expect to succeed at everything, can he?

Since tag questions are exceptionally easy to elicit during casual conversation, I was able to check for accuracy with no trouble. Informants tended to be unnervingly consistent through all parts of the test. Interestingly enough, there was very little variation between (32) and (33), and what variation there was did not take the form of a continuum.

The only was to describe these results is with a dialect model, positing four dialects: (1) acceptance of all such sentences; (2) acceptance of only those with non-neuter pronoun; (3) acceptance of only those with neuter pronoun; and (4) rejection of all such sentences. Within these dialects, if that is what they are, there may be continua. Furthermore, another factor seems to be in operation here. For instance,

> (34) A man with a brown plaid sports coat and blond hair of about shoulder length, who looks like he could be my brother, just ran by here, didn't he?

tends to be much better for informants in (3) and (4) than (32) and (33) are. The Zimmerpunkt may be effecting this, or increased description may simply render the subject NP more definite in the informant's mind. In any case, the continuum model is hopelessly inadequate to describe the data as it occurs here.

Our third alternative, then, is to assume that certain variations can be described by the continuum model, and that others require an assumption of dialect variation. Unfortunately, the limited research in this study is not sufficient to allow a prediction of which sorts of variation will fit which model.

In addition, a number of other questions must be answered, among them the question of whether informants tend to display the same behavior on tests for different items. I have been able to find a certain amount of correlation, although by no means as satisfactory as I would like, for certain informants; slightly more than 30% of my informants behaved consistently in three or more out of the seven structures tested; nearly all of them behaved consistently in at least two non-related structure tests. Consistent behavior is defined here as either consistent liberality (acceptance of all or nearly all the sentences), consistent conservatism (rejection of all or nearly all of the sentences in a set) and consistent middle-of-the-roadism (equal acceptance and rejection of sentences in a set). When sufficient detailed work is done on large numbers of related and unrelated structures, we may be able to determine if "dialects" (in the sense of discrete groups of informants who tend to behave similarly on large numbers of structures, either related or unrelated) do indeed exist. Again, it must be determined whether the variation can be defined in terms of standard syntactic rules or semantic hierarchies or whatever, or whether it is a product of some sort of as-yet-undreamt-of sentence acceptability ranking rules.

Since tag questions are exceptionally easy to elicit during casual conversation, I was able to check for accuracy with no trouble. Informants tended to be unnervingly consistent through all parts of the test. Interestingly enough, there was very little variation between (32) and (33), but what variation there was did not take the form of a continuum.

The only way to describe these results is with a dialect model, positing four dialects: (1) acceptance of all such sentences; (2) acceptance of only those with non-neuter pronoun; (3) acceptance of only those with neuter pronoun; and (4) rejection of all such sentences. Within these dialects, if that is what they are, there may be continua. Furthermore, another factor seems to be in operation here. For instance,

> (34) A man with a brown plaid sports coat and blond
> hair of about shoulder length, who looks like
> he could be my brother, just ran by here, didn't
> he?

tends to be much better for informants in (3) and (4) than (32) and (33) are. The Zimmerpunkt may be affecting this, or increased description may simply render the subject NP more definite in the informant's mind. In any case, the continuum model is hopelessly inadequate to describe the data as it occurs here.

Our third alternative, then, is to assume that certain variations can be described by the continuum model, and that others require an assumption of dialect variation. Unfortunately, the limited research in this study is not sufficient to allow a prediction of which sorts of variation will fit which model.

In addition, a number of other questions must be answered, among them the question of whether informants tend to display the same behavior on tests for different items. I have been able to find a certain amount of correlation, although by no means as satisfactory as I would like, for certain informants; slightly less than 30% of my informants behaved consistently in three or more out of the seven structures tested; nearly all of them behaved consistently in at least two non-related structure tests. Consistent behavior is defined here as either consistent liberality (acceptance of all or nearly all the sentences), consistent conservatism (rejection of all or nearly all of the sentences in a set) and consistent middle-of-the-roadism (equal acceptance and rejection of sentences in a set). When sufficient detailed work is done on large numbers of unrelated and related structures, we may be able to determine if "dialects" (in the sense of discrete groups of informants who tend to behave similarly on large numbers of structures, either related or unrelated. Again, it must be determined whether the variation can be defined in terms of standard syntactic rules or semantic hierarchies or whatever, or whether it is a product of some sort of as-yet-undreamt-of sentence acceptability ranking rules.

It is thus apparent that there are at least some sorts of syntactic and semantic variation among speakers which can be logically predicted, and which may be the result of some sort of ranking, producing a continuum model, or of gaps between dialects.

Of course, there are many other types of variation. For instance, Paul Neubauer has pointed out that there appear to be three primary dialects for the verb pretend, while Guy Carden has argued for the existence of the NEG-Q/NEG-V dialects. Unfortunately, when I tried to replicate their results, I found that informants did not remain consistent under repeated testing, either with different tests, or with the same test repeated after a lapse of at least a week. Both with the NEG-Q/NEG-V dialects and the pretend dialects, where a particular predicate has different limitations for different speakers, informants tended to switch when the alternative was pointed out to them; in some cases, their responses appeared to be almost totally conditioned by the particular environment they happened to be thinking of at the time. After repeated testing, I discovered that of the 13 informants who could not accept:

> (35) John is pretending that he is not Napoleon.

in the insane-asylum context (where John believes he is Napoleon, but knows he will not be released until he convinces the doctors that he does not believe that he is Napoleon), 9 had not been able to follow the story, or had refused to accept the context as a possible one; they therefore rejected (35) because "there is no way it could be a real-life sentence."

The additional problem that arises with dialects of the pretend sort is that there are any number of predicates in English, at least, that have varying connotations for different speakers. For instance choose, as in:

> (36) We may choose to be good, but we are incapable
> of being good, for we are mere linguistics
> people.

is unacceptable to many people for whom choose implies both the capability and the intention of performing the "chosen" act. There are entirely too many predicates with this sort of connotative difference, and if we are to accept all such differences as dialectally determined, we shall have to establish dialects (probably non-correlative) for a good many predicates in English. This is, I believe, undesirable, and will probably not add very much to a description of the language.

Furthermore, if we are to accept variations in which speakers do not maintain consistency from one test to another, we shall have to define our only possible informants as those who have never before been used in a linguistics study.

Variations like NEG-Q/NEG-V and the _pretend_ dialects do exist, of course, and must be studied. I doubt, however, that they constitute dialects, in the standard sense of the term.

This leaves us with the hypothesis that the only sorts of variation, as far as we now know, which maintain informant consistency are those based on relatively uninteresting syntactic and semantic matters, like those mentioned above. The very fact that consistency seems to be maintained only with such low-level, uninteresting, trivial stuff is in itself interesting, for, given more research, it may allow us to at least describe a few of the problems that occur when a linguist has gone to a great deal of trouble to set up a hypothesis, and a mere naive informant spoils the whole thing by invoking the bane of linguistics people--"Not in _my_ dialect, you don't!"

Footnotes

[1]In this scale, 2 is equivalent to ?; 3 to ?*; and 4, of course, to *.

[2]The purpose of (a) and (c) was primarily to test the informant's reliability on (b). In addition, there were a number of dummy sentences in (b), which were used to test the informant's comprehension of and compliance with the instructions. Of the original 39 informants to whom I gave the test, 7 were eliminated for unreliability (lack of consistency between the various parts of the test), and 2 were eliminated for incorrect responses on the dummy sentences on (b), e.g., giving a rating of 1 to a sentence violating a Ross Constraint.

[3]For reasons which I cannot discuss in this paper, I do not believe that naive informants are capable of making the four-level distinction required of them in Part (b) of this test. After several preliminary tests which I discussed with a different set of informants, I learned that they tend first to make the distinction only between acceptable and unacceptable, rather than between acceptable, unacceptable and odd; as a result, they view the primary pattern as being 1-2 versus 3-4, rather than 1 versus 2-3 versus 4. I have therefore reduced their responses on (b) to the simple form which appears on the chart. The minus, of course, indicates that the informant gave the sentence a rating of 3 or 4, while a blank indicates that he found the sentence acceptable, giving it a 1 or a 2.

[4](7) and (8) are at least minimally different, probably due to the appearance of the idiom let out in (8). In my speech, (7') is good, while (8') is only slightly better than (8) itself.

    (7') A performer who is dressing now will appear shortly.
    (8') I let the cats which were meowing out.

However, most of my informants saw no great difference in acceptability between (7) and (8); the difference is apparently sufficiently minimal to be ignored by most informants.