QUESTION-ANSWERING SYSTEMS

Emily Pope
University of Washington

Yes-no questions can be either positive or negative, and their answers, too, can be either positive or negative. This means that there are basically four categories of answers to such questions. However, most languages do not have four different answering morphemes or patterns. English, for example, uses "yes" for positive answers, whether the question was positive or negative, and "no" for negative answers, as in (1) and (2).

(1)	Q.	{Is it hot today? It's hot today, isn't it.}	
	A. 1.	Yes (it is hot today).	Positive Agreement (PA)
	2.	No (it isn't hot today).	Negative Disagreement (ND)
(2)	Q.	{Isn't it hot today? It isn't hot today, is it.}	
	A. 1.	No (it isn't hot today).	Negative Agreement (NA)
	2.	Yes, it is (hot today).	Positive Disagreement (PD)

This is a positive-negative answering system. (An answer is negative if it contains a sentential negation in its highest clause, and positive if it doesn't.)

Japanese, on the other hand, uses "hai" for a positive answer to a positive question and a negative answer to a negative question, and "iie" for a negative answer to a positive qeustion and a positive answer to a negative question, as in (3) and (4).

(3)	Q.	{Kyoo-wa atu-i des-u ka? Is it hot today? Kyoo-wa atu-i des-u ne. It's hot today, isn't it.}	
	A. 1.	Hai (it is hot today). Right (it is hot today).	PA
	2.	Iie (atuku-wa arimasen). Wrong (it isn't hot today).	ND
(4)	Q.	{Kyoo-wa atuku-na-i des-u ka? Isn't it hot today? Kyoo-wa atuku-na-i des-u ne. It isn't hot today, is it.}	
	A. 1.	Hai (soo des-u ne). Right (it isn't hot today).	NA
	2.	Iie, kyoo-wa atu-i des-u. Wrong, it is hot today.	PD

This is an agreement-disagreement system. (An answer is agreeing if it matches the question with respect to negativity, and disagreeing if it doesn't.)

These two types of systems are not the only ones that occur. Many languages have a special word for expressing positive

disagreement--for example, "doch" in German, as in (6A2), "si" in French, and "jo" in Norwegian.

(5) Q. Geht's dir gut?
A. 1. Ja. PA
2. Nein. ND

(6) Q. Geht's dir nicht gut?
A. 1. Nein. NA
2. Doch. PD

This phenomenon shows up in English and Japanese as an inability to reduce the answer as fully in positive disagreement as in the other categories (cf. (2A2) and (4A2)).

The reason positive disagreement so often has to have a special word is that it is the most semantically difficult or marked of the four categories. First let me explain what I mean by semantic difficulty.

Given that sentences have both syntactic form and semantic content, semantic difficulty is made up of the following: 1) semantic content, 2) incongruity of syntactic form and semantic content. We will regard disagreement as having semantic content, while agreement does not.[1] Thus we must assign ND and PD each one unit of semantic difficulty on account of their content. We will regard negativity as being congruous with disagreement, and positivity with agreement, while negativity is not congruous with agreement nor is positivity with disagreement. Thus we must assign NA and PD each one unit of semantic difficulty on account of their incongruity of form and content. This leaves PD, with two units of difficulty, the most difficult category.

Let me try to defend this definition of semantic difficulty by comparing it to an alternative definition--namely, one which says that semantic difficulty is simply made up of semantic content (disagreement) and syntactic form (negativity). This would make ND the most difficult category. The immediate argument for the definition we have chosen is that it gives the desired result--namely, that PD is the most difficult category, which we have seen to be the case. But I think that there is a deeper reason. Saying that disagreement is marked is a fairly normal sort of semantic rule. There is no doubt that the act of disagreeing is more marked than the act of agreeing. It constitutes a departure from what is expected. But agreement and disagreement are the semantic content of positive and negative. When this content has been extracted out and marked as being difficult or not, it is not surprising that utterances are not further marked, redundantly, as it were, for merely being negative in form. If a question is negative in form, the easy answer is the one that is also negative in form. Thus our second definition of difficulty would have little semantic justification.

The second component of semantic difficulty--incongruity of form and content--is just a way of expressing the semantic link among negativity, rejection, denial, and disagreement, as opposed

to positivity, acceptance, acquiescence, and agreement.

The notion of semantic difficulty has two uses. I predict that 1) the more semantically difficult a category, the later it will be acquired and used by children, and 2) semantically difficult categories will be the last to collapse morphologically or grammatically with other categories. My first contention is supported by the fact that children learn to use PA and ND sooner than they learn to use NA, and in particular PD. The data in Bellugi's thesis (1967) show this clearly.

When children first start using "yes" and "no" as answers, "yes" is used for PA, "no" for ND. NA is usually implied rather than expressed, and PD is expressed by an affirmative sentence, without "yes". The following examples are all from or slightly before Bellugi's period B (ibid.).

PA: Adult: Are you going to be little?
Adam: Yes.
Adult: When?
Adam: Friday.

Adult: Would you like to have your lunch right now?
Eve: Yeah.

ND: Mother: Is that my grape juice?
Eve: No, that Eve grape juice.

Father: Can I have my money back?
Sarah: No, my money.

NA Mother: Your pencil didn't break.
Eve: Only Fraser's.

Mother: Oh, we don't have any bread, Eve.
Eve: We hab buy some.

PD: Mother: Oh, I don't think you like the water.
Sarah: I like water.

Mother: You don't know how to swim.
Sarah: I how swim.

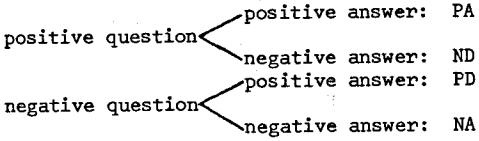
Bellugi herself points out that PD is the last category acquired. She says, of PD, "The mother says 'Oh, it doesn't'. The child counters with 'Yes, it does'. This is rare in the early periods. The children are much more likely to ignore, persist behaviorally, or implore, than they are to disagree verbally." This statement does not apply to ND, which is acquired rather early.

Now let me elaborate on my second contention--that semantically difficult categories will be the last to collapse morphologically or grammatically with other categories. This means that we can make some preliminary hypotheses about what question-answering

systems are possible--i.e., which ones could turn up in a human language and which could not. There are fifteen possible arrangements of the four categories, as follows (dashes mean that the terms they join are all represented by the same form):

A.	PA; PD; NA; ND	low probability
B.	PA; PD; NA-ND	very high probability
C.	NA; ND; PA-PD	very low probability
D.	PD; ND; PA-NA	very high probability
E.	PA; NA; PD-ND	very low probability
F.	PA; ND; PD-NA	impossible
G.	PD; NA; PA-ND	impossible
H.	PA-PD; NA-ND	high probability
I.	PA-NA; PD-ND	high probability
J.	PA-ND; PD-NA	impossible
K.	PA; PD-NA-ND	impossible
L.	PD; PA-NA-ND	impossible
M.	NA; PA-PD-ND	impossible
N.	ND; PA-PD-NA	impossible
O.	PA-PD-NA-ND	impossible

Let me again point out what the four categories mean:



Now, it is obvious that no language will express all four categories in exactly the same way, for this would rob answers of all semantic content other than "I respond" or some such. So (O) is not a possible system. Similarly, no language will express three of the categories in exactly the same way, for this would mean that although one type of question (positive or negative) would have two possible answers (agreeing and disagreeing), the other type of question (negative or positive) would have only one answer--again, an impermissible loss of semantic content. So (K), (L), (M), and (N) are impossible systems. By the same token, PD and NA cannot merge, for then negative questions would have only one answer, and PA and ND cannot merge, for then positive questions would have only one answer. So (J), (F), and (G) are impossible systems.

This leaves us with seven interesting systems--(A), (B), (C), (D), (E), (H), and (I). All of them are possible. Because PD is the most semantically difficult category, we predict that it will be the category least likely to merge with another. It will be more likely to merge with a category if the other two have merged than if they have not. On the other hand, a system with no mergers is a bit inefficient, conveying more information than is really necessary. These factors give roughly the following hierarchy: (B) and (D) are the least difficult systems, (H) and (I) are next,

then (A), and (C) and (E) are the most difficult or least likely to occur.

A brief survey of occurring morphological systems gave the following results:[2]

(B): German, French, Scandinavian, Tigrinya
(D): Latvian, Harari, Swahili
(H): English, Hebrew, Spanish
(I): Japanese, Hidatsa, Navajo, Amharic
(A): Chaha, 16th century English[3]
(C): none
(E): Soddo

this gives some inductive confirmation to my ranking of the system types in terms of probability. (B), (D), (H), and (I)-types are all fairly common, (A)-types less so, and (C) and (E)-types are virtually non-existent.

Actually, most systems will be more complex than indicated, making different distinctions in different ways. In English, for instance, the distinction made in (I) can be made intonationally, the distinction made in (H) is made morphologically, and the additional distinction made in (B) is made by the syntactic rule of tag deletion.

I have shown that the rules determining agreement and disagreement are rather simple-minded, straightforward, formal syntactic matching rules. I have pointed out that what I mean by agreement is that the negativity of question and answer is the same, and by disagreement, that they are different. Thus an agreeing answer to a yes-no question is phrased in the same way as the question, save for Subject-AUX Inversion, and a disagreeing answer differs from the question only by virtue of this rule and the addition (to a question which has none) or deletion (from a question which has one) of a sentential negation. This is true regardless of whether the questioner is really expecting a positive or negative answer. In English, a question like "Haven't you written some books?" expects a positive answer, yet a positive answer still counts as disagreement and should take the form "Yes, I have". I admit that the tag is less obligatory here than in the positive answer to "Haven't you written any books?", which expects a negative answer, but a plain "Yes" answer is still insufficient. In Swedish, for instance, the situation is quite parallel. Negative questions often clearly expect positive answers, but, in these cases, positive answers still take the PD form "jo" rather than the PA form "ja". The realities of the semantic situation, which do not always accord with those of the syntactic situation, enter in only at the level of degrees of grammaticality. The rules, then, determining whether an answer is agreeing or disagreeing are as follows, where Sneg means sentential negation.

(7) ∝Sneg question, -∝Sneg answer → disagreement
(8) ∝Sneg question, ∝Sneg answer → agreement

The rules which determine whether an answer is positive or negative, i.e., whether the sentential answer should be preceded by and reducible to "yes" or "no", are equally straightforward but allow latitude in a different way. The straightforward part is: the answer is negative if it has a sentential negation in its topmost clause, positive if it does not. The rules are (9) and (10).

(9) Sneg answer -> negative
(10) -Sneg (=∅) answer -> positive

According to these rules, the proper answers to the following questions are the two written beneath them.

(11) Q. Do you like nobody?
N. No (I like nobody) [or] (I don't like anybody).
P. ?Yes, I like somebody.
(12) Q. Does nobody like you?
N. No (nobody likes me).
P. ?Yes, somebody likes me.
(13) Q. Don't you like anybody?
N. No (I don't like anybody).
P. ?Yes, I like somebody.
(14) Q. Doesn't anybody like you?
N. No (nobody likes me).
P. ?Yes, somebody likes me.

There is, however, a second set of rules determining positivity and negativity, which often gives results contradictory to the first set ((9) and (10)). There are several reasons for the existence of this second set of rules. One reason is that the PD answers above are even more unwieldy than is usual, because they require not just the subtraction of a negative, but also the operation of an 'any -> some' discourse rule which involves a semantic change as well as a syntactic one.

A second reason is that the first set of rules treats all of the above questions as negative questions, with no real difference between (11) and (13) or (12) and (14). In other words, any negative that can be is interpreted as a sentential negation, and all questions with Sneg's are negative. However, it is well-known that many negatives in English are ambiguous as to scope. There is a reason for having both (11) and (13). The second set of rules takes these facts into account, and says that an answer is positive if there is no discourse sentential negation, negative if there is.

Here I must explain what I mean by discourse sentential negation. In a set of related sentences, the first sentence is an instance of discourse sentential negation if only the obligatory negative placement rules have applied to it. The obligatory rules are the one that attaches negatives to the AUX and the one that incorporates negatives into indefinites which precede the AUX. Since the second rule follows Subject-AUX Inversion, the first of the two rules is the only one that will apply to yes-no questions. Consequently, only negative

yes-no questions (those with a negative attached to the AUX) are instances of discourse sentential negation. This means that (15) but not (16) is an instance of discourse sentential negation.

(15) Isn't anybody home?
(16) Is nobody home?

Now, in determining whether the second sentence (the answer) is an instance of discourse sentential negation, the first sentence (the question) must be taken into account. Any negative which did not count as a discourse sentential negation in the first sentence does not count as one in the second, either. Any negative which did count counts in the second sentence as well, and is re-placed by the two obligatory rules. This means that "Nobody is home" is an instance of discourse sentential negation, and so introduced by "no", in (17A) but not (18A).

(17) Q. Isn't anybody home?
A. No, nobody is home.
(18) Q. Is nobody home?
A. Yes, nobody is home.

If this distinction seems a bit ephemeral, notice the marked unnaturalness of (20A) as opposed to (19A). Both are supposedly derived from (21). But in (20A), the "yes" forces the conclusion that the first negative ("no dogs") is not discourse sentential and so cannot be dissociated from the constituent to which it is attached, as it has been in the second clause. However, (19B) and (20B), derived from (22), are equally acceptable, because the second clause is ambiguous, and the negative can be interpreted as discourse sentential or not, as required.

(19) Q. Don't any dogs like you?
A. No, and neither do any cats.
B. No, and no cats do either.
(20) Q. Do no dogs like you?
A. ?*Yes, and neither do any cats.
B. Yes, and no cats do either.
(21) No dogs like me, and neither do any cats.
(22) No dogs like me, and no cats do either.

I have made the definition of discourse sentential negation dependent on the distinction between obligatory and optional negative placement rules. I said that in both first and second instances of discourse sentential negation, only the obligatory rules applied. When an optional rule applies in the question, the question is still acceptable, but it is not an instance of discourse sentential negation, and an answer with the negative in the same place is not one, either. Any negative which was not present in the question, but appears in the highest clause of the answer, is also discourse sentential.

With the explanation of discourse sentential negation that has been given, the second set of rules determining positivity and negativity can be given as (23) and (24) (DSneg = discourse sentential negation).

(23) DSneg answer --> negative
(24) -DSneg (=Ø) answer --> positive

These rules have some of the same problems with double negatives as the first set of rules. They can easily enough handle examples like (25A). Here a DSneg and a non-DSneg coexist, giving, by (23), a negative answer.

(25) Q. Do you like nobody?
A. No, I don't like nobody.

But consider an example like (26). In (26N), the DSneg that has been added to indicate disagreement has no way of moving onto the AUX (perhaps because of a crossover constraint), and so the reduced "no" answer is a bit strange, when interpreted according to (23). (26M) represents its meaning. In (26M), the DSneg and the non-DSneg have cancelled each other, but the "no" remains.

(26) Q. Does nobody like you?
P. Yes (nobody likes me).
N. ?No (not (nobody likes me)) $\Longrightarrow$ Nobody doesn't like me.
M. No, somebody likes me.

These examples show that, since they are generated in different positions, DSneg's and non-DSneg's may coexist in one S, at least in deep structure, before the negative placement rules start operating, and that the choice of "yes" and "no" is based on these early structures. However, two DSneg's may not coexist on the same S. There is only one "slot" to be filled, and it may only be filled once. Having two DSneg's on one S would be equivalent to having two "that" complementizers on one S, or two "the" determiners on one NP.

The way to "negate" a sentence with a DSneg is to take away the DSneg--to delete it. This results in a positive sentence, according to (24)--a sentence which take "yes". This is what has happened in (27P).

(27) Q. Doesn't anybody like you?
N. No (not (anybody likes me)) $\Longrightarrow$ Nobody likes me.
M. *(not not (anybody likes me)) $\Longrightarrow$ (Ø (anybody likes me)) $\Longrightarrow$
P. ?Yes, somebody likes me.

Here again, (27P) is just as strange as (13P) and (14P), and for the same reason. The switch from "anybody" to "somebody" is not really a smooth and automatic one. That is, although a double negative implies a positive and vice versa, they are not completely

equivalent linguistically. (28) seems to me to be much less tautologous than (29).

(28) It is not the case that nobody likes me. Somebody likes me.
(29) I haven't ever seen such a mess. Never have I seen such a mess.

Naturally, the switch from double negative to positive is easiest when a positive polarity item is involved; hardest when a negative polarity item is involved, as in (27). It is also more difficult when both negatives are, or would have been, DSneg's than when one or both are non-DSneg's. Thus the problem arises in fewer cases with the second set of rules.

I have given examples to show how the two sets of rules determining positivity and negativity work, even in the most difficult cases. We have seen that, as with agreement and disagreement, the times when confusion arises have to do with the questioner's bias. The two sets of rules conflict only in a subclass of the cases when the normal opposite bias of a question is reversed. Specifically, they conflict only in the answers to questions containing negatives which are ambiguous as to whether or not they are sentential negatives. In these instances, if we consider only questions with AUX-attached negatives to be negative, it is the first set of rules that acts like a positive-negative system, and the second set of rules acts like an agreement-disagreement system. Elsewhere, both sets act like positive-negative systems.

Navajo is like English in vacillating between a positive-negative answering system and an agreement-disagreement system. There, however, the favored system is agreement-disagreement. The paradigm is as follows:

(30)	Q.	Ch'éénísh díníyá?	Are you tired?
	A.	Aoo', ch'één déyá.	Yes, I am tired.
		Dooda, doo ch'één déyáada.	No, I'm not tired.
(31)	Q.	Doosh ch'één díníyáada?	Aren't you tired?
	A.	Aoo', doo ch'één déyáada.	Yes, I'm not tired.
		Dooda, ch'één déyá.	No, I'm tired.
	A'.	Aoo', ch'één déyá.	Yes, I'm tired.
		Dooda, doo ch'één déyáada.	No, I'm not tired.

The (A) answers are favored over the (A') answers.

It is interesting that Navajo has only sentential negation. One may ask "Isn't anybody coming?" but not "Is nobody coming?". The same is true of Japanese, which also has an agreement-disagreement system. The same is true of Hidatsa, which also has an agreement-disagreement system. I think that further research will bear out my finding that languages with only sentential negation tend to have agreement-disagreement systems, and languages with both sentential and NP negation tend to have positive-negative systems. The latter tendency is stronger than the former, for I know of counterexamples, such as Finnish, to the former. I think the closest thing to a universal here is (32), which may also be

phrased as in (33).

(32) Only languages with only sentential negation can have agreement-disagreement question-answering systems.
(33) Languages with both sentential and NP negation cannot have agreement-disagreement question-answering systems.

A reason for these results might be suggested by the second set of rules for determining positivity and negativity--(23) and (24). There we saw that when negatives not attached to the AUX are not regarded as sentential negatives, the system becomes more like an agreement-disagreement system. When a language has only sentential negation, it is always clear whether a question is negative or positive, i.e., whether it does or does not contain a sentential negation. And only when this is the case is an agreement-disagreement system a possible option.

We have seen that PD is the most semantically difficult of minimal answers to yes-no questions. This difficulty is reflected in many ways, showing that the way formal systems are constructed, and the order in which children acquire and use them, are both influenced by semantic considerations. We have also seen that one formal system, such as negation, may determine the shape of another, such as the question-answering system.

FOOTNOTES

1. It may be objected that, since what I call disagreement is syntactic disagreement, and not necessarily semantic disagreement, saying that disagreement has semantic content is not really justified. Those readers to whom this seems a serious objection may substitute "difficulty" for "semantic difficulty" and "incongruity of question and answer", or more specifically, "negativity switch between question and answer" for "semantic content" in the argument which follows.
2. All of these systems are discussed, and proper credit is given to my many informants, in my forthcoming book, Questions and Answers in English, Mouton and Co.
3. ENGLISH (16th Century)

	Agreement	Disagreement
Positive	yea	yes
Negative	no	nay

English has changed from a relatively difficult (A)-type system to a less difficult (H)-type system. A remnant of the old system may be seen in the fact that "yeah" is an optional variant of "yes" in modern English only in the case of PA, not PD, and "nah" or "naw" is a variant of "no" for ND. It is interesting that it is the answers to negative questions ("yes" and "no") that have survived.

BIBLIOGRAPHY

Bellugi, U., 1967, The Acquisition of Negation, unpublished Doctoral dissertation, Harvard University.

Gudmundsson, H., 1970, "Yes and No in Icelandic," in H. Benediktsson, The Nordic Languages and Modern Linguistics, Visindafelag Islendinga, Revkiavik.

Lakoff, R., 1969, "Some Reasons Why There Can't Be Any Some-Any Rule", Language, 45.3.

Leslau, W., 1962, "'Yes' and 'No' in the Ethiopian Languages," Language, 38.

Pope, E., 1971, "Answers to Yes-No Questions," Linguistic Inquiry,2.1.

Pope, E., forthcoming, Questions and Answers in English, Mouton, The Hague.