

- Nunberg, G. (1979). The non-uniqueness of semantic solutions. *Linguistics and Philosophy* 3(2), 143–184.
- Putnam, H. (1975). The Meaning of Meaning. In H. Putnam, *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Sadock, J. M. (1978). On testing for conversational implicature. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics*. New York: Academic Press. Pp. 281–297.

Toward a Taxonomy of Given–New Information

Ellen F. Prince

1. ON THE CONVEYING OF INFORMATION IN LANGUAGE

It is a truism that, when people use language naturally, they are usually attempting to convey information. Occasionally, the only information conveyed is of a purely social or ritual nature; such is the case in (1), where each speaker is doing little more than informing the other that s/he is aware of the other's presence, is not hostile, and will part ready to come into friendly contact again at some future time:¹

¹ Erving Goffman has pointed out (personal communication) that other types of information, in addition to ritual information, are conveyed in indirect ways—for example, meta-communication information. Furthermore, what makes ritual information ritual information is not the information itself but the indirect (ritualized) way in which it is presented. That is, if A makes the following utterance to B, A's utterance will exemplify what I shall call, for lack of a better term, "objective" information:

[A and B are passing each other on the street]

A: *I am aware of your presence, I am not hostile, and I shall part ready to come into friendly contact again at some future time.*

B: *Damn sociolinguists . . .*

The matter is obviously complex but, fortunately, not crucial to the subject of this chapter. See Goffman (1967) for a discussion of different types of information conveyed.

- (1) [A and B are passing each other on the street]
 A: *Hi.*
 B: *Hi.*

Utterances that convey only ritual information in the sense used here are typically short, (virtually) asyntactic, and relatively few in number; with most utterances, it is impossible for a speaker/writer to avoid conveying, in addition (or instead), "objective" information, although there is, of course, nothing to prevent that objective information from being banal or even silly.² Thus (2) conveys, in addition to the ritual information analogous to that conveyed in (1), the objective information that (A thinks that) it is a very hot day:

- (2) [A to B at bus stop on a summer day]
What a scorcher!

One presumably universal feature of natural language is that the objective information conveyed is not conveyed on a single plane. That is, there is an INFORMATIONAL ASYMMETRY in that some units seem to convey or represent "older" information than others. Given-new distinctions can be found on different levels—the sentence, the discourse, the participants' discourse-models—as will be seen in what follows. On all levels, however,—and perhaps this is not only universal, but also distinctive of human language—the crucial factor appears to be the tailoring of an utterance by a sender to meet the particular assumed needs of the intended receiver. That is, information-packaging in natural language reflects the sender's hypotheses about the receiver's assumptions and beliefs and strategies.

² There are some apparent counterexamples to this. One, the monologue, can however be construed as a sender conveying information to a receiver, the only aberration being that a single individual is playing both roles. It would be of interest to study naturally occurring monologues to see just how the information they convey is structured and distributed. Another is the utterance whose utterer believes that there is in fact no receiver, such as a lecture to an obviously distracted audience, a dissertation written for a presumed uninterested and lazy adviser (see Chomsky, 1975, p. 61 for a poignant description of such an experience). In such cases, the utterance itself, as an object (e.g., a lecture, a dissertation), tends to have ritualistic import, which accounts for why the sender bothers to produce it. And part of the ritual seems to be that the sender acts AS THOUGH there were a receiver. Other, less interesting, cases are those where the sender believes s/he is conveying information to a receiver but is in fact mistaken (e.g., unbeknownst to the sender, the telephone has been disconnected or God does not exist or an unopened letter is destroyed in a fire). Such cases are not relevant here because it is only the sender's assumptions at the time of producing the utterance that are relevant to the structure and distribution of given-new information, as we shall argue in what follows.

In what follows, I shall discuss briefly these three different levels of given-new information and some of the major literature associated with each (Section 2), and then I shall examine more closely one of these levels and propose a provisional taxonomy for it (Section 3). Following will be an illustrative analysis of two naturally occurring texts—one an informal oral narrative, the other a formal written didactic piece—and a comparison of the two with respect to the structure and distribution of given-new information in them (Section 4). Finally, I shall suggest further areas of research that seem particularly promising (Section 5).

2. "GIVEN-NEW"

The general notion of given versus new information figures prominently in much linguistic literature, under that name or under one of its aliases: old-new, known-new, presupposition-focus, and so on. It has been invoked both in the explication of many sentence-level phenomena (e.g., Gapping, Dative, Pronominalization, Left and Right Dislocation, sentential subjects, *it*-clefts, *wh*-clefts, Topicalization) and in the explication of how discourses are structured and understood. Bolinger, Chafe, Chomsky, Clark, Horn, Jackendoff, and Kuno are just a few of the linguists who have found the notion relevant to their work.

Unfortunately, however, this intuitively appealing notion has never received a satisfactory characterization that would enable a working linguist to not only invoke it but to actually put it to use. In fact, if one considers the definitions that have been presented, one discovers that there is not one notion involved but (at least) three.³ Of all the linguists who have used these terms or their synonyms, those that are perhaps the most strongly associated with the notion are Chafe, Clark and Haviland, Halliday, and Kuno. It is rather surprising, then, that, when their discussions on the subject are closely examined and compared, it becomes evident that no two of them mean quite the same thing, and that, in some cases, the differences are quite large. Space does not here permit a thorough investigation of the literature involved, so I shall simply discuss briefly the three levels of givenness that their notions may be thought of as belonging to, in order that the reader may have a clearer idea of what we shall then be examining.

³ As added evidence of the gravity of the situation, let me mention that the Old/New Information Workshop held at Urbana, Summer 1978, was quickly and quite appropriately dubbed the "Mushy Information Workshop."

2.1 Givenness_p: Predictability/Recoverability

In speaker-hearer terms, givenness in the sense of predictability/recoverability may be described as in (3).

- (3) **Givenness_p**: The speaker assumes that the hearer CAN PREDICT OR COULD HAVE PREDICTED that a PARTICULAR LINGUISTIC ITEM will or would occur in a particular position WITHIN A SENTENCE.

This type is represented by what Kuno (1972, 1978, 1979) calls "old-new" information and what Halliday (1967) and Halliday and Hasan (1976) call "given-new" information, although the two notions are defined differently and what is old for Kuno is not necessarily given for Halliday. In particular, Kuno defines old-new information in terms of recoverability: "An element in a sentence represents old, predictable information if it is recoverable from the preceding context; if it is not recoverable, it represents new, unpredictable information [1978, pp. 282-283]." Clearly, recoverability correlates with deletability, but it does not address many other phenomena that are felt to be related to given-new information, for example, the difference in status between a nondeletable pronoun and its antecedent, as in (4).

- (4) *Mary paid John₁ and he₁/ \emptyset bought himself_i a new coat.*

Compare (4) with (5):

- (5) *John₁ paid Mary and he₁/ \emptyset bought himself_i a new coat.*

If predictability/recoverability were the only criterion by which to judge the newness of an NP, then *he* in (5) would be old whereas *he* in (4) would be new, although, in some sense, they look very much the same age. Thus Kuno (1972) adds a second distinction, that of anaphoric-nonanaphoric, to which we shall return in Section 2.3.

For Halliday (1967), given-new is defined quite differently, in terms of intonation: An intonationally marked or unmarked focus identifies what is new; given is defined as the complement of a marked focus. Thus, in an information-unit with unmarked focus, nothing is given (p. 208). Halliday predicts that what is thus labeled new is "information . . . that the speaker presents . . . as not being recoverable from the preceding discourse [p. 204]." In Halliday and Hasan (1976), new is described as in Halliday (1967), and given is described as "expressing what the speaker is presenting as information that is recoverable from some source or other in the environment—the situation or the preceding text [p. 326]." It is still optional, that is, an information-unit may contain no given information.

Although Halliday's recoverability looks at first blush a good deal like

Kuno's predictability, the two will make different predictions if Halliday's restriction of given to sentences with marked focus is upheld.⁴ For example, if (5) were uttered with normal intonation (i.e., unmarked focus), *he* would be old for Kuno and neither given nor new for Halliday.

Halliday's notion does, however, relate directly to a certain class of otherwise mysterious phenomena, illustrated by the famous example in (6a):

- (6) a. *John called Mary a Republican and then SHE insulted HIM.*
b. *John called Mary a Republican and then she insulted him.*

Following Halliday, *SHE* and *HIM* in (6a) would be marked foci (assuming one may have more than one marked focus per information-unit), resulting in the remainder of the unit, *insulted*, being identified as given, that is, recoverable from the preceding context. Now the only way in which it could be recoverable is if it is somehow equivalent to *called a Republican*. Compare with (6b), where no such equivalence is inferred, and (7), where the contrast is even more dramatic than in (6).

- (7) a. *John_i called Sam_j a Republican and then HE_j insulted HIM_i.*
b. *John_i called Sam_j a Republican and then he_i insulted him_j.*

Here the addition of the marked foci effect two changes: First, because *HE* and *HIM* in (7a) are marked as being new, that is, unrecoverable (in the position which they occupy), they cannot refer to *John* and *Sam*, respectively. Being pronouns, they do, however, refer to some entities already in the discourse-model (about which more will be said), and, in this impoverished discourse-model, the only choices left are *Sam* and *John*, in that order. Thus the marked foci effect a change in coreference. Second, the marked foci cause *insulted* to be identified as given and thus equivalent to *called a Republican*, as in (6a). Therefore, we no longer need say that *called a Republican* in such sentences "presupposes" *insulted* in order to arrive at an account of the understanding this sentence receives. The situation is not quite so simple, of course; consider (8).

- (8) *John called Mary a Republican /and then SAM walked in /and they all started fighting.*

(Slashes indicate information-unit boundaries.) If *SAM* is the marked focus of its information-unit, then *walked in* must be given and should be understood as being equivalent to something in the preceding informa-

⁴ Halliday, himself, seems not to take the intonation criterion seriously when he remarks that anaphoric items are "inherently 'given' (1967, p. 206)." But perhaps the quotation marks around *given* implicate "not really."

tion-unit, which of course is not the case. Obviously, then, we do not want to say that *walked in* is given, but an analysis that considers only intonation as a primitive cannot distinguish between *walked in* in (8) and *insulted* in (6a) (unless, of course, the stress on *SAM* in (8) is considered to be of an identifiably different type from the stress on *SHE* and *HIM* in (6a)). What is needed is the inclusion of Kuno's predictability as a primitive, along with intonation, perhaps as part of a principle along the lines of (9):

- (9) **Parallelism Principle:** A speaker assumes that the hearer will predict, unless there is evidence to the contrary, that (a proper part of) a new (conjoined?) construction will be parallel/equivalent in some semantic/pragmatic way(s) to the one just processed.⁵

This is obviously a rough and inadequate approximation, but the idea is that, while extra stress always marks its item as new (in the sense of being unpredictable in a particular sentence-position), there are some situations in which the Parallelism Principle can still be invoked (e.g., (6a), (7a)) and others where other factors require its abandonment (e.g., (8)). Perhaps it is only in the former cases that the complement of a marked focus should be labeled as given. (Note that some version of the Parallelism Principle is needed in any event in order to account for the usual understanding of the pronouns in (7b).) Much research, including psycholinguistic experimentation, is required before any contentful version of (9) can be arrived at; what I have tried to do here is simply to isolate the notion of givenness in the sense of predictability/recoverability and to suggest that a consideration of speakers' hypotheses about hearers' beliefs and strategies must be a primitive.

2.2 Givenness_s: Saliency

In speaker-hearer terms, givenness in the sense of saliency may be roughly described as follows:

- (10) **Givenness_s:** The speaker assumes that the hearer has or could appropriately have some particular thing/entity/ . . . in his/her CONSCIOUSNESS at the time of hearing the utterance.

Chafe's (1976) notion of given-new information falls under this rubric; for Chafe, given information represents "that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance," and new, "what the speaker assumes he is introducing into the

⁵ See Kuno (1974) for a discussion of the notion of parallelism.

addressee's consciousness by what he says [p. 30]." He takes it to be a binary distinction; thus known items that are introduced into the discourse for the first time are as new as unknown ones; for example, the bold-faced NPs in (11a) and (11b) are equally new:

- (11) a. *I saw **your father** yesterday.*
b. *I saw **a two-headed man** yesterday.*

Furthermore, for an NP to qualify as given, its referent must have been explicitly introduced in the discourse or be present in the physical context or be categorized in the same way as a referent previously introduced or physically present (p. 32). Thus an inferentially related NP cannot be given, unless the inference is one of categorization. In the following examples (originally from Haviland and Clark, 1974), Chafe (pp. 41-42) calls *the beer* given in (12a) and new in (12b):

- (12) a. *We got some beer out of the trunk. **The beer** was warm.*
b. *We got some picnic supplies out of the trunk. **The beer** was warm.*

At the same time, however, Chafe states (p. 31) that only (but not all and only) given items can be pronominalized. Examples like (13) then suggest that inferentially related NPs in general might have to be included as possible given items:

- (13) a. *Pick two numbers, add six to the first number, and then multiply it by the second number.⁶*
b. *Harry threw up and Sam stepped in it.⁷*
c. *In New York **they** drive very differently from the way **they** drive in Philadelphia.*

Chafe's notion of givenness is the model for the notion used in Prince 1978a for the analysis of the function of *wh*-clefts, except that it was extended to allow for any inferences so long as they could be deemed situationally appropriate.⁸ Thus the bold-faced NP in (14) is given following Prince 1978a, but presumably new following Chafe 1976:

- (14) *If I write loosely of a noun as being in the status **GIVEN**, **what I really mean** is that the idea which this noun expresses has this status. (Chafe, 1976; p. 29.)*

⁶ I thank Bonnie Webber for this example.

⁷ See P.R.N. Tic Douloureux (1971).

⁸ New_s information is not discussed in Prince 1978a and the question of whether the given-new_s distinction is binary is not addressed.

That is, following Prince 1978a, the writer of (14) is purporting to assume that it is appropriate for the reader to have in mind, by the time s/he reaches the comma, that the writer really means something if he writes loosely of a noun as being in the status *given*.

There is of course a difference between the givenness of pronouns like the one in (15) and the givenness of *wh*-clauses like the one in (14):

- (15) [A₁ to B₁ as C_k passes by, in view and out of earshot]
How old do you think he_k is?

The difference is that, if the utterance is to be felicitous, the sender's assumptions about what is in fact in the receiver's consciousness must be correct in (15) but need not be in (14). That is, some markers of givenness_s (e.g., pronouns) correlate with a loss of explicit information such that understanding is greatly impeded if the item is not actually given_s for the receiver, whereas other markers (e.g., subject clauses) do not correlate with any loss of information and permit senders greater leeway in treating items as given_s. Flagrant misuse will indeed make the utterance sound inappropriate or bizarre, as in (16), but not incomprehensible:

- (16) a. [Customer opening service-encounter]
?What my friend bought here was a beautiful scarf.
b. [Newscaster opening evening news]
?That Carter will resign tonight has just come over the wire.

It seems that those items that must in fact be given_s for an utterance containing them to be felicitous are marked morphologically (pronouns, including deictics) [but cf. (13)], whereas those items that require only that their givenness_s be appropriate are marked syntactically (e.g., subject clauses). (See Horn, 1978 for a discussion of the givenness_s of subject clauses.) As in the case of givenness_p, much research is required, including psycholinguistic experimentation, to discover how language users actually go about marking and recognizing linguistic forms as representing given_s and new_s items.

2.3 Givenness_k: "Shared Knowledge"

In speaker-hearer terms, givenness in the sense of "shared knowledge" may be described as follows:

- (17) Givenness_k: The speaker assumes that the hearer "knows," assumes, or can infer a particular thing (but is not necessarily thinking about it).

This type is represented by what Clark and Haviland (1977) call given-

new information: given is "information [the speaker] believes the listener already knows and accepts as true" and new is "information [the speaker] believes the listener does not yet know (p. 4)." Whether the hearer knows the information directly for having been explicitly told it, or indirectly via inferencing ("bridging") is immaterial; thus *the beer* is given_k in both (12a) and (12b) for Clark and Haviland, as would be the bold-faced items in (18) (when uttered felicitously):

- (18) a. Have you heard from *Jane-Carol* recently?
b. We got some *beer* out of the trunk and it was warm.
c. Where were your *grandparents* born?

Kuno's (1972) notion of anaphoric-nonanaphoric, mentioned earlier, likewise falls under the rubric of givenness_k: An NP is anaphoric if "[its] referent . . . has been mentioned in the previous discourse" or is "in the permanent registry (p. 270)," where "the permanent registry" corresponds to what the speaker assumes about the hearer's assumptions. Presumably, Kuno's From-Old-to-New Principle (1979), which is posited in order to account for certain word-order phenomena in flexible word-order languages and for some syntactic constructions in English (Dative, Passive), is sensitive to givenness_k. That is, when we say that the fact that (19a) is better than (19b) is related to a tendency to put old information before new information, we are using *old* in the sense of "shared knowledge" (givenness_k) and not in the sense of predictability (givenness_p) or saliency (givenness_s), as *John* and *a boy* are equally unpredictable and unsalient and differ only in that *John* is usually given_k ("in the permanent registry"), whereas *a boy* is new_k:

- (19) a. *John* hit a boy on the head.
b. ?A boy was hit on the head by *John*.⁹

2.4 Relatedness of the Three Types of Givenness

Although different from one another, predictability, saliency, and "shared knowledge" are not mutually independent: If a speaker assumes that the hearer can predict that some particular item or items will occur in some particular position within a sentence, then the speaker must assume that it is appropriate that the hearer have some particular thing in his/her consciousness. And, if the speaker assumes that the hearer has some particular thing in his/her consciousness, then the speaker must assume that the hearer has some assumption or can draw some inference. Further-

⁹ These examples are from Kuno (1979, p. 6).

more, all three levels sometimes involve cases where some item is "given" for extralinguistic reasons, showing that all levels must ultimately relate to extralinguistic phenomena, in particular to what the speaker thinks is or should be or could appropriately be in the hearer's mind. For example, consider (20)–(22) in their most usual understandings:

- (20) a. \emptyset *Wanna fight?* (Given_p)¹⁰
 b. [Man to woman at ball, with appropriate body language]
Shall we \emptyset ? (Given_p)
- (21) a. [*A_i to B_j, as C_k walks by, in view and out of earshot*]
He_k's going to Austria. (Given_s)
 b. [Professor beginning first lecture of term]
What we're going to look at this semester is the world's indifference to the Boat People. (Given_s)
- (22) a. *Hi, I'm home.* (Given_k)
 b. *Where's Daddy?* (Given_k)

Thus, an understanding of givenness in the sense of "shared knowledge," though it reeks of things "nonlinguistic," is germane (and, perhaps, prerequisite) to an understanding of givenness in the other two senses, and I shall now focus exclusively on "shared knowledge," or givenness_k.

3. SO-CALLED "SHARED KNOWLEDGE"

3.1 Terminology

As a first step, let us discard the term SHARED KNOWLEDGE once and for all, for it has given rise to great confusion. If one thinks about it for a moment, one sees that all a speaker has to go on when treating something as given_k or "shared" is what s/he assumes the hearer assumes. The view that says that each individual has a belief-set and that, for any two individuals, the belief-sets may be overlapping, the intersection constituting "shared knowledge", is taking the position of an omniscient observer and is not considering what ordinary, nonclairvoyant humans do when they interact verbally.

At the same time, GIVENNESS, with or without a distinguishing subscript, is not a very convenient term either, as evidenced by its confusing

¹⁰ Although, in the rest of this chapter, the notion of givenness is reserved for entities (representable by NPs), what seems to be predictable and hence given_p in (20) are not entities (*Do you, dance*).

use in the literature. For lack of something better, I propose the term ASSUMED FAMILIARITY. At least, it lacks the unhelpful connotations of symmetry and fact and does not sound like anything else (see, however, Christopherson, 1939 and Hawkins, 1978.)

3.2 The Problem

We may now word the basic problem as follows. From the point of view of a speaker/writer, what kinds of assumptions about the hearer/reader have a bearing on the form of the text being produced, where that form is not uniquely determined by the "objective" information that the speaker/writer is attempting to convey? From the point of view of the hearer/reader, what inferences will s/he draw on the basis of the particular form chosen? We are, therefore, NOT concerned with what one individual may know or hypothesize about another individual's belief-state EXCEPT insofar as that knowledge and those hypotheses affect the forms and understanding of LINGUISTIC productions. Thus, although the problem is of general relevance to cognitive psychology in that it touches on matters of reasoning and knowledge, it is first of all a linguistic problem, if a goal of linguistics is to produce a theory of discourse that distinguishes between a random sequence of sentences and something we would intuitively call a "text."

The solution to the problem then may be seen as requiring three parts: (a) a taxonomy of linguistic forms, both morphological and syntactic; (b) a taxonomy of the values of Assumed Familiarity; and (c) an account of the correlation between the two. Structural linguistics and transformational grammar has provided us with the first part, at least for forms that are identifiable on the level of the sentence or less, and this chapter is an attempt to provide the second. Hopefully, once the two taxonomies have been arrived at, further research can determine the correlation.

3.3 The Taxonomy

One question that arises, then, is what the values of Assumed Familiarity are, that is, is it a binary distinction, a continuum, or something else? Consider the sentences in (22):

- (22) a. *Pardon, would you have change of a quarter?*
 b. *Noam Chomsky went to Penn.*
 c. *I got on a bus yesterday and the driver was drunk.*
 d. *A guy I work with says he knows your sister.*
 e. *Hey, one of these eggs is broken!*

It seems impossible to group the bold-faced NPs in (22) into two homogeneous and discrete sets: *you*, *Noam Chomsky*, *the driver*, *he*, and perhaps *one of these eggs* are all somehow felt to be familiar, that is, familiar to the hearer, yet they are all different. Likewise, *a bus*, *a guy I work with*, and perhaps *one of these eggs* are somehow felt to be new, or unfamiliar to the hearer, yet they are intuitively not the same.

To get a better idea of what I think is going on, I should like to propose a not very original analogy. Consider a text to be like a recipe. The goal of the recipe writer is to instruct the recipe reader on how to produce a particular dish. Two recipes for the same dish may be dramatically different; compare, for example, *Le Répertoire de la cuisine*, the very slim cookbook that professional French chefs use, where each recipe is just a few lines long, with a standard fat American cookbook, where major ingredients are listed, much of the equipment specified, and many of the processes spelled out.¹¹ Even then, however, some things are left unstated; even *The I Hate to Cook Book* does not tell us where to find boiling water. Why such differences and apparent inconsistencies? The obvious answer is that the writer of a recipe has a certain set of assumptions about what

¹¹ Following are two recipes for roast suckling pig, one from Rombauer, I. S., and Becker, M. R. 1931, *The Joy of Cooking* (p. 408), and the other from Gringoire, Th. and L. Saulnier, 1914, *Le Répertoire de la cuisine* (p. 81). They are cited here in full as a dramatic illustration of the effect that the sender's hypotheses about the receiver's beliefs has on the form of the text. First, *The Joy of Cooking*:

Roast Suckling Pig:

10 servings

Preheat oven to 450°.

Dress, by drawing, scraping and cleaning:

A suckling pig

Remove eyeballs and lower the lids. The dressed pig should weigh about 12 pounds. Fill it with:

Onion Dressing, page 457, or

Forcemeat, page 458

It takes 2½ quarts of dressing to stuff a pig of this size. Multiply all your ingredients, but not the seasonings. Use these sparingly until the dressing is combined, then taste it and add what is lacking. Sew up the pig. Put a block of wood in its mouth to hold it open. Skewer the legs into position, pulling the forelegs forward and bending the hindlegs into a crouching stance. Rub the pig with:

Oil or soft butter

(A cut clove of garlic)

Dredge it with:

Flour

Cover the ears and the tail with aluminum foil. Place the pig in a pan uncovered in the oven for 15 minutes. Reduce the heat to 325° and roast until tender, allowing 30 minutes to the pound. Baste every 15 minutes with:

the reader knows about ingredients, processes, and equipment, about what equipment the reader has available, and about what staples the reader keeps on the shelf.

Turning back to discourse, let us say that a TEXT is a set of instructions from a speaker to a hearer on how to construct a particular DISCOURSE-MODEL. The model will contain DISCOURSE ENTITIES, ATTRIBUTES, and LINKS between entities. A discourse entity is a discourse-model object, akin to Karttunen's (1971) DISCOURSE REFERENT; it may represent an individual (existent in the real world or not), a class of individuals, an exemplar, a substance, a concept, etc. Following Webber (1978), entities may be thought of as hooks on which to hang attributes. All discourse entities in a discourse-model are represented by NPs in a text, though not all NPs in a text represent discourse entities.

When a speaker first introduces an entity into the discourse, that is, tells the hearer to "put it on the counter," we may say that it is NEW. New discourse entities are of two types, however. In one case, the hearer may have had to CREATE a new entity, akin to going out and buying a suckling pig, in which it is BRAND-NEW. In the other case, the hearer may be assumed to have a corresponding entity in his/her own model and simply has to place it in (or copy it into) the discourse-model, akin to taking some staple off the shelf when its presence is suddenly taken for granted in a recipe (e.g., salt). Call this type UNUSED. In the sentences of (22), assum-

About 2 cups boiling stock and the pan drippings

Remove the foil from the ears and tail before serving. Place the pig on a platter.

Remove the wood from the mouth. Replace it with a small:

Apple, lemon or carrot

Place in the eyes:

Raisins or cranberries

Drape around the neck a wreath of:

Small green leaves

or garnish the platter or board with:

Cinnamon Apples, page 111, Apples Stuffed with Sweet Potatoes, page 111,

Apples Stuffed with Mincemeat, page 603, Tomatoes Florentine, page 307, etc.

Make:

Pan Gravy, page 322

To carve, place head to left of carver. Remove forelegs and hams. Divide meat down center of back. Separate the ribs. Serve a section of crackling skin to each person.

Now, compare with the recipe as presented in *Le Répertoire de la cuisine*:

Cochon de Lait Anglaise:

—*Farcir farce à l'anglaise. Rôtir.* (Translation: English Suckling Pig: Stuff with English stuffing. Roast.)

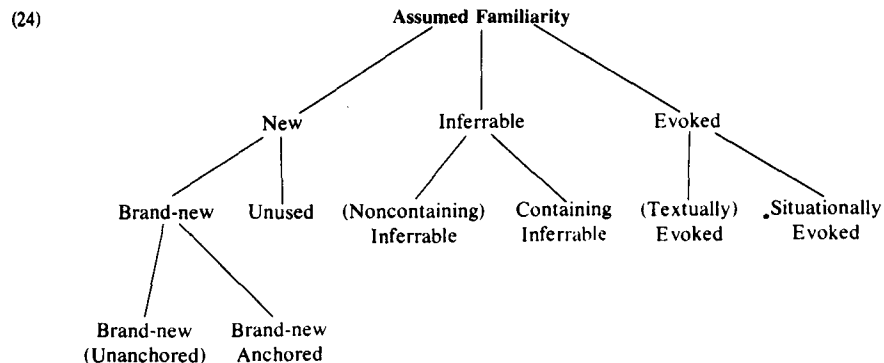
ing each is discourse-initial, *Noam Chomsky* is Unused, that is, assumed to be in the hearer's model, whereas *a bus* and *a guy I work with* are Brand-new and must be created by the hearer. Brand-new entities themselves seem to be of two types: ANCHORED and UNANCHORED. A discourse entity is Anchored if the NP representing it is LINKED, by means of another NP, or "Anchor," properly contained in it, to some other discourse entity. Thus *a bus* in (22c) is Unanchored, or simply Brand-new, whereas *a guy I work with* in (22d), containing the NP *I*, is Brand-new Anchored, as the discourse entity the hearer creates for this particular guy will be immediately linked to his/her discourse entity for the speaker. In the data, all Anchored entities contain at least one Anchor that is not itself Brand-new; that is, we find NPs like (23a) but not like (23b).

- (23) a. *a guy* $\left\{ \begin{array}{l} I \\ John \\ the plumber \\ a woman I know \end{array} \right\}$ *work(s) with*
 b. ?*a guy* $\left\{ \begin{array}{l} a woman works with \end{array} \right\}$

We shall return to this in what follows.

Now, if some NP is uttered whose entity is already in the discourse-model, or "on the counter," it represents an EVOKED entity. There are, grossly, two ways in which an entity can have come to be Evoked: Either the hearer had evoked it earlier, on textual grounds, by following instructions from the speaker—that is, it was once New or Inferrable (to be discussed later)—or the hearer knew to evoke it all by himself, for situational reasons. Call the first type TEXTUALLY EVOKED, or simply EVOKED, and the second SITUATIONALLY EVOKED. Situationally Evoked entities represent discourse participants and salient features of the extratextual context, which includes the text itself. In (22d), *he* is Textually Evoked, whereas, in (22a), *you* is Situationally Evoked.

The third and most complex type of discourse entity are the INFERRABLES. A discourse entity is Inferrable if the speaker assumes the hearer can infer it, via logical—or, more commonly, plausible—reasoning, from discourse entities already Evoked or from other Inferrables. In (22c), *the driver* is Inferrable from *a bus*, plus assumed knowledge about buses, that is, *Buses have drivers*. A special subclass of Inferrables are the CONTAINING INFERRABLES, where what is inferred off of is properly contained within the Inferrable NP itself; in 22e, *one of these eggs* is a Containing Inferrable, as it is inferrable, by a set-member inference, from *these eggs*, which is contained within the NP and which, in the usual case, is Situationally Evoked. Perhaps the diagram in (24) will make these distinctions somewhat clearer:



In addition, all seven types may occur with one or more new attributes, as the discourse-initial sentences in (25), (26), and (27) illustrate:

- (25) a. *I bought a beautiful dress.* (Brand-new + attribute)
 b. *A rich guy I know bought a Cadillac.* (Brand-new Anchored + attribute)
 c. *Rotten Rizzo can't have a third term.* (Unused + attribute)
- (26) a. *I went to the post office and the stupid clerk couldn't find a stamp.* (Inferrable + attribute)
 b. *Have you heard the incredible claim that the devil speaks English backwards?* (Containing Inferrable + attribute)
- (27) a. *Susie went to visit her grandmother and the sweet lady was making Peking Duck.* (Evoked + attribute)
 b. *Lucky me just stepped in something.* (Situationally Evoked + attribute)

4. ILLUSTRATION

Now let us consider and compare two naturally occurring texts, representing two different styles, to see how the taxonomy I am proposing works and what, if any, patterns appear. The first text we will look is an informal oral narrative; the speaker is a white, female, middle-aged, middle-class suburban Philadelphia college graduate. She has recently broken some bone in a fall and is wearing a cast. She is at home. The hearer is a friend, of the same sex and background, but with an advanced degree in linguistics and a tape recorder in her handbag. The speaker does not know that she is being taped. First, the text:

- (28) a. *Well, I have a friend of mine called me;*
 b. *a friend of hers who I know—two weeks—well,*
 c. *last week she called and said, "Well, you have*
 d. *company. Jan fell down four flight of steps." They*
 e. *have a house like this, and she was going to a*
 f. *luncheon and the women were honking the horn outside;*
 g. *she heard them, right? And usually she lets the door*
 h. *open but she didn't this time. So she comes*
 i. *running down the steps and she fell down four and*
 j. *landed on her side. Her right side's fractured. She's*
 k. *over at Holy Redeemer. She tried to get up and couldn't*
 l. *so she—and she realized—and they were still honking,*
 m. *see? So then after a while, she thought one of them'll*
 n. *have enough sense to come to the door, ring the bell,*
 o. *right? So she couldn't get up she said, she—it was a—*
 p. *so she crawled to the door and then finally one of the*
 q. *women came and rang the bell and she said to them, "I'm*
 r. *on the floor, I fell and I can't get up." So Nancy said,*
 s. *"Well, open the door," and she said, "I can't reach*
 t. *the knob," she said. "I'm really hurting." She said, "I*
 u. *can't pick myself up." She said, "My God, what did you*
 v. *do?" She said, "I don't know, I fell down the steps."*
 w. *So she said, "Well, can't you roll over on your other side?"*
 x. *and she said, "Honest to goodness, I can't do anything,"*
 y. *she said. "Can't you come in?" She said, "Well, how am I*
 z. *going to get in?" She said, "Well, try the kitchen window,*
 aa. *that—that's open." She said, "Tell Jane," who's a tiny*
 bb. *little thing (she's only about four feet ten), she said,*
 cc. *"Tell—can you hike Jane up and get her—go in the*
 dd. *garage and you'll find a stool or something for her to*
 ee. *get on and then hike her through the kitchen window."*
 ff. *So that's what they did. So she goes through and she says*
 gg. *she landed in the sink. Well, naturally, it's like*
 hh. *our kitchen. So she had taken her shoes off, right?*
 ii. *She had heels on and she took them off when they hoisted*
 jj. *her. She was on a step ladder but then they still*
 kk. *had to give her a little push, right? So she got in, she*
 ll. *said she sat right in the sink. So she had to work*
 mm. *her way out of that and she got in and here she opened*
 nn. *up the front door and it took the four of them to get*
 oo. *her up and she was screaming, she was in such pain. (Wolf-*
 son, 1976, pp. 160–61)

The notation is explained in (29), and the Assumed Familiarity analysis is given in (30):

- (29) **Notation**
 BN: Brand-new
 BN_A [—]: Brand-new Anchored [*Anchor*^{type}]
 U: Unused
 I(—)/—: Inferrable (*entity inferrable from*^{type})/inference-type
 I_C(—)/—: Containing Inferrable (*contained entity inferrable from*^{type})/inference-type
 E: (Textually) Evoked
 E_S: Situationally Evoked
 S.A.: Stereotypic Assumption
- (30) a. *I_i*: E_S (Discourse-participant = Speaker)
a friend of mine_i: BN^A[*i*^E]
me_i: E
 b. *a friend of hers_j who I_i know_k*: BN_A[*j*^E, *i*^E]
 c. *she_j*: E
you_i: E
 d. *company_k*: ?BN (if in fact entity)
Jan_k: E + A(ttribute)
four flight of steps: BN + A
they_{k+}: I(*k*^E)/member-set
 e. *a house like this*: BN_A[*this house*^{E*}]
she_k: E
 f. *a luncheon*: BN
the women_i: I(*luncheon*^E)/S.A. or frame
the horn: I(*car*^{luncheon}^E/frame)/S.A. (Luncheon frame:
 Cars have horns)
 g. *she_k*: E
them₁: E
she_k: E
the door: I(*house*^E)/S.A. (Houses have doors)
 h. *she_k*: E
she_k: E
 i. *the steps*: either I(*house*^E)/S.A. (Houses have steps)
 or E (from capsule statement at beginning)
she_k: E
four: I(*the steps*^E)/set-subset

- j. *her_k side*: I(*k*^E)/S.A. (People have sides)
her_k right side: E + A
she_k: E
- k. *Holy Redeemer*: U
she_k: E
- l. *she_k*: E
they_l: E
- m. *she_k*: E
one of them_{l_m}: I(*l*^E)/set-member
- n. *the door*: E
the bell: I(*the door*^E)/S.A. (Doors have bells)
- o. *she_k*: E
she_k: E
- p. *she_k*: E
the door: E
one of the women_{l_m}: E + A
- q. *the bell*: E
she_k: E
them_l: E
I_k: E
- r. *the floor*: I(*the house*^E)/S.A. (Houses have floors)
I_k: E
I_k: E
Nancy_m: E + A
- s. *the door*: E
she_k: E
I_k: E
- t. *the knob*: I(*the door*^E)/S.A. (Doors have knobs)
she_k: E
I_k: E
she_k: E
I_k: E
- u. *myself_k*: E
she_m: E
you_k: E
- v. *she_k*: E
I_k: E
I_k: E
the steps: E

- w. *she_m*: E
you_k: E
your_k other side: I(*k's side*^E)/S.A. (People have two sides)
- x. *she_k*: E
I_k: E
- y. *she_k*: E
you_m: E
she_m: E
I_m: E
- z. *she_k*: E
the kitchen window: I_C(*kitchen*^l(*house*^E)/S.A.)/S.A. (Houses have kitchens; Kitchens have windows)
- aa. *that*: E
she_k: E
Jane_n: I(*l*^E)/set-member
who_n: E
- bb. *she_n*: E
she_k: E
- cc. *you_m*: E
Jane_n: E
her_n: E
- dd. *the garage*: I(*the house*^E)/S.A. (Houses have garages)
you_m: E
a stool or something: BN
her_n: E
- ee. *her_n*: E
the kitchen window: E
- ff. *that*: I(preceding propositions)/propositions = act
they_l: E
she_n: E
she_n: E
- gg. *she_n*: E
the sink: I(*the kitchen*^E)/S.A. (Kitchens have sinks)
it: E
- hh. *our kitchen*: U
she_n: E
her shoes: I(*n*^E)/S.A. (People have shoes on)
- ii. *she_n*: E
heels: E + A

- she_n*: E
them: E
they_i: E (Actually, *they* = *l - n*)
 jj. *her_n*: E
 she_n: E
 a little step ladder: either BN + A or E + A (from *a stool or something*)
 they_i: E (actually *l - n*)
 kk. *her_n*: E
 she_n: E
 she_n: E
 ll. *she_n*: E
 the sink: either E (from capsule statement) or I(*the kitchen^E*)/S.A.
 she_n: E
 mm. *that*: I(preceding propositions)/propositions = situation
 she_n: E
 she_n: E
 nn. *the front door*: E + A
 the four of them_i: E + A
 oo. *her_k*: E
 she_k: E
 she_k: E

(Adverbials, expletives, dummies, and idiom-pieces have been omitted.)

Obviously, space does not permit a step-by-step discussion. A few points, however, are worth mentioning. First, nearly all of the subjects are Evoked, whereas less than half of the nonsubjects are. In contrast, one-sixth of the nonsubjects—but none of the subjects—are New. As for the Inferrables, they account for nearly one-third of the nonsubjects but only one-fifteenth of the subjects. The figures are presented in Table 1.

The types of inferences needed for the Inferrables are mostly culture-based, involving the Stereotypic Assumptions discussed in Prince 1978b in relation to existential presupposition (e.g., *Houses have doors*, *People have sides*, *Kitchens have sinks*). One set—*luncheon to women to horn* in (30f)—suggests that the much more complex notion of “frame” may in fact be necessary.

There are also some logical inferences: set to subset, as in *the steps to four* in (30i), and set to member, as in *they to one of them* in (30l)–(30m).

TABLE 1

Analysis of Subjects and Nonsubjects in Oral Text
(by Number and Percentage)

| | Subjects (76) | Nonsubjects (43) |
|-----------------|---------------|------------------|
| Evoked | | |
| E | 66 86.8% | 18 41.9% |
| E _s | 1 1.3% | 0 0.0% |
| E + A | 4 5.3% | 3 7.0% |
| Total | 71 93.4% | 21 48.8% |
| Inferrable | | |
| I | 5 6.6% | 12 27.9% |
| I _c | 0 0.0% | 1 2.3% |
| Total | 5 6.6% | 13 30.2% |
| New | | |
| U | 0 0.0% | 2 4.7% |
| BN _A | 0 0.0% | 3 7.0% |
| BN | 0 0.0% | 3 7.0% |
| BN + A | 0 0.0% | 1 2.3% |
| Total | 0 0.0% | 9 20.9% |

One inference is logical—member to set—but also culture-based: *Jan* to *they* in (30d), where the speaker must assume that the hearer assumes that people like *Jan* are members of some relevant set, presumably a family.

One other point bears mentioning. In Prince 1978b, it was claimed that, in colloquial discourse, unmodified and unexplained proper nouns are not used unless the speaker assumes that the hearer is already familiar with the referent. At first blush, this text seems to falsify that claim, and, in fact, the hearer assured me that she had never heard of *Jan*, introduced in (30d), or *Nancy*, introduced in (30r), and could not have been expected to by the speaker. One might discount the case of *Jan* as being part of direct discourse, but no such loophole exists for *Nancy*. Worse yet, other such occurrences appear in other data, none of which can be dismissed as performance errors. However, it turns out that, when whole texts are analyzed, it becomes clear that these putatively deviant proper nouns do NOT in fact represent what I am here calling Unused entities (which do obey the “constraint”) but represent rather Evoked entities, with the added attribute that the entity has a certain name. For example, by (30p), the hearer has already constructed an entity for the woman who had enough sense to get out of the car and come to the door. All that the NP *Nancy* in (30r) requires, then, is that the hearer assign the name *Nancy* to that en-

tity.¹² In contrast, *Holy Redeemer* in (30k), which cannot be analyzed as Evoked + Attribute, was in fact independently known to the hearer and exemplifies the more obvious use of proper nouns as Unused information that was at issue in Prince 1978b.

An interesting problem that requires more research for its resolution is the status of discourse entities introduced by what is referred to above as a "capsule statement," or "abstract" (Labov and Fanshel, 1977). That is, narratives frequently begin with a statement of the point or justification of the rest of the narrative, in which certain entities are represented. The question for us is whether those entities that are represented in such capsule statements are considered to have been evoked or not. In (30), there are two possible capsule statements:

- (31) a. *You have company. Jan fell down four flight of steps.*
 b. *So she goes through and she says she landed in the sink.*

Of the eight NPs in (31), five (*you, Jan, she, she, she*) are already evoked and do not present a problem. One, *company*, does not even clearly represent a discourse entity directly. (It seems to "prepare" the hearer to meet with a new entity; compare with *You're not alone*.) As for the remaining two NPs, *four flight of steps*¹³ and *the sink*, the next coreferential NPs (*the steps, the sink*, respectively) are ambiguous as to whether they represent Evoked or Inferrable entities. My own intuition inclines toward treating them as Inferrable, but the decision is not clear. If these next-mentions are in fact Inferrable and not Evoked, then we conclude that, although entities may somehow be "introduced" in capsule statements, they are not, however, actually evoked there. Analogous questions would no doubt arise in the treatment of entities represented by NPs in the titles of written texts, but of course it is quite possible that the two situations will be shown to work differently.¹⁴

¹² Note also that the entity represented by *Jan* in (28d) is originally introduced by a repaired string (*a friend of hers who I know*), which might have been edited out in a less faithful transcription. This raises important methodological and theoretical questions about the treatment of repairs in linguistic analyses of naturally occurring data.

¹³ *Four flight of steps* seems to be an irrelevant and unrepaired performance error; it should be *four steps*.

¹⁴ Gerald Prince has pointed out (personal communication) the problematic situation where some entity in a title (e.g., *A Fight*) may serve as a label for—or be the entity representing—the whole text that follows but does not correspond to any individual entity within that text (e.g., a description of two individuals debating). Titles then may serve not so much for the addition of entities and attributes to a discourse-model as for framing the discourse-model that is about to be constructed. Experiments reported in Bransford and Johnson (1972, 1973), where the presence of a title grossly affected the subjects' processing of a text, support this view of a title (and perhaps a capsule statement) as a metatextual object.

Analyses of this and other naturally occurring texts with respect to Assumed Familiarity reveal a number of patterns or tendencies. First, we find a preferred hierarchy or scale for what type of entity is used, more or less as in (32):

- (32) **Familiarity Scale**
 $\left\{ \begin{array}{l} \mathbf{E} \\ \mathbf{E}^s \end{array} \right\} > \mathbf{U} > \mathbf{I} > \mathbf{I}^c > \mathbf{BN}^A > \mathbf{BN}$

For example, consider the bold-faced NPs in (33) (in their most usual understandings):

- (33) a. *I bought a Toyota.* (E^s)
 b. *Ellen bought a Toyota.* (U)
 c. *One of the people that work at Penn bought a Toyota.* (I^c)
 d. *A person that works at Penn bought a Toyota.* (BN^A)
 e. *A person bought a Toyota.* (BN)

It seems that, if a speaker is in a position to say one of these on basis of his/her hypothesis about what the hearer knows and chooses instead to say one lower on the scale (to refer to the same individual), s/he will be seen, if found out, to have been deviant in some way (e.g., evasive, childish, building suspense as in a mystery novel). Put differently, we may say that the use of an NP representing a certain point on the scale implicates that the speaker could not have felicitously referred to the same entity by another NP higher on the scale. The recognition of such a scale permits this sort of implicature to be subsumed under the Gricean maxim of Quantity.

Note that the use of the scale must be relative to the speaker's hypothesis about the hearer's belief-set and cannot be construed as a statement about the relative probability of a particular type of NP occurring. For example, U is higher on the scale than I in spite of the fact that most texts contain more Inferrable entities than Unused ones. The reason is that most entities discussed are simply not previously assumed to be known to the hearer; for example, if the speaker of (22c) happened to know that the bus driver's name was Jake but assumed that the hearer did not know this, s/he could not substitute *Jake for the driver* and still be deemed coherent. If, on the other hand, the speaker thought that the hearer did already have an entity for this individual with the attribute of having the name *Jake* and simply said *the driver*, the hearer, by following the Familiarity Scale, would infer not-Jake and would, if s/he ever found out, feel that the speaker had "withheld information." The Familiarity Scale may fall under a more general Conservation Principle that says that hearers do not like to make new entities when old ones will do and that speakers, if

they are cooperative, form their utterances so as to enable the hearer to make maximal use of old entities.¹⁵

From the analysis of informal conversational texts like the one presented in (28), it seems that the same Familiarity Scale holds for NPs representing anchors, within BN^A entities. That is, the NPs in (34) represent the same scalar inequalities as those in (33), except that (34e) is virtually unacceptable:

- (34) a. *A friend of yours bought a Toyota.* (E^S)
 b. *A friend of Steve's bought a Toyota.* (U)
 c. *A friend of my neighbor's bought a Toyota.* (I)
 d. *A friend of a guy I know bought a Toyota.* (BN^A)
 e. ?*A friend of a guy's bought a Toyota.* (BN)

In fact, whereas the scale seems relevant in all types of discourse, the use of Brand-new Anchored entities, which is a means of UPGRADING an entity on the scale, seems to be characteristic of informal conversational discourse as in (28).

More special to informal conversational discourse, however, is the tendency to reserve subject position for NPs at the higher end of the scale. In fact, we find a variety of syntactic constructions that appear to be used at times solely in order to enable the speaker to do this, that is, to keep entities low on the scale out of subject position: existential *there*, *it*-clefts, Left Dislocation out of subject position, relative clauses, and a variety of so-called "run-on" sentences, for example

- (35) a. "*I had a little boy, black, about ten years old, HE used to go with me.*" —Terkel (1974; p. 132)
 b. "*There's some male beauty shops, THEY deal more in your feminine men and actors.*" —Terkel (1974; p. 317)
 c. "*There are some funerals, THEY really affect you.*" (Terkel, 1974, p. 661)¹⁶

¹⁵ Compare Sacerdoti's (1977) principle, "Use existing objects." Stenning (1978) has noted the same phenomenon and points out that traditional mystery novels violate this principle by causing the reader to create an entity for the murderer and for each other character, telling him only at the end to collapse the entity for the murderer with the entity for a particular other character.

¹⁶ Although such constructions usually result in the subject representing an Evoked entity, as in the examples of (35), they sometimes result in the subject being lower than Evoked on the Familiarity Scale. It is always the case, however, that the new subject is higher on the scale than would be the subject of the corresponding simple sentence. In (i), for example, the subject is an Inferrable, in contrast to the corresponding Containing Inferrable of (ii):

- (i) "*I had a pair of shoes, the soles were loose.*" (Terkel, 1974, p. 576.)
 (ii) *The soles of a pair of shoes I had were loose.*

Furthermore, deletion of subject relative markers appears to occur in just such cases, for example, in the first line of the narrative in (28) and in the sentences of (36):

- (36) a. "*We got a lot of fancy Cadillac cars don't tip.*" (Terkel, 1974, p. 298).
 b. "*I had a great-great-great-grandfather or something fought that Revolution.*" (Terkel, 1974, p. 42.)
 c. "*There was a piece of four-inch bone never mended.*" (Terkel, 1974, p. 52.)

On this point, Anthony Kroch (personal communication) has noted that, in his large corpus of oral discourse, ALL instances of subject relative marker deletion occur in sentences like those of (36), that is, where the syntactically main clause is informationally weak, having either an Evoked NP subject (usually *I*) or a dummy subject (*it*, *there*), and a semantically weak verb (e.g., *be*, *have*, *know*), and where the syntactically subordinate (i.e., relative) clause is highly informative. Thus it seems that the more formal counterparts to (36) are not (37), as one might think, but (38):

- (37) a. *We have a lot of fancy Cadillac cars which/that don't tip.*
 b. *I had a great-great-great-grandfather or something who/that fought that Revolution.*
 c. *There was a piece of four-inch bone which/that never mended.*
 (38) a. *A lot of fancy Cadillac cars don't tip.*
 b. *A great-great-great-grandfather or something fought that Revolution.*
 c. *A piece of four-inch bone never mended.*

In summation, it seems that we can distinguish a phenomenon that is perhaps true of discourse in general, the tendency to use an NP that is as high on the Familiarity Scale as felicitously possible, and one related phenomenon that seems especially pertinent to informal conversational discourse, a conspiracy of syntactic constructions resulting in the nonoccurrence of NPs low on the scale in subject position.

Now let us consider another text, the beginning of a chapter in Hymes 1974. Note that, unlike the oral narrative, this text is actually intended for us as receivers, insofar as we are linguists who read English, the group Hymes is presumably addressing. First, consider the written text and the rather crude analysis that follows:

- (39) a. *Linguistic Theory and Functions in Speech*
 b. *The late Uriel Weinreich (1966:399) observed:*

- c. "Whether there is any point to semantic theories which
 d. are accountable only for special cases of speech—namely
 e. humorless, prosaic, banal prose—is highly doubtful."
 f. "The purpose of this chapter is to generalize Weinreich's
 g. statement, and to remove the qualification: linguistic
 h. theories accountable only for such cases of speech cannot
 i. be consistently justified. I shall try to bring out the
 j. plurality, priority, and problematic (empirical) status
 k. of functions in speech.
 l. In speaking of "functions," I do not intend to raise here
 m. the many issues that attach to the notion of "functionalism"
 n. in the social sciences, and, more generally, in the philo-
 o. sophy of the sciences and humanistic disciplines. I
 p. use the term first of all because its use by the
 q. Prague School has associated it with the perspective
 r. developed here, and because it does seem the appropriate
 s. general term for a necessary idea. In their methodological
 t. reflections on worlds of human knowledge, scholars such as
 u. Ernst Cassirer and Kenneth Burke have found the question
 v. of function, and, in human action, the question of function
 w. known as purpose, indispensable. That the burden of proof
 x. lies with the advocate of the relevance of concern with
 y. such questions in linguistics today, does not reflect the
 z. nature of language, but the limitations of current linguistics.
 aa. The burden of proof ought to be, and I believe will come to
 bb. be, on those who think that linguistics can proceed
 cc. successfully without explicit attention to its functional
 dd. foundations. (Hymes, 1974, pp. 145–146.)
- (40) **Assumed Familiarity Analysis (TO BE TAKEN WITH LARGE GRAIN OF SALT):**
- a. linguistic theory: U (or U + A)
 functions in speech: U or I_C(speech^U)/?
 b. the late uriel Weinreich: U (or U + A)
 c. whether there is any point to semantic theories which are ac-
 countable only for special cases of speech—namely humor-
 less, prosaic, banal prose: I_C [(that) there is a point to
 semantic theories which are accountable only for special
 cases of speech—namely humorless, prosaic, banal prose]^U
 (semantic theories . . . U (or I_C. . .)]/theories have a point
 to them)/proposition can be true or false
 f. the purpose of this chapter: I_C(this chapter^{E₁})/chapters have
 a purpose (or U)

- Weinreich's statement: I(preceding propositions)/propositions
 which someone_i has observed/written/uttered = someone_i's
 statement
 g. the qualification: I(preceding propositions)/whether P is highly
 doubtful = qualification
 linguistic theories accountable only for such cases of speech:
 I(semantic theories . . . prose^E)/particular to general
 i. I: E_s
 the plurality, priority, and problematic (empirical) status of
 functions in speech: I_C(functions^{U (or E)})/?functions have at-
 tributes . . .
 l. "functions": E
 I: E
 m. the many issues that attach to the notion of "functionalism" in
 the social sciences, and, more generally, in the philosophy
 of the sciences and humanistic disciplines: I_C(notion of func-
 tionalism¹ (functionalism¹(functions . . . ^E)/objects to 'ism')/
 ism = notion)/issues attach to notions + A
 o. I: E
 p. the term: I(functions^E)/metalinguistic assumption
 its use by the Prague School: U (or: I_C(Prague School^U)/Prague
 School used/uses terms
 q. it: E
 the perspective developed here: I_C(here^E)/we are developing a
 perspective here
 r. it: E
 the appropriate general term for a necessary idea: I_C(an
 idea^{?BN} + A)/terms are for ideas + A
 s. their methodological reflections on words of human knowl-
 edge: I_C(worlds of human knowledge^{?U})/?people reflect on
 worlds of human knowledge + A
 t. scholars such as Ernst Cassirer and Kenneth Burke: I_C(Ernst
 Cassirer and Kenneth Burke^U)/member(s) to set (or: U)
 u. the question of function: U (or: I_C(function^E)/function is a
 question)
 v. the question of function known as purpose: I_C(the question of
 function^E)/things are known as other things
 w. that the burden of proof lies with the advocate of the rele-
 vance of concern with such questions in linguistics today:
 I_C(burden of proof¹ (notion of functionalism^E)/notions are
 such that there is a burden of proving them)/burdens lie with
 some person(s)

- y. *the nature of language*: U (or: $I_c(\text{language}^U)/\text{things have a nature}$)
- z. *the limitations of current linguistics*: U (or: $I_c(\text{current linguistics}^U)/\text{current linguistics has limitations}$)
- aa. *the burden of proof*: E
I: E
- bb. *those who think that linguistics can proceed successfully without explicit attention to its functional foundations*: $I_c(\text{that linguistics can proceed} \dots^U \text{ (or: } I_c(\dots) \dots^U \text{)})/\text{people think such things}$

Immediately obvious is the decrease in Evoked entities and the increase in complex Inferrable and Containing Inferrable entities, compared with the oral narrative presented earlier. The figures are given in Table 2.

In spite of (and contributing to) the shakiness of the analysis, a number of systematic differences between the two texts is apparent, all of which indicate additional complexity in the written text. First, in contrast to the situation in the oral narrative, we find here a high degree of METALINGUISTIC inferencing. Although it may be a straightforward matter to infer that what Weinreich observes, [i.e., the quoted sentence in (39c)–(39e)], constitutes *Weinreich's statement* in (39f)–(39g), it is less obvious that part of that quoted sentence—the *wh*-word and the VP *is doubtful*—constitute *the qualification* in (39g).

TABLE 2

Analysis of Subjects and Nonsubjects in Written Text
(by Number)

| | Subjects (12) | Nonsubjects (16) |
|-------------------|---------------|------------------|
| Evoked | | |
| E: | 5 41.7% | 2 12.5% |
| E _s : | 1 8.3% | 0 0.0% |
| E + A: | 0 0.0% | 0 0.0% |
| Total | 6 50.0% | 2 12.5% |
| Inferrable | | |
| I: | 1 8.3% | 3 18.8% |
| I _c : | 4 33.3% | 7 43.8% |
| Total | 5 41.7% | 10 62.5% |
| New | | |
| U: | 1 8.3% | 4 25.0% |
| BN _A : | 0 0.0% | 0 0.0% |
| BN: | 0 0.0% | 0 0.0% |
| Total | 1 8.3% | 4 25.0% |

Second, in addition to the metalinguistic problem, a great deal of complexity arises from the extreme ABSTRACTNESS of many of the entities involved. It is one thing to create an entity out of an already expressed proposition or set thereof, as the *that* requires in the oral narrative in (28ff) (*That's what they did*), but quite another to present a proposition as an entity, as is the case with the sentential subject in (39w)–(39y) in the written text.

Third, and more important, correlating with the large increase in the occurrence, in the written text, of Containing Inferrables is a BLURRING BETWEEN WHAT IS UNUSED AND WHAT IS INFERRABLE which does not arise in the oral narrative. There, the entities are all either assumed to be known to the hearer (e.g., *Holy Redeemer*) or not (e.g., *a friend of mine*), or inferrable on the spot (e.g., *one of them*). In contrast, is the reader of the written text assumed to have known, before reading *its use by the Prague School*, that a certain term was/is used by the Prague School, or is s/he assumed to have known simply that there is/was a Prague School and to infer from the NP that the Prague School used the term? Similarly, is the reader expected to have known, before reading *scholars such as Ernst Cassirer and Kenneth Burke*, that there is a class of scholars to which Cassirer and Burke belong, or is s/he expected simply to have known that Cassirer and Burke exist and to infer from the NP that there is a class of scholars of which they are members? Obviously, the answer is that there can be no sharp line in such texts: What is an Unused entity for one intended reader will be a Containing Inferrable for another; the latter reader, however, having fewer "staples" in his "closet," so to speak, will have much more work to do to construct the desired discourse-model.

A fourth difference between the two types of texts concerns the CULTURAL ASSUMPTIONS required for the inferencing. Note that both the oral narrative and the written text require certain assumptions; the two differ, however, in the nature of these required assumptions—for example, *Doors have knobs* versus *Notions are such that there is a burden on someone of proving them*. That is, like the entities themselves, the cultural assumptions required for inferencing in the written text are of a highly complex and abstract nature.

A fifth difference, and perhaps the one causing the greatest difficulty for the linguist analyzing texts like the written one, is not so much the abstractness of the entities or assumptions, per se, as the sheer SIZE OF THE ENTITIES. Many of the entities in the written text, but not in the oral text, are themselves made up of entities and attributes, the composing entities themselves often being complex.

Finally, when we consider the syntactic environments of the entities,

still other differences appear. In contrast to the oral narrative, just one half of the subject NPs in the written text represent Evoked entities, and only one-eighth of the nonsubjects do, as shown in Table 2. Further, over two-fifths of the subjects are Inferrables, mostly of the Containing Inferable type. No Brand-new entities occur at all, but a few Unused ones do, in subject as well as nonsubject position. (As suggested earlier, a more knowledgeable reader than I would of course have more Unused entities and fewer Inferrables.) Note, however, that, in both texts, subjects are more likely to be Evoked than Inferrable and more likely to be Inferrable than New.

5. AREAS FOR FURTHER STUDY

As this chapter is too inconclusive to warrant a conclusion, I shall instead close by outlining what I would like to see it lead to. The first step is obviously to refine/revise/replace the taxonomic model presented here. Although it is fairly adequate in its present form for texts like the oral narrative, it is far too crude for texts like the written one. The large and diverse number of NPs that now may fall under the category *Inferrable*—ordinary NPs, gerunds, *wh*-clauses, *that*-clauses, etc.—can perhaps be broken down into more manageable and meaningful subgroups.

One problem that has been skirted here is whether Assumed Familiarity is binary or, as in the taxonomy presented here, ternary, or something else. Although the data seem to indicate fairly clearly that it is not a continuum, it is not obvious at this point whether a binary or a ternary division is more appropriate. Intuitively, one can think of Inferrables as being classified under New—they were not previously in the discourse-model—and, at the same time, as being classified, along with Evoked, under some category Old—they are made up of old parts. Before we can motivate a decision for one or the other, however, further research is needed, included psycholinguistic research on the processing of discourse.

Once an adequate taxonomy of Assumed Familiarity is arrived at, it can, coupled with already existing taxonomies of linguistic form, lead to a deeper and clearer account of the correlation between form and understanding in natural language. Such an account can aid in several specific areas of research. In particular, it can shed light on questions of FUNCTIONAL SYNTAX by enabling analysts to describe explicitly, precisely, and, hopefully, in universal terms, the salient features of the environments in which different syntactic constructions occur.¹⁷ Certainly, it will

¹⁷ It is encouraging that Richard Rhodes reports (personal communication) that the taxonomy of givenness, presented here holds up nicely for Ojibwa, a language that appears on the surface to have a very different pattern of information distribution from English (see Tomlin and Rhodes (1979)).

aid in DISCOURSE ANALYSIS and, whether desired or not, it will force a distinction between different STYLES of discourse, by characterizing the nature and complexity of the entities and inferences involved and the morphology and syntax of the NPs representing those entities.¹⁸ Further, it might prove useful in research on THE DEVELOPMENT OF DISCOURSE COMPETENCE in children (and adults?) and on THE PROCESSING OF DISCOURSE by children and adults. Finally, one would hope that it would help us to arrive at a clearer understanding of the other kinds of givenness, predictability and saliency, discussed earlier.

In a somewhat different vein, an adequate taxonomy of Assumed Familiarity and the subsequent account of the form–understanding correlation will aid in the study of what reading and writing entail beyond decoding and encoding. Contra Olson 1977, where it is claimed that written language is explicit whereas oral language leaves much out, the analyses presented here suggest that just the opposite is the case: The comprehension of formal, literary discourse depends a great deal more on inference, quantitatively and qualitatively, than the processing of informal, colloquial discourse. More generally, such research will have practical relevance wherever the control of linguistic complexity is at issue, wherever it is the case that the sender has less-than-perfect intuitions about the receiver and wants to be understood.

ACKNOWLEDGEMENTS

Parts of this chapter appeared in Prince (1979), and an earlier version was presented at the San Diego State University Linguistics Colloquium, April 24, 1979. I should like to thank Cathy Ball, John Fought, Erving Goffman, Jeff Kaplan, Tony Kroch, Susumu Kuno, Bill Labov, Gerry Prince, and Bonnie Webber for their insightful comments and criticisms.

REFERENCES

- Bransford, J. D., and Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717–726.
- Bransford, J. D., and Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.
- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and Topic*. New York: Academic Press. Pp. 25–55.

¹⁸ Note, for example, that the sort of “blurring” between Unused and Containing Inferable entities that arises in the written text seems to covary with a text’s having more than one intended receiver and/or a text’s sender not knowing the receiver(s) well. See Rubin 1978 for a discussion of these and other parameters along which discourses may differ.

- Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon.
- Christopherson, P. (1939). *The articles: A Study of Their Theory and Use in English*. Copenhagen: Munksgaard.
- Clark, H. and Haviland, S. (1977). Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse Production and Comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates. Pp. 1-40.
- Goffman, E. (1967). On face-work. In E. Goffman, *Interaction Ritual*. Garden City, New York: Anchor Books. Pp. 5-46.
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English. Part 2. *Journal of Linguistics*, 3, 199-244.
- Halliday, M. A. K., and Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Haviland, S. and Clark, H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Hawkins, J. A. (1978). *Definiteness and Indefiniteness*. Atlantic Highlands, N.J.: Humanities Press.
- Horn, L. (1978). *Présupposition, thème, et variations*. Unpublished Manuscript.
- Hymes, D. (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.
- Karttunen, L. (1971). Discourse referents. In J. McCawley (Ed.), *Syntax and Semantics 7: Notes from the Linguistic Underground*. New York: Academic Press. Pp. 363-386.
- Kuno, S. (1972). Functional sentence perspective. *Linguistic Inquiry*, 3, 269-320.
- Kuno, S. (1974). Lexical and contextual meaning. *Linguistic Inquiry*, 5, 469-477.
- Kuno, S. (1978). Generative discourse analysis in America. In W. Dressler, (Ed.), *Current Trends in Textlinguistics*. Berlin and New York: de Gruyter. Pp. 275-294.
- Kuno, S. (1979). *On the Interaction between Syntactic Rules and Discourse Principles*. Unpublished manuscript.
- Labov, W., and Fanshel, D. (1977). *Therapeutic Discourse*. New York: Academic Press.
- Olson, D. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review*, 47, 257-281.
- Prince, E. F. (1978a). A comparison of *wh*-clefts and *it*-clefts in discourse. *Language*, 54, 883-906.
- Prince, E. F. (1978b). On the function of existential presupposition in discourse. In D. Farkas, W. Jacobsen, and K. Todrys (Eds.), *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*. Department of Linguistics, University of Chicago. Pp. 362-376.
- Prince, E. F. (1979). On the given/new distinction. In W. Hanks, C. Hofbauer, and P. Clyne (Eds.), *Papers from the Fifteenth Regional Meeting of the Chicago Linguistic Society*. Department of Linguistics, University of Chicago. Pp. 267-278.
- Rubin, A. D. (1978). *A Theoretical Taxonomy of the Differences between Oral and Written Language* (Technical Report No. 35). Bolt Beranek and Newman, Inc. Cambridge, Mass.
- Sacerdoti, E. (1977). *A Structure for Plans and Behavior*. New York: Elsevier-North-Holland.
- Stenning, K. (1978). *On why making reference out of sense makes it so hard to make sense out of reference*. Paper presented at the Sloan Workshop on Indefinite Reference, University of Massachusetts, Amherst, December 2.
- Terkel, S. (1974). *Working*. New York: Avon Books.
- Tic Douloureux, P. R. N. (1971). A note on one's privates. In A. Zwicky et al. (Eds.), *Studies Out in Left Field: Defamatory Essays Presented to James D. McCawley on the Occasion of His 33rd or 34th Birthday*. Edmonton and Champaign: Linguistic Research, Inc. Pp. 45-52.

- Tomlin, R., and Rhodes, R. (1979). The distribution of information in Ojibwa texts. In W. Hanks, C. Hofbauer, and P. Clyne (Eds.), *Papers from the Fifteenth Regional Meeting of the Chicago Linguistic Society*. Department of Linguistics, University of Chicago. Pp. 307-320.
- Webber, B. L. (1978). Jumping ahead of the speaker: On recognition from indefinite descriptions. Paper presented at the Sloan Workshop on Indefinite Reference, University of Massachusetts, Amherst, December 3.
- Wolfson, N. (1976). *The Conversational Historical Present in American English Narrative*. Unpublished doctoral dissertation, University of Pennsylvania.