LANGUAGES AND LINGUISTICS SYMPOSIUM
1975

April 7-8, 1975
Brigham Young University
Linguistics Department
and
Language Research Center
# Table of Contents

A Second Look at Unpassives ...................................... Joel Hust ................................. 1.1

Transformational Grammar and Junction Grammar:
Contrasting Metaphors ........................................... Melvin J. Luthy .............................. 2.1

A List of Verbs Most Needed For Communication
In A Foreign Language Context ................................. Thomas F. Rogers .......................... 3.1

Innovative Techniques of Russian Instruction
In The United States 1970-74 .................................... Donald K. Jarvis ............................ 4.1

Literary Computing: Some Methodological Pitfalls ........... Steven Sondrup .............................. 5.1

Chaucer's "Underweighted" Line: Some Computer-Assisted Steps Toward a Reassessment .......... Karen Lynn ................................. 6.1

Usage Patterns of Thou, Thee, Thy and Thine
Among Latter-Day Saints ....................................... Don Norton ...................................... 7.1

Languages of Cultures ........................................... V. Lynn Tyler ................................. 8.1

Toward a Language Instructional Model ......................... Robert W. Bush ................................ 9.1

Some Practical Applications of Junction Grammar
In Computational Linguistics .................................... Charles D. Bush ............................ 10.1

Presentations by the following participants in the Languages and Linguistics Symposium are not included in this publication. Those interested in obtaining copies of or information about these presentations may contact the authors at the addresses listed below:

**The Study of Comparative and Historical Linguistics**
- Raimo Anttila
  Department of Linguistics
  UCLA
  Los Angeles, Calif.

**Languages and Bilingual/Bicultural Education:**
- What the Government and Others are Doing
  - LeRoy Walser
    2001 N. Kenilworth
    Arlington, Virginia 22210

- "A Letter From Israel Whiton, 1851"; poem and linguistic analysis
  - Clinton F. Larson
    A-224 JKBA
    Brigham Young University
    Provo, Utah 84602

**Discourse Analysis: Some Major Strategies**
- Willis L. Pitkin Jr.
  Department of English
  University of Utah
  Logan, Utah 84322

**American Dialects**
- Harold B. Allen
  162 Klaeber Court
  University of Minnesota
  Minneapolis, Minn. 55455

**Teaching a Portion of Skills Needed for Reading Comprehension in a Foreign Language**
- Marion R. Pomeroy
  333 MCKB
  Brigham Young University
  Provo, Utah 84602

**English Analysis: A Synergistic Approach**
- Daryl Gibb
  150 Fletcher Building
  Brigham Young University
  Provo, Utah 84602
A SECOND LOOK AT UNPASSIVES

Joel Hust

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
A SECOND LOOK AT UNPASSIVES

Joel Hust

In her paper, "Nonsources of Unpassives," Dorothy Siegel (1973) presents a number of problems relating to sentences such as (1).

(1) a. The president's blunder was unreported (in the press).
b. The brass knuckles were unpaid for.
c. Antarctica is uninhabited (by man).
d. The Garbage went uncollected.

Siegel refers to such sentences as "unpassives." Unpassives are characterized by the presence of an auxiliary (either be or go) and a past-participle prefixed with un-, which Siegel refers to as the "unparticiple." However, the most striking feature of unpassives, Siegel maintains, is that there exist no well-formed active sentences from which they could plausibly be derived. Note the impossibility of the sentences in (2), which one would be apt to posit as the underlying active counterpart of the grammatical sentences in (1).

(2) a. *The press unreported the president's blunder.
b. *Someone unpaid for the brass knuckles.
c. *An Antarctica is uninhabited.
d. *A (went) uncollected the garbage.

Siegel convincingly argues that even though the present theory would allow the derivation of (1) from underlying structures similar to (2), in spite of the fact that the sentences in (2) are ungrammatical, such as analysis is to be rejected on a number of grounds. Hence, one is faced with the task of determining an alternative source for unpassives, and more specifically, as I see the problem, determining the source of the unparticiple.

Siegel considers three possible sources for unpassives and presents arguments which lead her to reject each of them. First, she considers the possibility that the prefix un- is a surface realization of the same NEG that underlies not. Then unpassives could be derived from active sentences by a series of transformations, including the passive, as well as a transformation which relabels the participle resulting from the passive as an adjective, and a transformation which attaches NEG to the adjectival participle. This analysis is rejected due to the vast differences in the syntactic patterning and behavior of un- and other realizations of NEG, such as not. Siegel then considers the possibility that unpassives are not derived at all, but rather that the unparticiple are deep adjectives and unpassives are generated directly by phrase structure rules. I will return to Siegel's argument against this possibility below. Finally, Siegel considers the claim attributed to Ross, that unparticiple are the surface realization of without-being + PARTICIPLE phrases: that is, the team went unbeaten derives from the team went without being beaten. This claim is refuted by the lack of parallelism in the distribution of unpassives and without-being + PARTICIPLE phrases. Siegel's well-considered presentation makes it abundantly clear that we are dealing with a range of data which is not only highly interesting and complex, but also extremely resistant to satisfactory analysis. Siegel goes on to outline the essential features which she believes the correct analysis of unpassives must incorporate. Her proposals will be taken up later in the course of this paper.

At this point it would be well to take a closer look at the surface properties of unpassives and unparticiple. Recent work in generative grammar has shown that in most cases surface structures are very similar to their corresponding deep structures. Thus, a close look at the surface properties of unpassives should give us some clues as to their deep structure source. First and most importantly, it is to be noted that unparticiple are surface adjectives, not verbs. Siegel notes this fact and presents a substantiating argument, which I will briefly summarize.

The argument rests on the fact that negative un- appears only as a prefix to adjectives or words which are derived from adjectives. Thus, we have adjectives like unkind, untrue, nouns like untruth, unkindness, and adverbs like ungracefully, unendingly, as well as participles like uncollected and uninhabited. On the other hand we find no cases where un- is attached to underived nouns, e.g. *unhonor, *unlove, to nouns derived from verbs, e.g. *unintegration, *unarrival, *unresistance, or to underived adverbs or prepositions, e.g. *unnow, *unthrough. And of course there are no cases where un- is attached to verbs, e.g. *to unlove, *to unknow, etc. Thus, in the simplest analysis, Siegel concludes, un- is attached only to adjectives. The labeled bracketing of untruth must therefore be (un(true)adj).adj-th). Likewise the labeled bracketing of uninhabited must be (un(-((inhabit)v -ed)adj)adj). Rather than (un-(inhabit)v -ed)adj. Otherwise unpassives constitute the single exception to the labeled bracketing of words prefixed with un-. While Siegel's basic observation that un- occurs only as a part of adjectives and their derivatives is correct, there remains some question regarding the labeled bracketing of unparticiple since there are instances of unparticiple, the positive counterparts of which are not adjectives. Consider the following pairs:

(3) a. The thief was seen.
b. *A seen thief stole the cookies.

(4) a. The thief was unseen.
b. An unseen thief stole the cookies.

But leaving the question of the labeled bracketing of unparticiple aside there seems to be no doubt that unparticiple function as surface adjectives. In fact the case exemplified by (3) and (4) above lends further support to this conclusion. Typically adjectives can occur in the environment Art ___ N, whereas verbs cannot. Above unseen is grammatical in this environment, whereas seen is not. To further illustrate this point consider the sentence in (5).

(5) a. The red book was on the table.
b. The unread book was on the table.
c. The read book was on the table.

There is on the other hand an environment in which only verbs can occur and adjectives cannot, as
Illustrated by (6).

(6) a. The book remains to be read.
b. *The book remains to be red.
c. *The book remains to be unread.

Notice that in both (5) and (6) the distribution of the unparticiple parallels that of the adjective, not that of the verb. A further case which demonstrates that unparticiples behave more like adjectives than like verbs is illustrated in (7) and (8).

(7) a. The package was open.
b. The package was opened.
c. The package was unopened.

(8) a. *The package was being open.
b. *The package was being opened.
c. *The package was being unopened.

As Siegel notes, unpasses can't take progressive aspect. Note that in the above examples the unparticiple and the lexical adjective behave the same way with respect to the progressive, both being ungrammatical. The related verbal participle, however, is grammatical with the progressive.

The morphological, distributional and selectional properties of unparticiples thus lead us to the conclusion that they are adjectives and not verbs, at least in surface structure. In view of this it would seem natural to assume that unparticiples might be lexically, rather than transformationally derived and inserted directly at the level of deep structure. In fact, if one accepts the assumptions of the Extended Lexical Hypothesis as presented in Jackendoff (1972), an extension of the Lexical Hypothesis developed in Chomsky (1970), then this is the only possibility, since within this framework processes of derivational morphology are strictly restricted to the lexicon.

Before attempting to ascertain whether such a standpoint can be maintained, however, let's consider some of the problems any solution which involves deriving unpasses from active sentences will have to face. This is the standpoint adopted by Siegel. She suggests that unpasses derive from underlying active sentences containing verbs which aren't prefixed with un- and that un-, which is present in deep structure as an independent lexical item, is attached to the verb at some point in the derivation by a transformational rule. Siegel states that this attachment transformation would have to follow the passive and also a transformation relabeling the participle which results from the passive as an adjective. Her analysis thus does away with the need to posit unoccurring abstract verbs bearing the prefix un-, such as those in (2), and also avoids positing NEG as the underlying source for un-.

The primary problem with such an analysis is the fact that unpassivization doesn't apply freely in sentences which otherwise freely passivize. The range of complements is severely limited in unpasses. Most striking is perhaps the fact that many unpasses are unacceptable when they contain by agent phrases. Consider the sentences in (9).

(9) a. Tom painted the door.
b. The door was painted by Tom.
c. The door wasn't painted by Tom.
d. *The door was unpainted by Tom.

If unpasses are derived from actives by a process which includes the passive transformation one would expect by-phrases, the presence of which is a crucial part of the structural description of the transformation, to be perfectly normal and in fact to be expected, rather than questionable or unacceptable.

Siegel points out three types of verbs which can passivize but cannot unpassivize. One type is ruled out by a morphological constraint which prohibits the stacking up of negative prefixes. The other two types are relevant to the discussion at this point, as they demonstrate further environments in which verbal participles can occur but unparticiples cannot. The first type consists of verbs which take two prepositional complements, e.g., make away with, put up with, etc.

(10) a. The old law was done away with.
b. The old law was undone away with.

The second type consists of verbs sub categorized for NP PP. This includes idioms like take care of X, take advantage of X as well as verbs which take indirect objects, give X to Y, and after dative movement, give XY.

(11) a. Susie will be taken care of.
b. *Susie will be untaken care of.
c. *Care will be untaken of Susie.

(12) a. The gift was given to the school.
b. *The gift was ungiven to the school.
c. The gift was given the school.
d. *The gift was ungiven the school.
e. The school was given the gift.
f. *The school was ungiven the gift.

Siegel attempts to account for these two types of verbs by her generalization (24), a constraint on surface structures. I quote: "(24) In surface structure, there may not be a more highly stressed word in the VP than the word to which un- is attached."

(p. 311) However, this generalization is false. Many of the example sentences Siegel gives as instances of grammatical unpasses, in fact, constitute counterexamples to (24). I will list a small sampling.

(13) The tests were uncomplicated by mishaps.
(Siegel's 18a.)

(14) Sam was unstayed by minor obstacles.
(Siegel's 20a.)

(15) The tragedy was unreported in the press.
(Siegel's 23a.)

The stress patterns indicated above are both predicted by the NSR (Chomsky and Halle, 1968) and confirmed by every native speaker to whom I've presented this data. In all cases the most highly stressed word in the VP is not the word to which un- is attached, but rather the object of a PP complement at the end of the VP. Hence generalization (24) cannot be maintained, and the cases illustrated by (10)-(12) remain unexplained. However, if one assumes that unparticiples are lexically derived adjectives the facts illustrated by (10)-(12) follow as an automatic consequence since verbs, but not adjectives, are subcategorized for PP PP, NP PP and NP.

Furthermore, unpasses don't take participial complements, even when the associated verb does.
Unpassives don't take infinitival complements, even though the associated verb does.

Examples (16)-(18) indicate that there is a further range of sentences which passivize but don't unpassivize.

Further problems for a derivation like that outlined by Siegel are illustrated by the data presented in (3) and (4), above, which indicated that there are verbs from which unparticipial adjectives are derived, but from which participial adjectives are not. I will repeat the examples below.

(19) a. The thief was seen.
   b. *A seen thief stole the cookies.
   a. The thief was unseen.
   b. An unseen thief stole the cookies.

Other verbs of this type are, e.g. liked, noticed, cared for, read, heeded, etc.

A further problem for a transformational approach to the derivation of unpassives is that they cannot take progressive aspect. Siegel attributes this to the stativity of unparticiples. All unparticiples are stative, even if the verbs from which they are derived are nonstative. At the moment I see no way to prevent sentences like (21) if the underlying verb is nonstative (short of global statement to the effect that a verb can't take progressive aspect if at some later time in the derivation it will become stative [due to the affixation of un-]).

(21) *The West was being unexplored by trappers.

If, on the other hand, unparticiples are inserted directly they will already be specified as stative and no global statement would be necessary.

A related problem has to do with the fact that get may co-occur with regular participles but not with unparticiples:

(22) a. Sam got acquitted by the jury.
    b. *Sam got unacquitted (by the jury ).

(Siegel's examples)

It seems that get can co-occur only with nonstative verbs and adjectives and a certain subset of stative adjectives. I don't know exactly how to characterize this subset of stative adjectives at present, but they seem to denote the result of a just completed process. Compare the examples in (23) with (24).

(23) a. Sam got red (because he was embarrassed).
    b. Bill got sick.
    c. Mary got smart in a hurry.

    b. The door got open.
    c. *The babies got alive.

The adjectives in (24) seem to denote simply an existing state and not one that has just resulted from some change or action. However these adjectives are to be specified (perhaps [-perfective] or [-effective]), the unparticiples, if they are underlying adjectives, would carry the same feature and the parallel in the selectional restrictions of this subset of stative adjectives and the unparticiples would be automatically accounted for. That is, the unparticiples would belong to this subset. Otherwise the unparticiples would have to be assigned this feature later in the derivation when un- is attached to the underlying nonstative verb.10

Two more properties of unparticiples, which seem related to one another, may follow as an automatic consequence from the fact that they are inserted directly as stative adjectives. The first is the fact that subject oriented adverbs don't co-occur with unparticiples, as Siegel illustrates with the following examples.

    b. *Nixon's gaffe was enthusiastically unpublicized.

(26) a. The proposal was enthusiastically defeated.
    b. *The proposal was enthusiastically undefeated.

Notice that an adverb like enthusiastically presupposes a deep structure human or + agent perhaps subject and a nonstative predicate. If we take the structure underlying the passive (26a.) to be roughly (27) (with many details omitted), a more or less traditional formulation, the reason that a subject oriented adverb is possible is apparent.

(27)

\[
\begin{array}{ccc}
  \text{S} & \text{NP} & \text{VP} \\
  \text{someone} & \text{adv} & \text{defeated} \\
  \text{enthusiastically} & \text{the proposal} & \text{by Pass.}
\end{array}
\]

The verb, defeated, is of course nonstative and the subject, someone, is [+human] and hence can take an adverb like enthusiastically. (If X or delta or "nothing" is posited as underlying subject a [+human] interpretation assigned by semantic rule is at least possible, it seems to me.) The underlying structure of (26b.), on the other hand, would have to be something like (28).

(28)

\[
\begin{array}{ccc}
  \text{S} & \text{NP} & \text{VP} \\
  \text{The proposal} & \text{cop} & \text{adj.} \\
  \text{was} & \text{adv.} & \text{undefeated} \\
  \text{enthusiastically}
\end{array}
\]
Since the subject in (28) is +human and the predicate adjective is +stative an adverb like enthusiastically could not occur. Since (29) has a nonstative verb and thus one would predict that a subject oriented adverb like enthusiastically could not occur. This prediction is borne out by the ungrammaticality of (29).

The second property of unpassives which I believe can be similarly explained is the fact that when unparticiples co-occur with modals, like can, the only reading for the modal is the epistemic one. This is illustrated by Siegel with the following examples.

(30) a. The garbage can be uncollected. 
   (1 reading) 
   b. The garbage can't be collected. 
   (2 readings)

Since the root sense of the modals is subject oriented, as opposed to the epistemic sense, which is speaker oriented, modals in the root sense must have a human subject as well as a nonstative predicate. Since (30b.) has a nonstative verb and can have a human subject at an underlying level, deriving from a structure analogous to (27), the root sense of the modal is possible. If the unparticiple in (30a), however, is inserted directly at the level of deep structure, roughly analogous to (28), then (30a) has no +human subject and no nonstative predicate at any point in the derivation and hence the subject oriented (root) sense of the modal is impossible.

Before turning to Siegel's arguments against direct insertion of unparticiples, a comment on one more problem with the approach suggested by her seems appropriate. Siegel's class of unpassives seems to me to be somewhat artificially limited. Siegel states:

If unpassives were not directly related to verbs through the unpassive transformation, then we would expect to find X was unY'd (by Z) where Y is a verb which cannot passivize.

If there were a root in an unpassive such that the root could not passivize, then this would be strong evidence against deriving unpassives from actives. 6

Apparent counterexamples such as uncollected (by), unaccustomed to, unexampled by, and unabashed by are not real unpassives, for they do not occur in all environments in which true unpassive participles can occur. For example, all unpassives can co-occur with go. Thus, we get the garbage uncollected but not the discovery went unparalleled.

First of all, I disagree with Siegel's grammaticality judgement with regard to her sentence containing unparalleled. The sentence, this discovery went unparalleled in the history of science, not only sounds normal to my ear, but in fact sounds SO normal as to be a cliché. However, Siegel is right in observing that there is no active counterpart with the same meaning. Furthermore, it seems to me that a generalization would be missed if examples like those given in Siegel's footnote 6 are not explained in the same manner as "true unpassives," at least as regards the source and attachment of un-. The similarities far outweigh the differences and hence I feel that such examples do in fact represent the "very strong evidence against deriving unpassives from actives" which Siegel speaks of above.

To summarize briefly, we have seen that a lexical derivation and direct insertion of unparticiples as stative adjectives is to be preferred to a syntactic derivation for the following reasons:

1. The occurrence of un- is peculiar to adjectives.

2. The distribution of unparticiples parallels the distribution of adjectives more closely than that of verbs.

3. The full range of passivizable actives don't unpassivize.

4. The subcategorization of unparticiples is not identical to that of the corresponding verbs. Specifically:

   a. By-phrases are often unacceptable.

   b. Unparticiples can't occur in the environments: PP PP NP PP and NP, even though their corresponding verbs can. Furthermore this fact is unexplained by Siegel's generalization (24), but follows automatically if unparticiples are adjectives.

   c. Unparticiples are impossible with participial and infinitival complements, which indicates further subcategorizational differences.

5. There exist unparticiples for which there is not corresponding positive participial adjective.

6. Selectional properties of unparticiples don't correspond to those of the corresponding verbs, e.g. stativity.

7. Subject oriented adverbs and subject
oriented modal readings are impossible with unpassives, but not with the corresponding passives.

8. There exist unparticiples which derive from verbs which don't passivize, or have no positive active counterpart, e.g. unparallelled.

If unparticiples are directly inserved as adjectives points 1-8 cease to be problematic.

Now let us turn to Siegel's arguments against the direct insertion of unparticiples. Above we have seen that there are extensive differences between unparticiples and the verbs to which they are related. There are, to be sure, many similarities, however. Most of Siegel's arguments against a direct generation of unpassives are based on the fact that such similarities exist, which could not be accounted for if unparticiples were not transformationally derived from their corresponding verbs. However, Chomsky (1970) suggests that similarities and regularities among derivatives of a common root should be accounted for by redundancy rules in the lexicon. Such a rule (modeled on Chomsky's informal statement on p. 213) could state (very roughly) that if a V₁ is specified with the feature + NP, then a lexical item (un-[(V₁)-ed]) adj. is automatically specified with the feature +NP, be ___. It seems to me that such rules could in principle easily account for the dissimilarities between unparticiples and their related verbs. In fact, Jackendoff (1972) claims that such rules not only allow us to capture the notion of "separate but related lexical items" (p. 23) but in fact make somewhat different and more satisfactory predictions than transformations in a number of cases.

For purposes of this paper let us assume the existence of the lexical redundancy rule roughly and informally described above. Given such a rule, can Siegel's arguments against the direct insertion of unparticiples be overcome? In order to attempt to answer this question let's take a look at some of the specific arguments Siegel presents. Her first argument involves verbs which make propositional complements. With many of these verbs, like tamper with, hint at, accede to and pay for, passivization can take place giving sentences like (31). (Siegel's examples.)

(31) a. Our wiretaps have been tampered with.
       b. The convention was paid for by big corporations.
       c. The convention was unpaid for.

In such cases we also find grammatical unpassives.

(32) a. Our wiretaps were untampered with.
       b. The convention was not paid for.

However, in other cases passivization is impossible.

(33) a. These duties devolve on the Vice President.
       b. The Vice President is devolved on by these duties.
       c. This result follows from your hypothesis.
       d. *Your hypothesis is followed from by this result.

In just those cases where it is impossible for the prepositional objects to prepose, unpassivization is also impossible.

(34) a. *The Vice President is undevolved on by these duties.
       b. *Your hypothesis is unfollowed from by this result.

Siegel maintains that if unparticiples are directly inserted a generalization is lost. If unparticiples are transformationally derived then the grammaticality of (32) follows from the grammaticality of (31), since in both cases the object can be preposed. Likewise the grammaticality of (34) follows from the ungrammaticality of the starred examples in (33), since the passives can't apply in these cases and the objects can't be preposed. On the other hand, Siegel maintains, if unpassives aren't derived from underlying activities then the examples in (34) must be ungrammatical for some reason unrelated to the ungrammaticality of (33b. & d.). The ungrammaticality of (34) can be related to the ungrammaticality of (33b. & d.) within a lexicalist framework, however.

In order to see this let's consider the passive transformation. Chomsky (1973) formulates the passive with the S. D. (35) with the condition that VY must form a semantic unit.

(35) X, NP, VY, NP, Z

Let us, however, try to make the stronger claim that the third term of (35) also forms a syntactic unit, that is, that it is a syntactic constituent. We will then analyze tamper with and pay for as main verbs (V) with the substructure verb + preposition (v p)₁² and the subcategorizational feature + NP, and we will analyze follow and devolve as main verbs (V) with the substructure (v) and the subcategorizational feature + PP. Our wiretaps and the convention are then analyzed as deep structure direct objects of verbs and not prepositional objects. On the other hand from your hypothesis and on the President are prepositional phrases. There is independent evidence that an analysis along these lines is correct. For example, Emonds (1972) points out that only NP's and PP's can occupy the focus position in cleft sentences. Note that the sentences in (35) are grammatical but those in (36) are not.

(35) a. It's from your hypothesis, that this result follows.
       b. It's on the President, that these duties devolve.

(36) a. *It's with our wiretaps, that the enemy tampered.
       b. *It's for the convention, that big corporations paid.

Such sentences indicate that the strings with our wiretaps and for the convention are not PP's but from your hypothesis and on the President are. We can now give the passive transformation the formulation (37).

(37) NP Aux V NP X by Pass 4 2 be en 3 0 5 6+1

Furthermore, we will assume (38)-(40) to be partial expansions of a portion of the phrase structure rules of the base.
Such a formulation makes the claim then that the constituent structure of the string underlying pseudo-passives must be (41), not (42), otherwise the S.D. for passive is not met.

The ability of a string to pseudo-passivize is then simply a function of its constituent structure. If this claim is correct it makes accounting for pseudo-passives with an interpretive device like Jackendoff's Thematic Hierarchy Condition unnecessary. (Jackendoff, 1972) This point is perhaps worth a slight digression. Consider the sentences in (43).

a. The thief slipped into the closet.
b. The soap slipped into the closet.
c. The closet was slipped into by the thief.
d. *The closet was slipped into by the soap.

The claim is that we are dealing with two distinct (but related) lexical items, slip and slip into, and furthermore slip into is selectionally restricted to a +human (or +animate or +agent) subject. (43c.) would have the underlying structure (44).

Now note that our redundancy rule relating unpassive and their corresponding verbs is specified in terms of the feature +NP. We have seen that the verbs in (32) have this feature, but those in (33) to not; they are subcategorized +PP. Thus the absence of corresponding unparticiples is accounted for. Such an account in fact seems superior to Siegel's because the verbs in (33) now don't have to be exceptionally specified as not undergoing the passive. They don't passivize, simply because they don't meet the S.D. for passive. And furthermore we have been able to show that the impossibility of passivization for these examples and the lack of related unpassives is due to the same fact, namely that the underlying verbs are not specified +NP.

Siegel's second argument is overcome in similar fashion. It concerns verbs like care for, which cannot take optional object deletion.

a. Sam cares for the animals.
b. *Sam cares for.
c. The animals were uncared for.
d. *The animals were uncared for the bananas.

Siegel argues that animals must be the deep object in (46c.) otherwise we can't account for the impossibility of (46d.). However, the subcategorization relations of (46a.) and (46d.) are accounted for by our redundancy rule and the fact that uncared for the bananas
is ungrammatical is due to the fact that adjectives (uncared for) are not specified for NP complements.

Siegel's third argument against the direct insertion of unparticiples has to do with agent phrases. Siegel notes that most unpasses sound best without agent phrases. This was cited as an argument against a derivational approach earlier in this paper. However, there are cases, as Siegel points out, where unpassive participles require the presence of agent phrases. Consider the sentences below.

(47) a. The tests went uncomplicated by mishaps.
   b. *The tests went uncomplicated.

(48) a. Sam was unstayed by minor obstacles.
   b. *Sam was unstayed.

Siegel concludes that since verbs, but not adjectives, are subcategorized for the manner adverb by NP, it follows that the verbal roots underlying the unparticiples in the a. examples must be responsible for the appearance of the agent phrases. Siegel further points out that it is sometimes the case that agent phrases in passives contain prepositions other than by.

(49) a. The tragedy was reported in the press.
   b. The problem was discussed among his friends.

In just these cases where a passive can take an agent phrase introduced by a preposition other than by the unpassive can also.

(50) a. The tragedy was unreported in the press.
   b. The problem was undiscussed among his friends.

Once again, however, these facts can be handled by redundancy rules in the lexicon, probably a reformulation of our approximation at a rule, formulated earlier. Such a rule would specify that if there exists a V1 with the feature +NPj Aux NP, a lexical item of the form (un-((V1)-ed)) specified with the feature +NP Aux be adj by NP. Now what is striking about such a redundancy rule is that it looks so similar to the passive transformation. And if the lexicon contains rules which just duplicate syntactic transformations this would seem at first glance to constitute a proliferation of rules and a loss of generality. Joe Emonds (personal communication) has pointed out to me that if a strictly lexicalist position is to be maintained the lexicon must contain rules which are very similar to syntactic transformations, and the fact that redundancy rules are so similar to syntactic transformations may actually represent a gain in generality, rather than a loss. For example, the similarity between lexical redundancy rules and syntactic transformations might be utilized to make maximally concise statements in the lexicon. That is, part of the entry of a transitive verb might have some very rough similarity to (51)

(51) V
   : (features)
   : adj. un- + passive

Furthermore Jackendoff (1974) posits lexical redundancy rules which refer to the rules of the base. So there seems to be evidence emerging that there is an overlap in rules operating in the syntax and those operating in the lexicon. And upon reflection this doesn't seem unreasonable. One would expect, for example, that the constituent structure of complex lexical items, like idioms for example, would be similar or identical to structures generated by the base, and one would expect rules operating on those structures to be similar to those operating on the output of the base. Katz (1973) presents an extremely interesting discussion of idioms, their constituent structure and insertion into the base, which makes clear the necessity of parallelism of lexical and syntactic structures.

In summary, it seems that a good case can be made for the direct insertion of unparticiples. However, attempts at an analysis of structures such as the unpassive make it clear that much more work concerning the structure of the lexicon and its interaction with the rules of the syntactic component is necessary before analyses of such complex data can be evaluated with confidence.

NOTES
1 These examples and those in (2) below, including parentheses, are Siegel's. The agent phrases are in parentheses because many people find unpasses with agent phrases unacceptable.

2 Concerning the un- in question I quote Siegel's footnote 3:

The un- under scrutiny in this paper is negative un-, not the privative un- which shows up in words like undress, unbutton, and unsaddle. Privative un- implies a reversal of the action specified in the verb. Aside from its semantic distinctness from negative un-, privative un- differs in another way from the negative un-. Only privative un- shows up on verbs: to unbutton, *to uncollect. There are yet other differences. For example, privative un- can take progressive aspect, whereas negative un- cannot:

(iii) a. The blouse was being unbuttoned.
   b. *The garbage was being uncollected.

Also, privative un-, unlike negative un-, can co-occur with subject selected adverbs like enthusiastically:

(iv) a. The blouse was unbuttoned enthusiastically.
   b. *The garbage was uncollected enthusiastically.

The existence of two un-s predicts that there should be ambiguities in un-words in which the action implied by the verb can be both reversed and negated; and in fact, such ambiguities exist: unbuttoned, unzipped.
It is important to keep this distinction in mind throughout the text.

3 The necessity for this step within Siegel's framework will become apparent below.


5 The "Uncola" is the exception which proves the rule.

6 Such cases as this and 5c are somewhat problematic, since when read or seen are modified in certain ways they can appear prenominally.

i. A well read book lay on the shelf.

ii. That is a seldom seen occurrence.

Note, however, that the unparticiple are bad in such environments.

iii. *That is a seldom (often) unseen occurrence.

I'm not sure what the principle involved here is. For a discussion of some cases which seem related see Bever and Langendoen (1973).

7 Siegel points out that for many people unpassives with by-phrases are unacceptable. For people who allow by-phrases acceptability varies according to the specific content of the by-phrases as well as the specific unparticiple involved.

8 Siegel assumes the following S.D. for the passive.

X NP Aux V (p) NP Y by+PASS

9 This example seems problematic because there are adjectives which take PP with to, e.g. he was good to me. The fact that (12b) is ungrammatical probably has to do with the stativity of unpassives, which I will discuss later in the text. Note that the stative *he was green to me or *he was alive to me are unacceptable in the sense where the PP is a complement of the adj. and not the VP.

10 Siegel wishes to account for the facts exemplified by (21)-(24) (as well as others) by selectional restrictions on her lexical item un-.

11 For arguments that this assumption is not quite correct see Newmeyer (1970) and Jackendoff (1972).

12 For a discussion of sub-structures of complex lexical items and idioms see Katz (1973).

13 This claim has not yet been tested against a very wide range of data and hence the possibility exists that it cannot be maintained. For the types of examples I have considered in terms of this analysis the constituent structure proposed, taken in conjunction with Emonds (1972) analysis of particle shift, does not seem unreasonable. However, it is unclear whether it can be refined to accommodate examples like those given in Chomsky (1973). However, even if the claim cannot be maintained and the S.D. of the passive transformation must be (35), I believe that the correspondences Siegel points out can be accounted for along lines of argumentation similar to, but probably more sophisticated than, those presented below.

14 Note that there is a difference in the meaning of ((slip into x)) and (slip into x) which is just what we would expect if separate lexical items are involved.

REFERENCES


TRANSFORMATIONAL GRAMMAR AND JUNCTION GRAMMAR:

CONTRASTING METAPHORS

Melvin J. Luthy

Languages and Linguistics Symposium

April 7-8, 1975

Brigham Young University
Metaphorical communication is so much a part of man that he sometimes fails to separate the metaphor from the concept to which it refers. Consequently, in language study, some persons find it difficult to separate linguistic models from the language systems they represent. Like metaphors, linguistic models draw attention to certain features which are felt to be important by the person using them—but they neglect others. Each different grammar or metaphor, then, may call attention to different aspects of language.

It is the ideal goal of the linguist to someday discover a model that will do justice to every aspect of language. Today there is widespread agreement that the first necessary quality of such a grammar is that it be generative; that is, it must model man's creative and intuitive linguistic processes relating form with meaning. To the degree that linguists are able to make explicit approximations of these processes, others, including psychologists and teachers, become interested in using the linguist's description as a basis for their studies of cognition, perception, and language acquisition. It is clear that their studies can be no more valid or insightful than the metaphors upon which they are based. Thus, our modern grammars are a mixed blessing—both aiding our thinking and limiting it, both revealing vital information and obscuring it.

Transformational Metaphors in Stylistic Analysis

Among those hoping to benefit from the use of modern grammars are students of literature, and one of their primary expectations is for linguistics to aid them in stylistic analysis. Such a goal should prompt us to consider the kind of metaphorical "looking glass" that our grammars provide for such study. By looking at language through the metaphor called transformational grammar, we first see categories and lexical items generated in constituent hierarchies, modeling, in a semantically interpretable way, the meaning of the sentences, or the conceptual structure in the mind of an ideal speaker-hearer. This we call deep structure. Next, we envision deep structure strings being manipulated in various ways, added to, deleted from, and finally given acceptable linear order and phonological value. This we call surface structure. Continuing the metaphor, we often consider paraphrase to be the case in which the meaning of one deep structure has been expressed with differing surface structures; and we view ambiguity as the case in which different deep structures are retrievable from the same surface structure.

This concept of deep structure, surface structure, and paraphrase has allowed students of literature to look at stylistic analysis as the study of the recurrent or habitual use of transformational options, and the preferred choice of paraphrase. This approach, like the metaphor behind it, rests in part on the assumption that valid paraphrases do exist. Attacks on the validity of this assumption do not dissuade researchers from relying on it. After all, paraphrase, like sentence, may elude our rigorous definition, but it remains a useful concept. Notice, however, that its definition rests on the transformational model positing deep and surface structures.

Junction Grammar Metaphors in Stylistic Analysis

Understanding that metaphors have limitations, we may profit from asking how our view of paraphrase may differ if we viewed language with another linguistic description. To do so we may turn to another model called Junction Grammar, a generative model briefly outlined in Eldon Lytle's A Grammar of English Subordinate Structures. In this model, the generative process consists of three primary operations: adjunction, or the type of operation which generates a sentence from a subject and a predicate, or a predicate from a verb and its object; conjunction, or the type of operation which combines like categories without changing the category of their common dominating mode; and subjunction, or the type of operation which subordinates modifiers and recategories constructions, as in the nominalization of sentences, predicates and verbs, or the adjectivalization of verbs, etc.

The metaphor of generation in this model does not proceed from a single node branching into constituent structure to form strings. Rather, it begins at the bottom of the inverted metaphorical tree where grammatical categories (and sememes) are related to each other by the operations mentioned above, such that higher dominating nodes are generated, which in turn are related to each other by one or more of the operations available until a single dominating node is generated. Thus, the following subset of rules (selected from a larger finite set of "junction rules") generates the sentence, "George and Alice help students who need counsel." (Categories and sememes chosen to be related by generative operations: N = George, Alice, students, counsel; V = help, need.)

Rules
1. V + N = PV
2. N + PV = SV
3. N + N/SV = N
4. N & N = N

Rule #1 adjoins (adjunction) a verb with a noun, resulting in a category PV (predicate).

Figure A.

```
  +
V  N
```

- need counsel

Rule #2 adjoins a noun with a predicate, resulting in the generation of the category SV (sentence).

Figure B.

```
  +
N  PV
```

- students

- need counsel
Rule #3 subjoins (subjunction) two identical N's such that the second N is the N dominated by an SV, and the resulting category of the subjunction is N. This type of subjunction can also be called interjunction, since it consists of the intersecting of two sets at the point of their common member. The point of intersection in this case is marked with the relative pronoun who.

Figure C.  
\[ \text{SV} \quad \text{N} \quad \text{students} \quad \text{who} \quad \text{N} \quad \text{need} \quad \text{counsel} \]

Rule #1 can reapply to generate a new PV category from the adjunction of a V and the new category N.

Figure D.  
\[ \text{PV} \quad \text{V} \quad \text{heap} \quad \text{students} \quad \text{who} \quad \text{V} \quad \text{N} \quad \text{need} \quad \text{counsel} \]

Rule #4 conjoins two N's (neither shown in the tree at this point), and generates an N dominating them both.

Figure E.  
\[ \text{N} \quad \text{N} \quad \text{N} \quad \text{George} \quad \text{Alice} \]

Rule #2 can reapply to adjoin the dominating N generated by Rule #4 with the dominating PV generated by the last application of rule #1. The result is a well-formed sentence with a single dominating node, SV:

Figure F.  
\[ \text{SV} \quad \text{N} \quad \text{George} \quad \text{Alice} \quad \text{help} \quad \text{students} \quad \text{who} \quad \text{V} \quad \text{N} \quad \text{need} \quad \text{counsel} \]

Indeed, many kinds of recategorizations without first generating separate sentences underlying each, as would be the case in standard transformational grammar. Also, the relative pronoun in this model is seen as the marker of the point of intersection of two sentences, rather than the result of a transformational rule changing the second identical NP into a wh-PRO form.

The concept of interjunction can be further illustrated by adding a fifth rule interjoining object N's.

5. \[ N + N/PV = N \text{ (N subjoined to the identical N of a PV, with the resulting dominating category of N)} \]

Figure G.  
\[ \text{SV} \quad \text{N} \quad \text{manager} \quad \text{hired} \quad \text{(the)N} \quad \text{girl} \quad \text{whom} \quad \text{loved} \]

The recursive applications of rule #3 will also generate a relative clause subordinate to the subject N.

Figure H.  
\[ \text{SV} \quad \text{N} \quad \text{manager} \quad \text{who} \quad \text{hired} \quad \text{(the)N} \quad \text{girl} \quad \text{whom} \quad \text{loved} \]

In each case the relative pronoun marks the intersection of two sentences. Word order is not specified at this point in the model since the relationships shown in the tree are considered universal, while word order is language specific. (This nonordered representation is another contrast with transformational grammar.)

In transformational grammar the pronoun is the result of a second occurrence (under given conditions) of an identical (co-referential) NP which assumes the wh-PRO feature and then is moved to the beginning of the sentence of which it is a part.

We may turn now to the active-passive question which, in standard transformational grammar, is often used to illustrate the concept of paraphrase. Those using transformational grammar for stylistic analysis will posit one deep structure for the basic thought, and then relate that deep structure via transformational rules to different surface structures, or paraphrases. (See Figure I. on following page)

In junction grammar, transformational rules are not used in sentence generation. Consequently, each active and passive sentence will have its own phrase marker. (See Figure J. on following page.)
The passive metaphorical tree suggests that the dominating SV in Figure K is not transitive since the object category is empty (E). Yet, there is a transitive relationship in the predicate publish book, which has been recategorized as ADJ through the operation of subjunction, i.e., published book. The highest ADJ category then functions like any other adjective following the be verb, and in this case is modified by agentive prepositional phrase subordinated by the same type of rule that subordinated relative clauses. Since the subject of the sentence is identical with the object of the adjectivalized predicate, the object N, although intuitively felt, is not spoken in English sentences.

If we view stylistic analysis through this metaphor we begin to concentrate on the difference between active and passive rather than on their sameness. Our attention is drawn to other generative operations: we notice the writer's recategorization of PV. Instead of leaving it as a PV as in Figure J, he prefers to recategorize it as an ADJ, and then modifies the ADJ with the agentive prepositional phrase. Consequently, in this model, the sameness in syntactic deep structure cannot be part of our definition of paraphrase since active and passive paraphrases do not share the same deep syntax. Paraphrase, in junction grammar, therefore, must be defined in yet-to-be-understood semantic terms. This is a major difference with far-reaching implications for those wishing to use a linguistic model for a stylistic analysis.

Another significant area of difference in these metaphors is the manner of modeling modification. As one studies the available transformational texts, he is struck by the lack of information regarding adverbial modification, as if it were not of concern in a model of linguistic competence. Of course it is of concern, but the model does not lend itself well to clear explication of it. Adjectival modification is implicitly shown with the embedding of S's and NP's. Sentence modifiers are shown as parts of Pre-Sentence constituents, but other adverbial modifiers are not clearly distinguished from other instances of concatenation:

The junction grammar model makes a general claim that each instance of modification involves the operation of subjunction. Consequently, each instance of modification, adjectival or adverbial, is shown in essentially the same way:
In each case, a modifying structure has been subjoined at a specific level of modification in the phrase marker: the sentence modifier has been subjoined at the sentence level; the relative clause has been subjoined at the noun level; and the adverb has been subjoined at the predicate level, since it modifies the entire predicate.

This method of modeling modification allows for the explication of nested modifiers in a way unavailable in standard transformational grammar. Consider the further modification of the adverb well by the word very. The phrase structure rules in a standard transformational grammar would consider this a concatenated adverb sequence, obscuring the intuitive knowledge we have of nested modification. The junction grammar model would simply subjoin very to well.

Whether or not we are consciously aware of level of modification is often a function of the metaphors we use. From my own experience, I find that students completing their first semester of transformational grammar are seldom aware of the problems of modification, since the model they study does not clearly focus on them. More advanced students trying out their newly acquired transformational wings begin to work with structures that so bewilder them. For example, they learn that in some way a distinction must be made between the sentences, "The fact that was obvious bothered him." and "The fact that the fact was obvious bothered him." They learn that perhaps the most satisfying way of handling the problem is to show that the subject noun phrases in such sentences are different. The relative clause is embedded with the rule $NP \rightarrow NP + S$; whereas, the complement is embedded with the rule $NP \rightarrow (ART) N + S$. This makes it possible for an NP in the relative $S$ to be identical with the preceding NP. This condition must be met for a relative embedding. Since no NP exists to the left of $S$ in the complement embedding, this cannot be construed as a relative embedding.

The rub comes when a student wants to further modify the noun fact in the complement embedding by saying, "The astonishing fact that the fact was obvious..." At this point he returns to the standard transformational claim that adjectives are derived from deep structure relative clause embeddings. Thus, the astonishing fact must have the following structure:

---

![Diagram](image_url)

---

Placing this modification in the tree, and hoping to follow it with a complement embedding is perplexing, indeed, since the sequence $N + S$, necessary for the $S$ to be interpreted as complement embedding, is not present when the noun is modified:

---

![Diagram](image_url)

---

The method of showing modification in the junction grammar model is free from these limitations: First, the "that" complement is shown as a full subjunction in which the entire complement sentence intersects with the preceding noun, rather than with only the subject or object noun, as is the case with a relative clause:

---

![Diagram](image_url)

---
Relative (interjunction)

This leaves the antecedent noun unencumbered to be modified in any appropriate way:

Figure R.

For junction grammar as a metaphor, this is clearly an advantage; it draws our attention to important concepts of modification that we have neglected because our transformational metaphor does not illuminate them.

On the other hand, as a metaphor, the transformational model has many other advantages. For example, it illustrates the regular syntactic dependencies involved with English auxiliaries, as it metaphorically draws our attention to a systematic movement of affixes. Since the junction grammar metaphor is not transformational it does not focus our attention on auxiliaries in this way.

What we see, then, when we briefly contrast the operations of two generative grammars is shifting metaphors, each calling attention to specific properties and processes: The transformational model with a syntactic base generating deep structure sentences later modified with transformational rules in various ways to describe sentence relatedness, ambiguity, synonymy, and to suggest a definition of paraphrase; junction grammar with a syntacto-semantic base generating through adjunction, subjunction and conjunction deep structure categories leading to sentence descriptions in which modification is clearly shown.

In addition to the obvious need to apply evaluation criteria to both models, there is a subtle, yet equally important need to recognize that whatever model is used by the researcher, his insights will be aided and limited, illuminated and obscured by his chosen metaphor. Without that constant awareness he may one day find himself in his ivory tower—with the door locked.
A LIST OF VERBS MOST NEEDED
FOR COMMUNICATION IN A FOREIGN LANGUAGE CONTEXT

Thomas F. Rogers

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
A LIST OF VERBS MOST NEEDED FOR COMMUNICATION IN A FOREIGN LANGUAGE CONTEXT

Thomas F. Rogers

While in Yugoslavia recently I kept a list of those verbs most frequently required in particular verbal transactions or communication contexts. My needs and activities were those, I believe, of the average foreign tourist.

I was pleased to learn that, over an entire summer, I could express most every personal need by learning some ninety verbs which appeared to fit eight or nine distinct semantic categories. In the process, moreover, I was surprised to discover that a number of verbs commonly encountered in most every language grammar--i.e. "to sit," "to dress"--while describing actions common to us all, were not normally called for in rudimentary conversation (and, when so, mostly in the imperative).

It is my contention that my ninety "indispensable" verbs would, with but slight variation, prove essential for most any foreign traveler in any language culture. As verbs are the "heart" of expression in any language, moreover (and often the most difficult part of speech)--I further contend that, by mastering the forms of those represented by my list, one could at least "make himself understood" in any particular language culture after only weeks or even days of grammar and lexical study. This would also enable the traveler to move within brief periods from one unfamiliar language culture to the next and learn enough of each particular language code to "communicate" to the natives on a fundamental level.

My list is admittedly subjective and pragmatically arrived at, and--as far as a paper is concerned--I have relatively little more to say than this prospectus already explains. However, I would be pleased to provide copies of that list in the form of a handout and to entertain the response of those who would care to look at and discuss it with me.

BASIC LIST OF VERBAL SEMEMES
(FOR ALL LANGUAGES)

(1) MODALS--(a) EXISTENTIAL,
(b) ATTITUDINAL and EMOTIONAL STATES

- have to/must
- hope
- fear
- try
- remain
- begin
- end
- like
- love
- need
- want to
- can
- repeat
- continue

(2) PERCEPTION & COGNITION

- listen
- hear
- think
- know
- believe
- understand
- remember
- see
- look at
- forget
- notice
- find out/learn
- agree
- read

(3) DISCOURSE

- ask (question)
- ask (favor)
- say/tell
- speak
- answer
- explain
- announce
- promise
- consult with
- (agree)

(4) SOCIAL AMENITIES

- invite
- suggest
- introduce
- congratulate
- thanks
- excuse

(5) CIRCUMSTANTIAL POSITION & STATUS

- work
- make
- do
- be located
- learn/study
- live
- stay
- meet
- give
- use

(6) LOCOMOTION

- arrive
- depart
- go
- travel
- bring
- enter
- exit
- walk
- ride
- fly
- swim
- sail
- run

(7) CREATURE COMFORTS

- sleep
- eat
- breakfast
- sup
- bathe
- wash
- dine

(8) TRANSACTIONS

- get (obtain)
- pay
- find
- shop for
- look for
- send
- buy
- receive
- sell
- accept
- order
- exchange
INNOVATIVE TECHNIQUES OF RUSSIAN INSTRUCTION
IN THE UNITED STATES 1970-74

Donald K. Jarvis

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
INNOVATIVE TECHNIQUES OF RUSSIAN INSTRUCTION IN THE UNITED STATES 1970-74
Donald K. Jarvis

Introduction

This paper is an excerpt from a longer study commissioned by the Center for Applied Linguistics, Arlington, Virginia. The study, entitled "The State of the Art of Russian Instruction in the U.S.A.: First and Second Year College Level," was in turn the third section of several commissioned recently by the Center to survey "The State of the Art of Russian Instruction in the U.S.A." An abstract of this Section III is included by way of introduction to the present paper.

This paper surveys the state of the first two years of college level Russian instruction during the early 1970's. Data was gathered via a survey of recent literature and a questionnaire returned by 52% of all U.S. programs.

Most U.S. students of Russian begin and end their study in the first two years of college work for a total contact time of ca. 350 hours. By contrast, Soviet students may have over 600 hours contact time with a foreign language when they begin college-level work. Despite the importance of efficiently using this short time, the profession has paid scant attention to this level judging by the criteria of dissertations, articles, and professional preparation. However, recent signs indicate increased interest.

During the first year of instruction, more than a third of the programs use an eclectic approach, one third report an audio-lingual approach, and less than a third describe their technique as the more traditional grammar-translational approach. Texts used reflect this preference. Second year instructors favor grammar-translational texts supplemented by readers.

Innovations include 1) computer-based and computer-assisted instruction, 2) speech delay, 3) individualized approached, 4) decoding courses, and 5) the Lipson technique. These seem at first glance to have little in common, but closer inspection reveals some common factors: 1) all develop a high level of student involvement; 2) all but speech delay encourage student initiative and expression of individuality, long recognized in psychological literature as motivating factors; 3) the Lipson technique, speech delay, and decoding courses emphasize focus on the content as well as the form of the language, an aspect recently advocated by Birkenan; 4) all but the Lipson technique involve limitations in oral work.

CAI-CBI

Computer-assisted instruction (CAI), in which the computer serves as an adjunct to regular class work, has recently been reviewed by George Kalbouss and Edward Purcell.2 Negative aspects of CAI include 1) the fact that expenses for nine months could run over $5,000 for computer time and one terminal, 2) non-marked, non-standard Cyrillic keyboards and 3) the present limitation to practice of written skills only. Nevertheless, Kalbouss notes CAI's value in vocabulary and grammar drills: the machine can explain a student's errors to him while he is doing his exercises and can tailor exercises to his needs. Kalbouss also notes the computer's value in materials preparation: the machine can easily help the instructor control vocabulary use in exercises and can record and analyze student errors for both student and instructor's benefit. CAI was reported by two programs responding to our survey: University of Southern California and Gallaudet College (for the deaf) in Washington, D.C.

Computer-based instruction (CBI) denotes reliance on the computer to do the bulk of the instruction. Sophisticated, successful programs teaching translation by computer at the University of Illinois have been described by Constance Curtin et al., as has a branching CBI program at Stanford.

Speech Delay

Speech delay is a technique in which beginning classwork omits speaking practice in favor of meaningful (content-centered) listening and writing practice. Visual aids and active student response are also key elements. Homework may include reading and writing practice at early stages. The value of delays improves listening comprehension was convincingly presented over a decade ago by James Asher, but only recently has the principle been applied to another technique than Asher's Total Physical Response Method. Valerian Postovsky reports impressive empirical evidence of the method's superiority in developing a base for all four skills: speaking skills were surprisingly enhanced in Defense Language Institute subjects who began with a six weeks speech delay. Donald Dragt has reported favorably on its use at Michigan State, and Fyank Ingram has reviewed literature on the area. Daugherty's transformational-based "structure" course at the University of Colorado also emphasizes active oral skills, but does not emphasize listening comprehension as much as the above program does. In addition, the University of Rochester and Idaho State University report courses stressing passive skills and thus may be considered variants of the speech delay movement.
Individualized Instruction

Despite massive interest in individualization in other languages, Russian instructors have been extremely hesitant to experiment with it. Based on the premise that speed is a crucial variable in instructional success, individualized programs allow each student to proceed at his own rate through a given set of instructional materials. Teachers are viewed as course organizers, reference individuals, and evaluators. Little convincing empirical evidence has been presented to indicate self-pacing techniques' superiority in foreign language instruction, especially at the college level, where students have some control over total course load, and are usually tolerant of, if not dependent on the instructor's pressure to maintain speed through course material. Patterson reports tentative but generally favorable results with an individualized college Russian program at the University of California at Davis. Staffing and materials preparation were a problem, student progress through material was less than normal, but motivation was increased and attrition rates reduced.  

James Connell describes a less structured but apparently highly successful individualization scheme for a small college program. M. Keith Meyers gives a rather full description of an individualized Russian program at Earlham College (Richmond, Indiana), but reports no conclusions from it.  

In an empirical study conducted at Purdue University, William Buffington found significant advantages for the self-paced mode in advanced students' learning and recall of Russian vocabulary, and he convincingly argues for further study of self-pacing and self-evaluation.

Decoding Courses

Beginning courses which focus on the single skill of reading technical Russian--frequently called "decoding" courses, since merely understanding a written text is usually the goal of such courses--are hardly new, but recently have been the object of renewed interest as part of a general trend to limit class focus for achievement of a useful skill.

The CBI course at Illinois is a decoding course and has been described above. A more conventional program at Brigham Young University has recently been described and apparently shares some common points with a Pennsylvania State course described over a decade ago: drastic limitations on required vocabulary together with increased emphasis on word derivation, deducing dictionary forms, and understanding participles and other deverbative forms. Students are allowed to begin reading in their own fields as soon as possible with individualized aid from the instructor.

The Lipson Technique

Six of our respondents listed their methods as "Lipson technique." This method is characterized by: 1) class dynamics emphasizing role-playing, whimsy, creativity, and humorous recombination of carefully controlled elements; 2) visual symbols to avoid translation; 3) a strong oral emphasis as a basis for later reading; 4) inductive presentation of grammar rules which are then explicitly set forth; 5) a careful presentation of the single-stem verb system. A published version of the Lipson technique is now available.

FOOTNOTES

11. Ibid.
15. Sanford Couch, "Return to the Language Lab," Russian Language Journal, XXVII (Spring 1973), 40-44.


18. Probably the most widely used of these supplements is Robert L. Baker's Workbook for Russian for Everybody, Part I: Lessons 1-21, and Workbook for Russian for Everybody, Part II: Lessons 20-40. Preliminary editions are available, at this writing, through BYU Press, Marketing Division, Brigham Young University, Provo, Utah 84602.


37. David A. Hanson. Personal letter to the author from this Brandeis University professor who has used the Lipson technique for eight years.


43. Aronson, "Why Aren't We Fluent?" 437-47.

LITERARY COMPUTING: SOME METHODOLOGICAL PITFALLS

Steven Sondrup

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
LITERARY COMPUTING: SOME METHODOLOGICAL PITFALLS

Steven Sondrup

Since literary critics and historians have long fancied that they dealt exclusively with matters of the spirit and of human intuition, there was considerable resistance on the part of many when computers were first introduced as a tool for literary analysis. Some regarded the computer as an unwelcome interloper that had escaped from the technological, if not the military-industrial complex. The simple fact of the matter, though, is that quantitative judgments, which computers facilitate, have long been part and parcel of literary criticism and analysis. In historical but more particularly in stylistic studies, quantitative judgments have been used to explain many aspects of style and stylistic change. Observations like "The mood of this novel is created by the preponderance of adjective." or "The dynamism of this poem results from the concentration of verbs." are frequent and typical. For such evaluations to be accurate and statistically significant, the relevant stylistic element must be counted and compared to some kind of standard or norm. Typically, though, the judgment is subjective and approximate. Although the intuition of the researcher or critic may in the end be absolutely correct, we deserve and should expect more specific and concrete evidence: we should expect that such subjective observations will be supported by facts.

In the past the expense and tedium of actually counting any stylistic feature has precluded furnishing this kind of specific data, and the public has generally been rather indulgent and understanding. Computers, however, can quite literally in seconds make these counts and comparisons with unrivaled accuracy. It should be noted too that this is essentially all that computers can do: they cannot interpret the data, they cannot explain the significance of the data. In short they cannot replace the mind and soul of the sensitive critic. Computers can only make the perceptive and understanding critic's work easier, more accurate, and more penetrating.

Although there are many things that computers cannot do, what they can do, they do very well. The great precision and exactitude that computers bring to literary analysis, while a welcome antidote against unsupported approximations and intuitive guessing, are awesome. Many traditional stylistic and grammatical categories and definitions are too inexact or ambiguous to be used in conjunction with the great accuracy afforded by computer-aided tabulation. The critic, is, therefore, necessarily faced with the need to rethink and reconsider some of the most basic categories and concepts of his discipline, lest the precision of the computing techniques be diluted by the ambiguity of poorly defined or ill-considered categories.

One of the first steps in any computer-aided study is the generation of a word-frequency chart. The procedure is simple from a computing point of view; it involves counting how many times each word in the corpus is used and then computing what percentage of the total each word represents. Although this is a simple, straightforward procedure from a technological standpoint, it requires the literary critic to ask himself some very probing questions. What is, for example, meant by the term "word" in this context and how does one "word" differ from every other "word"? Although Chomsky and Halle provided a penetrating and precise definition within the framework of transformational grammar, the literary critic must nonetheless consider how, within the context of his particular study, this crucial term is to be defined. If an unedited text is submitted to the computer, and the computer is programmed to count how many times each word in the corpus appears, the results supplied will be based solely on orthography. All items that have exactly the same spelling will be grouped together and counted together, and conversely variant spellings, inflected forms, and abbreviated forms of what is usually considered the same lexical item will be grouped separately and counted separately. For a study concerned with a particular author's orthographic habits, such a count would be useful, but for studies concerned with more subtle elements of style, such statistics would be only marginally applicable, if not completely misleading. The first step in going beneath the orthographic surface should involve coding the most obvious homographs, so that the computer will be able to distinguish between them, list them separately, and count them separately. Thus book in the sense of a printed volume should manually be distinguished from book in the sense of making a reservation. Or in German der Arm--the arm--should be distinguished from the adjective arm--poor--in some way.

Closely related to this segregation of the most obvious homographs, but perhaps slightly more subtle, is the question of words that in common parlance seem to have a wide range of overlapping meanings but syntactically and perhaps grammatically behave very differently. In many cases it is a matter of two very different deep structures emerging in the same surface structure. Consider for example the German word auf meaning roughly "on". It can function as an ordinary transitive preposition--that is a preposition that takes an object--as in the sentence: Der Hund springt auf den Stuhl. (The dog jumps onto the chair.) The same lexical item--the word auf--can also function as a verbal participle as in the sentence: Die Musik hört sofort auf. (The music stopped immediately.) Now the inclination of some might be simply to regard the word auf as a single item that has two or more different meanings, and in some situations such a procedure might be perfectly satisfactory. But the matter is more complex and it appears on the surface. The basic question is not one of conceptual meaning in the ordinary sense of the word, but rather a matter of syntactical behavior of the

two items. If the computer-assisted analysis of any literary text is going to be useful in making judgments about the grammatical patterns or syntax of a particular poet, such syntactic subtleties must be reflected in the preparation of the text. In this case auf as a preposition and auf as a verbal participle would necessarily have to be distinguished in such a way that the computer would be able to tabulate them separately.

Perhaps an even more telling and extreme example can be found in the English word have. It is such a common word that the complexities beneath the surface are often overlooked. Have can be used as the main verb of a sentence: I have a house in the forest. In such sentences have is the only verb in the sentence and takes an object as normal transitive verbs do. Have though can also be used as an auxiliary verb to form the past tenses of other verbs: They have already come. Although the transitive form of have and the auxiliary form of have in many respects seem to be the same word, syntactically and grammatically they are two very different words: their spelling and some aspects of the functioning appear to be similar, but they derive from very different deep structures and just happen to look alike at first glance. The very different nature of these words is clearly reflected in other languages. Italian, for example, in some stylistic respects illustrates aspects of this difference, but in Spanish the distinction is clear and obligatory. Have as a main verb requires the use of the tener: Tengo dos cabillos. (I have two horses.) Tiene una casa en Mexico. (He has a house in Mexico.) As an auxiliary verb Spanish requires the use of haber: Ha hablado. (He has spoken.) or Ya ha comido. (He has already eaten.) Although the difference between the two words is not as apparent in English, it is nonetheless as real and as critical. In English moreover it is not only a matter of accurately representing the syntactic structure of any particular literary text: there are a number of phonetic implications that play a role in any metrical analysis of the text in question. Have and its conjugated forms --has and had-- when used as main verbs in clause and stressed, but as auxiliary verbs, they are not stressed. Computers have recently proven extremely useful in scanning large quantities of poetry and suggesting the favored metrical patterns of different poets, but in order to accomplish this task effectively, the distinction between the stressed and unstressed forms of have necessarily must be indicated.

Although in English there are relatively few other words that function in the same way, in other languages the verb to be can function both as a main verb and as an auxiliary verb. Consider the French Il est ici. (He is here) and Il est venu aujourd'hui. (He came today.) A French text would thus have to be examined for occurrences of etre both as a main verb and as an auxiliary. The matter is even more complicated in the case of German: in addition to the corresponding forms of have and be--haben and sein--werden, the verb corresponding roughly to become, can be both an auxiliary and a main verb. Consider, for example, Er wird bald König. (He soon becomes king,) and Er wird bald hier sein. (He will soon be here.)

Just as the computer-assisted tabulating techniques require precision in distinguishing one word or lexical item from another, great care must be exercised in classifying words. It would be of considerable interest to know whether a given poet, for example, used more adjectives in his verse than is typical in prose. It would also be very enlightening to know whether nouns rhymed with nouns or with verbs or with adjectives that modify them. It would be useful to know what class of word most frequently appeared in stressed positions and what class appeared most frequently in unstressed positions. But before this kind of information can be provided a working definition of the various parts of speech must be established. Too many studies have been based on the traditional largely Latin-based eight parts of speech. Though these eight parts of speech may be a useful point of departure, they are generally too poorly defined to provide significant information. In endeavoring to overcome this difficulty, some researchers have carefully defined the parts of speech in terms of the structure of the relevant language, in one case producing twenty-four different parts of speech. In another case the emphasis was placed on simplicity: only four parts of speech plus a large class of undefined or unspecified words were established. Both the approach of defining categories with great precision and thus multiplying their number as well as that of simplifying and thus working with a relatively small number of classes have advantages and disadvantages. The crucial factor is simply establishing a system that will yield the kind of information sought, whether it be very general or highly specific.

Although it would be very difficult if not entirely impossible to establish criteria that would apply in all cases, general guidelines can be suggested. Perhaps the first important consideration to be borne in mind is that categories should be established that are useful in analyzing the language in question. Although this may seem obvious, critics with a distinctly traditional, literary rather than linguistic background have in the past applied categories that make eminently good sense in Latin to language where they do not fit well at all. If, for example, the language in question does not distinguish between adjectives and adverbs, it makes little sense to set up the categories; a general category modifier would probably make more sense. Secondly, the categories should be defined specifically in terms of the kind of data that is being sought. If only questions of a very general nature are being asked, then quite obviously general categories will suffice. If, however, more specific questions are to be asked, more precise categories would be required. Ideally, though, more than surface structure should be reflected in both cases. Nominalized adjectives, for example, in many respects are nouns, but in terms of the deep structure they modify a substantive
that does not emerge in the surface structure. How then should such adjectives be counted? Ultimately the answer will depend on the kind of information being sought, but the linguistic facts of the matter should at least be kept in mind.

While the general goal of computer-aided literary analysis is to bring a degree of precision and exactness to a field where these qualities have been notably lacking for many years, it must be remembered that language and especially literary language is not mathematically exact. Ironically therefore the final suggestion for setting up a framework within which the usefulness of the computer can be optimized will in many respects necessarily relativize and perhaps temper the foregoing pleas for exactitude. Since language and literature are full of ambiguities—matters that ultimately must be left open to individual judgment and interpretation—any critical framework that is established to analyze literary language must leave room for these ambiguities. There is much poetry that is based on the tentative and at times inexact nature of language. To resolve ambiguities in one direction or another, even if this is done consistently and with great care, introduces a most unwelcome element of arbitrary judgment and destroys some of the fundamental meaning of the passage, especially if the poet's intent is particularly ambiguous. Examples of poetic ambiguity are legion, but by way of example consider the opening lines of Gerard Manley Hopkins' "Spring":

The glassy pear tree leaves and blooms,
they brush
The descending blue...

What are the words leaves and blooms? They can be regarded as verbs whose subject is pear tree, but they are also the antecedents of they and therefore necessarily nouns. Hopkins quite intentionally introduces this element of ambiguity, and in resolving it one way or the other, an important element of poetry would be destroyed.

Thus in an ironic way, the precision that computer-aided techniques provide leads to an awareness of the ambiguous and approximate nature of poetic diction. In spite of the irony involved, it is certainly preferable to pursue new standards of accuracy and precision with an awareness of the irresolvable factors that will be encountered than to work in the dark, unaware of this critical aspect of the nature of language and especially poetic diction.
CHAUCER'S "UNDERWEIGHTED" LINE:
SOME COMPUTER-ASSISTED STEPS TOWARD A REASSESSMENT

Karen Lynn

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
Perhaps it's true that we don't miss great authors we never had, just as we don't miss friends who've never been born. But it's difficult to imagine Middle English without Chaucer. He seems to fill an important emotional niche for medievalists; I think our personal stability requires him. When a medievalist is discussing Chaucer, he knows that it's one of the times he won't have to be on the defensive quite so often; he won't quite so often have to make a dash for that escape tunnel he by now knows so well: that is, whenever the middle ages has left us something unappealing and inexplicable that we must read and deal with anyway just because it has survived, the medievalist accounts for this unattractive relic by muttering something about "the tastes of another day" or the necessity for "approaching the work through conventions by now completely foreign." Valid though the point may be, the suggestion that "well, they might really have liked this in the Middle Ages" always has a ring of condescension to it. That's why Chaucer is so indispensable; he's the proof, Exhibit A whether or not there's an Exhibit B, that people six centuries ago could speak for themselves, and could speak to us.

It's important to begin with that this statement of agreed-upon medievalist preferences because this paper has as its point of departure an assumption that's admittedly an intuitive one. Though the Middle English period, like every other literary period, is full of un distinguishable verse by a host of minor poets, only Chaucer sounds like Chaucer; just as no one but Milton can sound consistently Miltonic, whatever that term means, no other poet can turn out a Chaucerian line, except as an occasional happy accident. We sense the difference in Chaucer; we're convinced of it; we can recognize Chaucer's work immediately; so whatever Chaucerian means, it must mean something. It's here that the student of poetry and the student of linguistics are brought together; as the linguist sorts out and gives names to the elements of language and the processes of linguistic construction, the student of poetry is particularly concerned with the behavior of these elements and processes as the poet employs them.

How can we best approach the question of what makes a Chaucerian line Chaucerian? An adequate analysis of any one poet's handling of any one verse form must fulfill two requirements: it must give a satisfactory description of the underlying verse patterns as the poet conceives it, perhaps comparable to linguistic competence; and it must provide an account of the variations and alternatives within the prosodic craft of the particular author which to him constitute the acceptable actualizations--his possible performances--of this pattern. The alert student of Chaucer would surely see the value of working toward this kind of linguistic analysis; it's not enough for him to read the Prologue to the Canterbury Tales, shake his head in admiring disbelief, and murmur "isn't that wonderful!" or "isn't that Chaucerian?" Specifcally, the intuitive point about Chaucer's verse to which we will attempt to give some kind of non-intuitive substance has to do with one widely-shared, pleasurable, positive reaction to his poetry; it's light; it doesn't weigh too much; it moves better than anyone else's verse that we know of in Middle English. Eleanor Hammond, in her book English Verse Between Chaucer and Surrey, notes that one of the characteristics of Chaucer as a writer of verse is what she calls the "underweighted" flavor of his verse. This "swift fluid narrative," she says, comes from his tendency to include only a small number of stressed, grammatically important words in any one line. She feels that he was greatly aided by the Middle English syllabic final -e, certainly not a sound that is going to slow a line down, even if it is pronounced; thus, she reasons, Chaucer was able to fill the required number of syllable-slots in each line without using as many words as a modern poet would have to; or, to put it another way, the ratio of syllables to words is higher than would be possible in Modern English. A Modern English rendering of Chaucer, if it attempts to keep the same meter, must therefore fill up the slots with something else--more words or longer words--and is thus heavier. And Paull Baum is another critic who speaks of Chaucer's practice of "allowing in the place of the normal five stresses four, three, or even two rhetorical emphases," meaning presumably primary lexical stress.

This all sounds very logical, but it's difficult to believe that the secret of Chaucer's light line rests entirely upon his avoidance of stressed syllables; even with the help of the final -e, that most convenient and unobtrusive of syllables, good poetry is not usually made up of prepositions, conjunctions, and auxiliaries. In 1966 Morris Halle and Samuel Jay Keyser suggested a new system, or at least a new notation, for describing a verse form and for separating metrical from unmetsrical lines according to this description. They described iambic pentameter as a line of 10 positions, the odd-numbered positions weak, receptive to stress, the even-numbered ones strong, receptive to stress. They defined an entity which they named stress maximum: a fully stressed syllable occurring between two unstressed syllables within the same syntactic constituent within a line of verse.4 The Halle-Keyser theory was the basis of a computer-assisted study of samples from five Middle English poets, including Chaucer. This investigation pointed toward somewhat different explanations for Chaucer's light line, at the same time that it pointed up, incidentally, some possible shortcomings of the Halle-Keyser theory itself.

From verse samples that had been marked for binary or contrastive stress, the computer tallied, among many other features, the number of stressed syllables, the number of stress maxima, and the comparison of these counts to the total number of syllables. The first item to be noted is that Eleanor Hammond and Paull Baum have missed their guess: Chaucer's rapid line movement is in fact not the result of fewer stressed syllables, not even by comparison with the lines of John Lydgate, again an intuitive choice but this time as the man who wrote

Karen Lynn

Chaucer's "Underweighted" Line: Some Computer-Assisted Steps Toward a Reassessment

6.2
the heaviest, most non-fluid lines in Middle English. Thirty-two percent of Lydgate's syllables were stressed syllables, whereas in fact thirty-four percent of Chaucer's were. Nor does a comparison of stress maximum percentages yield any clues as to lightness, either in the ratio of stress maxima to stressed syllables, nor in the total stress maximum percentages. These figures were very much the same for Chaucer, Hoccleve, Lydgate and Dunbar.

The source of Chaucer's lightness must lie elsewhere. The data gathered by the computer shows two interesting distinctions which may constitute at least a partial explanation. The first item of interest has to do not with the number of stresses—we have seen that these figures tell us nothing useful. Instead, it's the placement of stresses that seems to be a key difference between Chaucer and poets not noted for their light lines. According to Halle and Keyser, as long as the syllable-count of a line falls within appropriate limits for a particular meter, the only way that the line can be classified as unmetrical is if a stress maximum—again, a stressed syllable flanked immediately on both sides by a syllable of lesser stress—appears in a weak position, an odd-numbered position or syllable-slot. In the case of iambic pentameter. So a poet can write a line with stressed syllables in all five weak positions, if he wants to. The line passes the test of metricality with any number of weak-position stresses, just as long as no weak-position stress qualifies as a stress maximum. And how does a poet neutralize a potential stress maximum in a weak position? In other words, how does he write variant but metrical lines? He surrounds the weak-position stress with adjacent stress or an adjacent syntactic break; the potential stress maximum is thus changed into a simple stress, allowable in any position in the line.

Chaucer certainly used stress syllables in odd-numbered positions; in fact, positions 1, 3, 5, 7, and 9 each contain just over 1% of his total stressed syllables. Long lines neutralizing elements surrounding these odd-position stresses make these lines, though complex, still metrical. But compare this percentage with Hoccleve's distribution of 2 or 3% in each odd-numbered position, and Lydgate's of as many as 5%. The implications are clear. Though Hoccleve and Lydgate wrote metrical lines, and though they could have found a precedent for this weak-position stress in Chaucer's verse, it's the frequency of this feature, not just its existence, that's important. Chaucer, Hoccleve, and Lydgate each used roughly one stressed syllable for every three they wrote, but Hoccleve and Lydgate spread these syllables out into weak positions as well as strong ones, making more frequent and more cumbersome use of this allowable deviation from the basic pattern.

It might be useful to express this same point, in general terms, from another approach. Chaucer's reluctance to use stressed syllables in weak positions meant in addition that most of his strong-position stresses, since they had not stressed neighbors, would be actualized as stress maxima. And in fact, the odds turn out to be almost nine in ten that if Chaucer placed a stressed syllable in a strong position, this syllable would emerge as a stress maximum. With Lydgate, the actualization of stress as stress maximum in these positions drops to only eight in ten. It seems to me that in this evidence may lie, in part, the real nature of Chaucer's much-discussed "lightness." When a stress or stress maximum occurs in a normal place in the line, the expected stress, set up by the metrical pattern, does not combine with the actual lexical stress to throw the word into greater relief; rather, the reader's or hearer's anticipation of the stress lessens the obtrusiveness of the stressed syllable, and the line is therefore lighter. Whether it was conscious or unconscious, Chaucer's realization that the features which neutralize stress maxima are the same features which load weight upon the line was possibly very influential in his artistic handling of the iambic pentameter form.

Halle and Keyser's system for assigning a complexity count to each line is a good index of the features discussed above, since complexity depends upon weak syllables in strong positions and stressed syllables in weak positions. As the complexity count approaches 10 or 12, the line may not easily be apprehended as iambic pentameter, for example, on the first reading. Chaucer's average complexity is very low—only 1.8, in fact, as opposed to the average complexity of 2.6 for Hoccleve and 3 for Lydgate.

The second significant distinguishing feature of Chaucer's verse which might help account for his fluid movement is the number of possible assignments of syllable to position. The iambic pentameter pattern presupposes ten syllable-slots; traditionally, one slot may go unfilled at the beginning of a line, and one or two extra slots may be added to the end. If an eleven-syllable line needs to fit into a ten-syllable pattern and has only one elision, that's simple enough; there's no debate. But if a line needs two elisions, and there are five possible sonorant sequences or optional syllabic -e's, then the question is more complicated, because there are ten possible ways, not all of them metrical of course, of distributing these twelve syllables over the ten slots. When a long line contains a number of unstressed syllables and optional final -e's, and word-initial and word-final consonants like s, m, n, r, and h—all of which Chaucer seemed happy to slide over—there may be a great many possibilities indeed. The first version of the computer program was designed to accommodate as many as thirty possible distributions for any one line, but this limit turned out to be a naïve estimate, to say the least. For one basic fourteen-syllable line from the Wife of Bath's Tale—"Somm seyde honour, somme seyde holyness"—there are 126 possible ways of making it a 9, 10, 11 or 12-syllable line. This isn't just a game; it tells us something important about Chaucer's phonetic choices. His lines were somewhat longer than the next highest sample, but only by a tenth of a syllable, an insignificant difference which might vary with another sample. And yet for each of his lines there are ten or the averages 4.5 possibilities of syllable to position; no other poet had even half as many. Lydgate again is at the other extreme, with only 1.3 possibilities for each line; there are 1.5 for each line of Dunbar, 1.6 for Hoccleve. Consciously or unconsciously, Chaucer wrote highly elidable lines—often more elisions than he needed. These vowels and consonants elide in traditional poetic practice because they are the least obtrusive.
the easiest spoken in combination with other sounds. They move quickly, and may well be an important element in Chaucer's fabled lightness.

In discussion of stress and light-weight lines, it seems more and more that the traditional prosodic distinction of simple contrastive stress is perhaps not sufficient. As one indication, Chaucer's critics praise him for his skill in not weighing down his lines with stresses in every stress position; but when the critics perceive what seems to be exactly the same technique in Hoccleve, it's called "Hoccleve's clumsy forced stress." Are the two manifestations really the same, colored only because we hate to say anything bad about Chaucer? Or is there really a difference, if we were to examine finer distinctions among the stresses, in the syllables that Chaucer uses in these positions? How about compounds and tri-syllables? There may be significant differences in Chaucer's placement of varying levels of stress, even though his percentage of contrastive stress does not, as we have seen, set his verse apart from anyone else's. So it may be well to remark, in conclusion, that Halle and Keyser's failure to treat these stress distinctions is one deficiency in their theory that Paul Kiparsky has cited recently. One advantage of his new system is that it accommodates a four-fold distinction in stress, carrying a more exact and meaningful accounting of lexical stress over into the comparison with traditionally binary metrical patterns. Like Halle and Keyser's, Kiparsky's system is methodical and eminently programmable, alluring, in fact, to the student who wants to know how poets write poetry but who realizes that art is long and life is short.

FOOTNOTES

6 PauU Baum, p. 85, for example.
USAGE PATTERNS OF THOU, THEE, THY AND THINE
AMONG LATTER-DAY SAINTS

Don Norton

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
 USAGE PATTERNS OF THOU, THEE, THY AND THINE AMONG LATTER-DAY SAINTS

Don Norton

Until a couple of years ago, I had always assumed that the counsel in the Church to use the respect pronouns thou, thee, thy and thine applied only to converts, children, and a few inattentive adults. On listening more closely, however, I find cause for general concern.

The illusion that we use these pronouns freely and correctly stems, I think, from the fact that two of these pronouns, thou and thee, are common usage: thy presents few problems; the object case these occurs predictably in a few convenient syntactic formulas. Not many modern ears require thine, even before a word whose initial sound is a vowel.

By contrast, correct use of thou and its appropriate verb form is rare among the majority of Church members, except in a few scriptural formulas, such as "Our Father who art in heaven." Yet thou is difficult to avoid, since the structure that calls for it occurs normally in modern English. Thus it is common for a person to attempt the structure, then find himself in a syntactic corner. The consequence is widespread incorrect use of thou and the archaic verb form, even among educated, experienced church members. For example, a BYU religion faculty member whom I often hear pray, says, "We pray that thou wilt," rather than "will." In the face of this difficulty, many people I talk to about the use of the respect pronouns wish that the Church would abandon the "quaint" usage, or at least make its use optional, as it is in impromptu ordinances and blessings. Their arguments are simple and sobering: few Latter-day Saints master the usage; it poses special problems to converts; other churches have dropped the usage; and because it is in so little heard and used, it no longer signals the sacred deference it once did.

At the present time, however, Church leaders counsel to the contrary. Their desire is that the Saints continue to use in their public prayers the archaic pronouns of respect.

In an article in the current New Era (April, 1975), Sharon Jones, writing on the usage of thou, thee, thy and thine, cites a policy statement in The New Era of September, 1972: "The Presiding Bishopric expresses concern about some incorrect uses of the language that seem to be gaining widespread momentum among the youth of the Church." After citing a mispronunciation of patriarchal, and the misuse of thy in church talks (... in the name of thy Son...), the note concludes: "It has also been noted that as they pray, many people are incorrectly using the terms you and your instead of thee, thou, and thy." (p. 7)

The most extensive discussion of this problem appeared in the Melchizedek Priesthood Manual of 1972-3, excerpts from President Joseph Fielding Smith's Answers to Gospel Questions.


Q. Is it important that we use the words thy, thine, thee, and thou in addressing Deity; or is it proper when directing our thoughts in prayer to use the more common and modern words, you and yours?

A. Our Eternal Father and his Only Begotten Son, Jesus Christ, should never be approached in prayer in the familiar expressions so commonly used in addressing human beings. The Father and the Son should always be honored in our prayers in the utmost humility and reverence. The changing in the wording of the Bible to meet the popular language of our day, has, in the opinion of the writer and his brethren, been a great loss in the building of faith and spirituality in the minds and hearts of the people.

President Smith attempts in his answer to justify historically the usage of the respect pronouns. Although his conclusions are accurate (that the respect pronouns show sacred deference), his language facts are not true. He implies that disuse of the respect pronouns in prayer stemmed from increased worldly mindedness during the century following the translation of the Bible—a neglect to show deference to deity. In fact, there appears to be little historical relationship between common and sacred uses of these pronouns, until during more recent years, when churches in general have given up the archaic usage. Until then, thou, thee, thy and thine seem always to have been the usage of English writ and worship, irrespective of their other uses. In daily usage, these pronouns at times reflected different degrees of deference, and finally disappearing altogether, except in isolated dialects.

What is the correct usage of thou, thee, thy and thine? I found no handy reference that answered this question. One may exist, but I could not locate it in several hours of searching through language history texts and grammar/usage references. Finally, from several sources, I inferred the following rules, which are surprisingly simple. I will start with the possessive pronoun, then take up the objective case thee, and finally explain thou, the most difficult form, since it requires a special affix on the verb.

Thine is the noun-substitute form, equivalent to your, theirs, mine, etc. One needs it only very infrequently in prayers.

Thy, thine are possessive determiners, thine occurring before words whose initial sounds are vowels—thine eyes, thine ear, etc. Thy quite comfortably replaces thine in modern usage. (Biblical usage in some words is divided: thy/thine hand, heart. The h sometimes was silent, apparently.)
The usage of *Thee*, the object case pronoun, is quite straightforward.

*Thou*, as subject case pronoun, presents few problems (occasionally, a confused person will use substitute *thou*—"We pray that thee will..."), but the verb that follows is unfamiliar usage, occurring only in prayers and in the scriptures. (Unfortunately, the custom of extensive oral reading of the scriptures is not common among us.)

Yet even the verb forms are highly regular, there being but a few, commonly occurring exceptions:

General rule: *thou* + verb base (present or past form) + *-est* or *-est*

<table>
<thead>
<tr>
<th>thou</th>
<th>art, wast, wert (subjunctive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;</td>
<td>hast</td>
</tr>
<tr>
<td>&quot;</td>
<td>wilt</td>
</tr>
<tr>
<td>&quot;</td>
<td>shalt</td>
</tr>
<tr>
<td>&quot;</td>
<td>must</td>
</tr>
</tbody>
</table>

Do takes two forms, the emphatic auxiliary *doest*, and the regular principal verb form, *dost*:

Thou dost sustain us.
Thou doest many great works.

The past form presents a special problem, such as is evident in the verb *built/builled*:

*built*st, *builtest*, *builled*st, *builled*est

The solution is simple: Thou *did*st built.

One's ear is the best guide in the formation of the past forms.

How closely does LDS usage conform to these rules? As I have suggested, not very closely. My students of usage consistently confirm what I myself have observed: Very few Latter-day Saints use *thou* and the required verb forms confidently and in accord with the rules. These are the usage patterns we observe.

1. Correct usage is limited largely to the repetition of a very few scriptural or formulistic phrases. These are most common:

   - Wilt thou... Wilt thou... Wilt thou... Wilt thou... Wilt thou... Wilt thou... Wilt thou...
   - Thou hast. . . Thou hast. . .
   - etc.

2. There exists a strong impulse in most Latter-day Saints to avoid risking the use of *thou*; and consequently most prayers omit that usage, substituting
   a. the imperative:
      (Please) bless us that. . .
      Grant us... Give us...
   b. or other indirect constructions:
      We're thankful for. . .
      We express thanks to thee for. . .

3. The uncertainty of the saints is evident not only in the avoidance of non-formulistic structures, but in the many errors one hears, and in the striking quality of prayers where the *thou* is used freely and correctly.

The errors are predictable, sometimes embarrassingly deviant:

We pray that thou will/would. . .
We ask that thee will... . .
We ask thee to... We ask thee to...
We thank thee for thy many blessings and we pray that you... We ask that ye will...
We ask thee or you... We ask that ye you might...
We thank thee for the gospel that thou gaven us...
. . .whatever thy seest we stand in need of.

The causes of the disuse of the hallowed forms are easy to identify. First of all, the respect pronouns occur only in prayers and in the scriptures, which constitute only a fraction of our linguistic experience. And fluent spoken usage is the product of ear training. Because usage of *thou*, *thee*, *thy* and *thine* is so occasional, we should not expect fluency without special vocal rehearsal as compensation for sparse ear training.

I think the ultimate problem, however, is the loss of the sense of the sacred, such a marked characteristic of life style in the modern secular world. People to whom the sacred is the greater reality find it natural to use special linguistic signals to mark the importance of their communication with that transcendent reality.

This point is the essence of President Smith's argument to continue the usage of the respect pronouns:

As mankind gets farther and farther away from the worship of the True and Living God, the less respect and reverence men will give to God. Therefore, this reverence being weakened or missing, the less inclined are men to look upon the Supreme Being with awe, humility, and reverence. (p. 184)

The solution to teaching the saints to use the hallowed pronoun forms is first to teach them reverence for sacred; then to provide intensive, systematic programs which call for the frequent use of these pronouns
as they would be used in prayers--scripture
reading in the family; a charge to parents to
learn the usage *themselves, and then to guide
their children in proper usage; and finally,
outlines and exercises which will serve to es-
tablish more securely the usage which the author-
ities in the Church exhort us to observe.

Without a deliberate program toward this
end, I predict that fluent and correct usage will
not become general.

The objection to the respect pronouns my
students usually offer is that it is sincerity
that counts, not strict adherence to grammatical
rules. My reply is that the right kind of sin-
cerity will quickly prompt a Latter-day Saint
to learn and use the language which traditionally
has been used to mark that deference to the Lord
which true sincerity implies. I am confident
that the Lord does not reject sincere prayers
that are ungrammatical; but I also know that once
a true saint receives counsel on a point such as
this one, he will desire to follow that counsel,
whatever the cost in effort. I believe that the
respect pronouns, when used properly, genuinely
facilitate worship; for that is the function they
have served for centuries. One of my students, on
hearing an elderly church member offer a fluent
prayer in which thou occurred frequently and
correctly, stated succinctly the crucial point:
"His prayer was vivid and beautiful, very impres-
sive."
LANGUAGES OF CULTURES

V. Lynn Tyler

Languages and Linguistics Symposium
April 7-8, 1976
Brigham Young University
LANGUAGES OF CULTURES
V. Lynn Tyler

How well can you recognize what constitutes cultural "language"? Examples:

A. In "X-l" society, smooth human relationships are emphasized so much, and open disagreement and overt opposition are suppressed so much that there has been developed a way of saying one thing while meaning something quite different. One crucial thing about learning to best communicate in "X-l-ese" is knowing what people mean without saying it in so many words, with or without using certain gestures, or even through configurations of silence. In "Z-2" society, on the other hand, the norm is saying exactly what is meant, without mincing words or using confusing code systems. Remember that in "X-l" it is the other way around. Individuals who in any way speak out violate the social norm and tend to be ostracized. [100.C:6. See Bibliography/References.]

B. The stranger listened carefully, trying to distinguish the whistling sound that came from a long distance. All of a sudden, as if in concert, the people in the nearby group began doing apparently purposeful but, to the observer, somewhat strange things. They then left him, amazed and alone, as they strode off in the direction from which the sounds of whistling had come. [73:15] Language?

C. Syllables needed to translate the Gospel of Mark are, approximately: ENGLISH: 29,000; SLAVIC: 36,500; INDO-IRANIAN: 43,000. [33:314] Language?

D. "No! No! No!" he returned. Each word was punctuated by a stamping foot and a faster rate of speech and figure-eight head-shaking. [78] Language?

E. For at least an hour, the deaf children sat enthralled, enjoying "sounds" of silent signing which wove picturesque thought models in Ameslan. Language?

F. The disparity of meaning in the domain of the term FAMILY is for the tested Korean families only .15 that of the Americans. Students and workers, however, are high with .22 and .25 respectively. [93; pers. notes.] Language?

G. ...some seven hundred thousand distinct elementary gestures can be produced by facial or postural expressions, by movements of the arms, wrists, fingers, etc., and their combinations. Such a blank of silent language structures is drawn upon in the linguistic science of pactorsiology. [73:19] Language?

Possibilities: A. "manner of language" B. whistle language C. syllabic codes D. gestures + negative non-verbal E. affective signing F. statistics and psychocultural concepts G. silent languages of gestures

According to Peter Farb, in his intriguingly readable treatise on what happens when people talk [and write and read -- or otherwise communicate]:

Every speech community has definite ideas about the situations in which various topics can be broached, particular words employed, or even pronounced, and certain tones of voice used.

Every [communication] situation is made up of a series of brief...events which are clearly separated from one another by the employment of different strategies, by the change in social interactions, or by a switch to a different topic of conversation. [33:38]

In reference to the complexity this implies, though from another context, Paul cautioned the Corinthians,

For if the trumpet give an uncertain sound, who shall prepare himself to the battle?

So, likewise ye, except ye utter by the tongue words easy to be understood, how shall it be known what is spoken? for ye shall speak into the air.

There are, it may be, so many kinds of voices in the world, and none of them is without signification. (1 Cor. 14:8-10) Emphasis added.)

Seen, sound, or silent "voices" can be described significantly. There can be better communication with every people. Everyone can experience, study, or in some way profit from the multitudinous forms and functions of what we call, oftimes limedly, "language." [53:80, 62:33]

A versification of this idea appears with this treatise. (See Cultural languages.)

The Brigham Young University Language Research Center joins many people and institutions [12, 30, 49, 52, 60, 62, 86, 88, 92, 95] in an attempt to adequately identify significant language differences that make a real difference, and the unique similarities that can contribute to more effectively understanding and using the languages of cultures in our world today and tomorrow.

We are cautiously aware of the immensity of the challenge of such an undertaking. Please note that we are not seeking to describe all impossibly complex linguistic properties, nor each and every miniscule cultural detail. We are only (!) researching communicational "differences that make a difference" and "significant similarities," -- particularly in situations of high affect: what consistently "turns people on or off" in intercultural encounters.

[97-99]
All of us should recognize that the peoples of every speech community have what to them is an acceptable and--we anticipate--identifiable system for communicating certain culture-bound ideas or feelings. Such result from the use of conventionalized marks, signs, gestures, or other codes. Situations, objects, actions, or conditions also "speak or tell" discernible and associated ideas for other people--and, hopefully, for us. They also can convey comprehensible feelings.

Systems and means, conventions and situations, modes of behavior and of expression, and acceptable levels of meaning vary from each other and, most extensively, between differing peoples. Thus we have "LANGUAGES OF CULTURES." [10, 26, 29, 33, 34, 35, 37, 38, 40, 42, 44:90-93, 45, 50, 52, 53, 59, 63, 88, 90, 91, 96, 106. Note concern­ing bibliographic references follows.]

Those of us who attempt to investigate the scope, parameters, and researchable ramifications of the mysteries and riddles of the languages of cultures continually are jarred with come-up-ance in awe at the communicative processes that can and continually do take place. [4, 6, 56]

As the renowned cultural-linguistic analyst, Eugene Nida, expresses it,

The most intriguing aspect of language is meaning, but we have only begun to explore the intricacies of its structures and its relations to [intercultural] communication.

In some respects we are like the scuba diver who, when he first visits a coral reef, is amazed, bewildered, and intrigued by the abundance of life forms and their intricate inter­relations.

At first we are almost overwhelmd by what we discover in the semantic structures of languages, but once our explorations have begun, there is no turning back. Each discovery is only a prelude to more varied and greater discoveries. [75:9; see also pp. 68, 78.]

Challenges in categorizing each such discovery, and making each translatable [76] for use by other investigators and reporters, were proposed some twenty years ago by Dr. Nida. He suggests at least these language fields to be studied: (1) ecology, (2) material culture, (3) social culture, (4) religious culture, and (5) linguistic culture. [75:68-78; see also 68, 69, 73, 74, 97-99.]

He views languages as basically a part of culture, indicating:

...words cannot be understood correctly apart from the local cultural phenomena for which they are symbols.

This being the case, the most fruitful approach to the semantic problems of any language is an ethnological one. This involves investigat­ing the significant of various cultural items and the words [and/or other language codes] used to designate them. [75:78. Emphasis added.]

The vast new research frontier comprehending the languages of cultures boggles the mind. [7, 28, 53, 60, 97, 98] Which "languages" will you learn to use, as they relate to your own or other languages? Of the probably 6,000 spoken [1600+ written] languages of the world today, English, French, Chinese, Russian, and other so-called "sophisticated" languages each have millions of ide­lects, or potential ideas and their communicators. [39,53]

Farb says:

The existence of speech situations and speech events demonstrates why no one can adequately [meaning: completely?] learn a foreign language by instruction [alone]. A course of study teaches merely the vocabulary and grammar, not the [full complexity of] appropriate situations in which to use the alternative ways of saying something that every language offers. [33:39]

Farb uses his whole book to develop it in a broad range of concerns considered to be valid. He suggests many fields of needed research, in order to enhance meaningful communication. And we recognize that there are many today who are making significant investigations; yet, there is so much yet unknown. [80, 98]

It would be fascinating to make various applications of the 15 million plus comparative findings of Charles Osgood and others, on the basis of only about 600 terms dimensionally weighted, from 30 or so representative cultures. They call their work a WORLD ATLAS OF AFFECTIVE MEANING. [79; see also 68, 69, 71, 72] Consider what it would entail to comprehend all idelects for all cultures!

Or, we might like to probe the depths of subjective and linguistic culture with Harry Triandis
and associates, seeking out roles, attributions, habits, and like dimensions. [52:17-23]

Lorand Szalay’s pioneering work on psychocultural dictionaries stretches the mind with great potential for avoiding uncertainties in human communication between far-ranging patterns of thought and expression. [52:1-16; 93; personal notes.]

Para-normal (PSI-) languages also confront us on new extremities of attractive research challenges. [70:96-104] Reaching those who are in many ways communicationally unreach is another frontier that for some has become almost obsessing. [84] To mention these horizons is, for many of us who feel thin in thick things, [28] an intellectual exercise in deep frustration that we cannot yet perceive at all.*

We hope to entice or persuade many able people to join us and our colleagues, to delve deeply into new discoveries of how people “make sense” to each other. [28, 36, 66, 67, 78, 83, 104, 105] Any who will can aid in considering, for example, the multifaceted and myriad motions and intriguing insights of “micro-momentary expressions,” [64, 102] the cues of “cautions with cultural contrasts,” [99, 106], or even the limits of “languishing languages.” [53:91-95] Such possibly intriguing topics can be as contributory as those of socio- and psycho-linguists who determine languages of character and other forms of micro-, meta-, and para-communication. [50, 54, 75, 90]

In the intercultural arenas of business, government, and even in education, there is a challenge to aid with conditioning languages and disturbing non-senses, or: how and why some communications do or do not succeed in given circumstances. [1, 5, 11, 12, 16, 86, 89, 101]

Time, the vehicle and task-master of language, has hardly been touched, investigationally speaking. [Some beginnings: 37, 40, 41, 44:154-7; 59, 97.] Those of us who are Latter-day Saints (Mormons) have a scriptural injunction to do something about what as yet few of us do all that well:

Study and learn and become acquainted with languages, tongues, and peoples. (See Doctrine and Covenants 90:15.)

That is our challenge, and we for our part are trepidatiously willing to accept as much of it as we can. We welcome co-workers from any discipline, for there are few if any in our world today who are not or will not be using increasingly diverse languages of cultures. [6, 17, 21, 45, 49, 80, 90, 98, 105]

As a brief vista of one of the kinds of languages of cultures, I will conclude with a few samples from a hurried consideration of:

WAYS TO AVOID “SAVING NO!”
(Intentionally or Un-intentionally)
With At Least 40% Less Intelligibility

Herbert H. Clark, and others, [22, 23, 95] have found that a negative statement—depending on its dimensions and scope, of course—on the average takes about 40% longer to understand than a positive statement. At least this seems to be so for the participants tested in one cultural sampling.

We might conclude that when a simply stated "NO!" is unacceptable in a given cultural circumstance, it might take even longer to understand if other factors are added to the refusal, denial, or contrariness. Cultural examples abound; we can take time for but a few. [See 1, 23, 19, 27, 40, 41, 42, 46, 51, 52, 65, 90, 94, 95.]

SPEECH-LESS

1. SILENCE, as NO! (This may be different for children, youth, adults; by sex.) With or without anticipating a reaction, this negative response may mean: "I do not care!" "I do not know." or "I choose not to respond." Or, it may be intended as an angry insult, or as a sign that saying NO! verbally isn’t worth all that much effort.

Then, there are South Indian Paliyans, Quakers, and New Englanders and others, of course, who often use silence as a cultural way of limited response, indicating a variety of intended meanings, many of which are NO! [33]

2. Stylistic GESTURES as NO! (The speed of each also has a range of meaning.) To express a negative response, a head may be moved up and down. (In our culture this may mean no or "I am sleepy." or "I am listening.") In other cultures, a nod back and forth, right and left, or in a sort of figure eight motion can mean NO! or NO WAY! Or, even more confusingly, it may say "I am not certain either you or I understand."

Then there is the switching finger—which can be a sign of shame, or—in some cultures—a call to a pet, or worse. We are familiar with the uncommittal shoulder shrug for "maybe not" or "I am not sure." In other places in

* The expanded bibliographic references for this treatise give an almost [new resources reach regularly] up-dated sample of examples of creative thinking for communications principles only sporadically touched in this review of some of the languages of cultures. Research suggested by the questions and challenges presented here can be enhanced by using findings from the broad range of intercultural, linguistic, and language specialists cited herein—as well as from textual references and sources quoted in each of the texts themselves. [See also 21, 52, 53,F, 78, 85, 92, and current LRC bibliographies.]

[See "Bibliography and References" following.]
the world a tilt of the head must accompany this motion to encode a non-verbal negative. Caution is wise in use of gestures which can appear to be defiance when affection is intended, or where finger codes (o.k.) of the English world are interpreted obscenely.

3. Expeditious EXITING can be more than subtle NO-ing! This is especially so with moods of disgust, with vocal "harrumphs" accompanying. This might convey less than 48% in its quick intelligibility. In fact, such motion is often quite plain in its negative connotations.

4. STARING as a negative may be much more effective than a simple NO! Or, a turn to the right when a request for left is made says, possibly, "I am not about to buy your directions" (which may be less than a subtle way of indicating a negative rather than positive response). Stopping or speeding up sometimes gets a negative message across, like a stuck out tongue does in some cultures. (Watch out, that could mean "I am thinking!")

5. Almost speaking is the MOUTHED (but voiceless) NO! This can be subtle or direct. It works, too, from a little distance. (That may be best in threatening circumstances.)

SPOKEN (Often in company with gestures, paralinguage, etc.)

6. TONAL: The "Un'-uh!" or "uuuuummm." (Usually with a shake of the head, motion of a finger or hand, or shoulders). Sometimes only a grunt will do.

7. REVERSE-QUESTION says NO! "Why ask some odd-ball thing like that?" (Or more politely:) "Do you want I should believe that?" Or, "Must I, really?"

8. Some peoples make an art (?) of the SARCASTIC or CRITICAL negative statement. "Oh, come off it!" "You're nuts!" "Go jump in a lake." Or some APOLOGIZE: "I don't think I can accept that." "I'm sorry; I did not get it." = NO!

9. TANGENTIAL STATEMENTS as NO! This is performed by politely or rudely changing the subject, repeating an objectionable question in a different tone or manner, or by putting someone off. (The latter often happens in many cultures as parents somehow say to children: "I will think about it later." Or, "Let's talk about that the second Tuesday of February.")

10. HIDDEN NEGATIVES appear in most speech patterns. In English, for example: Stop = Do NOT go (on)! Try a positive translation for UNimportant, DISsable, absent, hardly, exclude, scarcely, doubt, few-if-any, etc. All of these potentially-negatives can be more confusing than their positive counterparts. It may depend on a context-particularly so when used in conjunction with any of the previously discussed modes of NO!

Rather than double this ten to twenty forms of NO!, I will add a couple more, take a quick look at two or three other similar: "languages"—then conclude.

Consider the negative (or positive) cultural connotations of such antinomies as: ALL SENTENCES USING THESE EXACT WORDS ARE FALSE; or ALL MEN ARE LIARS; AS A MAN I AM ONE OF THE BEST. [33:129-133]

Open PREVARICATION can be NO! Outright lying can be just "funnin'" or "white lying" (perhaps to save face or to avoid embarrassment for somebody), or can be bold-faced lying—either as offensive or defensive NO! Too, an oblique truth or half-truth can be prevaporation as well, and often appears as part of a cultural code which tells when or when not to be acceptable. When someone asks how are YOU feeling, what do YOU say? [Farb devotes almost a whole chapter to this form of negative response: 33:129-133. See also: 5, 7, 18,23, etc.]

We could have as easily used the language of YES! Possibly that word has been at least 48% more understandable. Or we could have reviewed languages to console, or to show care—as these vary from culture to culture. (DO YOU console your employer the same way as you do your spouse, or the truck driver whose fender you just bashed, or that child whose favorite toy is broken?)

By now it should be more than obvious that, in the languages of cultures, we are dealing in far more than vocal idelets or printed terminology. Think of American political language. Who understands all the implications? We are certainly obligated to try to understand. It seems to me that the more metatalk we hear, the more we have new challenges to try to solve the riddles of intercultural communication: the languages of cultures.

If we are to decipher these riddles, and we can do so, I am sure we must be more precise and incisive than we now are. We do, with many others—hopefully including YOU—accept the challenge.

BIBLIOGRAPHY and REFERENCES


56. Laird, Alarcon. The Miracle of Language.
   A. USA - Mexico
   B. Anglo-Naive
   C. USA - France


100. United States-Japan Trade Council, Washington, D.C.
   A. Japan and America: We're Talking, But Are We Listening?, March, 1974.
   C. Our Two Countries: Perspectives on our Views of Each Other. 1972.

101. Wilson, Herbert B. "Cultural Literacy Inventory". The Cultural Literacy Laboratory, 1974. Multicultural Education Center, University of Arizona.


Different PEOPLES have discernably distinct ways to THINK in, WRITE in, SING in, and otherwise USE (or even to abuse) LANGUAGES -- as BRIDGES to Understanding, Feeling, Believing, and Doing.

While no temporal languages always are consistent within themselves, nor None wholly uniform with others; Though all languages are complex, and each is some ways unique; Yet all are of inestimable value to those who employ and enjoy them consistently or uniquely.

There are languages of PRAYER, and to SHARE, And some are used to reveal CARE. Languages of EMPATHY and TRUST, Languages of SYMPATHY and MUST.

Languages by AGE, and varied SITUATION, Languages SAGE, of MULTI-COMPUTATION; Of FORESIGHT, HINDSIGHT, and DELIBERATION; Of INSIGHT, FORTHRIGHT, and of MEDITATION.

Languages of TOUCH and SMELL; Of MUCH, of DWELL; Languages of SPACE and TIME, of RACE and RHYME, Of range from GRACE to CRIME.

Languages of PLACE, and FACE, and CHASE, And HASTE; of TASTE, and WASTE; Of PRIME, and GRIME, and PANTOMIME.

Language Patterns are of THOUGHT -- or NOT! Most can be TAUGHT but rarely BOUGHT. Some are HOT, or of choice by LOT. Languages of SWEARING, DARING, and of BEARING; Of COPING, MOPING, and of HOPING; Of MATH, and WRATH, and for a BATH.

Languages of DRAMA, TRAUMA, and of ART; Of MAN, and CLAN, and of the HEART; Of BAIT, and RATE, and to TRANSLATE; And some INNATE, that do RELATE.

TRADE languages, STAID languages, And some POLITICAL; Some to EVADE, to make AFRAID, To UPBRAID, to PERSUade, or be CRITICAL.

There are languages UNIQUE, others OBLIQUE; Some TABU, while others ESCHEW; Some play GAMES, and another that SHAMES; Some to ACCLAIM, yet others DEFAME. Some CONTROL, or CAJOLE; Some seem to BORE, or to CEASE NEVERMORE.

There are languages to SPURN, DISCERN, ADJURN; Some of VERSE -- or, like this, worse; Some can CURSE, DISBURSE, or Try to DESCRIBE the UNIVERSE.

Some languages are BRIGHT, and INVITE; Others UNITE, DELIGHT, IGNITE, INDICT; Some MAKE LIGHT, or SLIGHT; yet Others seem "RIGHT." Some CONSOLE, some are DROLL; Some TAKE a TOLL, some HEAL the SOUL.

People's LANGUAGES are systematic means of/for communicating ideas and feelings -- using Conventionalized signs, marks or gestures; or The suggestion by objects, actions or conditions of Associated ideas or feelings.

There are languages of MUSIC, And of FUN things CHIC From every kind of BAILIwick; Of PEDIGREE, and LAW DECREE; Of what is FREE; of HE; of SHE!

There is language to INHIBIT, or to EXHIBIT, And frequently some used to PROHIBIT. Language of SONG, of WRONG, of what may be STRONG; Of CHEER, and SMEAR, and what seems DEAR; To SEE, to HEAR, to FEEL, to FEAR. Languages THEORETICAL by some seem HERETICAL, or SLICK, or THICK, or lean too hard on RHETORIC. Languages that are PLAIN may be for DISDAIN, or To RESTRAIN, or to ORDAIN, or declare INSANE. There are languages of WORK, or SHIRK; of SMIRK; Of LOVE, and HATE; of JOY, of FATE.

Some languages are BOLD, and others are COLD; Some are for DANCE, or CHANCE; Some come from GLANCE, or TRANCE; Some do ENHANCE, or aid ROMANCE.

There are MISTALK, MIXTALK, METATALK, and "NO!" And SQUANK, and BALK, and SHOCK, and SHOW; And DRESS, and "YES!" and MAYBE, and "GO!"

This cultural look at language, hardly started, With little insight yet imparted, To keep the mind ahead, on top, Now must stop!

(...you know this could go on all year! But, if it did, we might cause fear That language bridges could not get built; Instead there'd only be a sense of guilt. So, let's get on back to doing work -- To find for languages each quirk!)

This 'poetry' at first may seem less grand, Until YOU, too, write: try your hand:
TOWARD A LANGUAGE INSTRUCTIONAL MODEL

Robert W. Bushman

Languages and Linguistics Symposium
April 7-8, 1975
Brigham Young University
TOWARD A LANGUAGE INSTRUCTIONAL MODEL

Robert W. Bushman

The overall goal of language instructional design is, of course, the optimization of language learning. Though we most frequently discuss subordinate or peripheral concerns, the paramount issue in the back of our minds is how to best go about teaching language. We all alert on experimental results, nascent developments or even pure speculation which seems to offer some promise of a better way. The object of this paper is not to propose some better way, however, but to look at how we go about finding the better way. I don't pretend to have any grand answers to the question, how should language be taught, but I am concerned with the processes we use to go about answering that question. Somehow, it seems that the process of language curriculum development should occupy our attention, as well as the content, methodology, or results of that curriculum; for it may well be that our success with a language curriculum is as much a function of the developmental process itself as of our skill in applying that process. Indeed, as I consider the various processes which have been or are now being used, it seems that in general they are quite ad hoc -- that there is no single, considered system. Moreover, the processes are better characterized as intuitive than scientific. Perhaps the artistic endeavor is just as well, but there seems to be an idealistic need in me that reaches for some system that can actually generate answers to the question, how should language be taught. Perhaps that ideal reflects my need for structure, prediction or certitude in a domain that is far too complex to be dealt with in that way. Nevertheless, the ideal offers hope, and that hope leads me to explore here the possibility for success of a prescriptive model of language curriculum development.

To begin, I will model three general development procedures and assess them against the criterion of prescriptive power, the most demanding requirement for a model of this type.

SIMPLEX MODEL

The first, most familiar and elemental model I have labeled the simplex model. If I may be somewhat judgmental, it takes a generally simplistic notion of what language is, and uses whatever instructional methodology that notion seems to suggest, prima facie. Some examples of the simplex model would be the puristic forms of grammar—translation, the natural or direct method, audio-lingualism, total physical response, St. Cloud, community language learning, etc.

Grammar—translation assumes that the essentials of a language have been adequately captured by a given grammar and lexicon, and therefore if one learns that grammar and lexicon, he will have essentially learned that language. It is very simple logic.

The direct method sees native language learning to have occurred not from academic exercise, but from a great deal of exposure, mimicry, and interaction. Therefore, the way one should learn another language is by being bombarded with it, by mimicking it, and by being forced to cope with it. Direct logic, direct method.

Strict audio-lingualism holds language as a set of psycho-motor habits. Therefore, language should be taught through a sophisticated process of psycho-motor conditioning along the classical stimulus-response paradigm.

The St. Cloud, or audio-visual method, believes that conversational communication necessarily takes place in both auditory as well as visual media. Therefore, graphic illustrations are to be used in addition to aural-oral experiences.

Total physical response observes that children hear a great deal of language, indeed, follow a great many commands in their mother tongue, before ever uttering a word of it themselves. Obviously, then, language students should build up a large latent competency from listening comprehension before being required to produce utterances of their own.

In the view of community language learning, language is learned as a response to felt psychological and societal needs. It then lets students interact from those needs, querying an outside source for the foreign language data necessary to do so.

If these rather gross simplifications can be forgiven, these and other like methodologies may be characterized by the existence of some underlying philosophy of language, language use, or language learning from which is derived some face valid instructional approach. In this paradigm are two assumptions: 1) that the essence of language learning has been adequately captured by the philosophy, and 2) that the indicated methodology is truly valid for that philosophy. While it is typical that underlying philosophies are constantly challenged, the second assumption of method validity escapes much attention, though both are critical to overall validity of the model.

It should be said that whatever methodology of the simplex variety is used, students usually do learn language therefrom. But because an evaluation procedure is usually not built into the process, neither of the above assumptions are tested. It is therefore not known whether students learned whatever language they learned because of or despite the method used. The simplex model will generally be able to say, trivially, that language learning will result from an application of Method X, but is incapable of higher orders of questions, such as, "Does Method X work better than Method Y?" or, ultimately, "What is the best method?"

PRAGMATIC MODEL

The next model I label the pragmatic model, because it is not necessarily founded on any particular philosophy of language, learning, or instruction, but dedicated to what works. It is
essentially an evaluation model, aimed at answering the question, "Does Method X work better than Method Y?" The model does not prescribe instruction, but from trial and error can identify the best available options. On the loose end of the line, this is contemporary eclecticicism, using a bit of this here, and whatever seems to fit over there. On the tight end, this is experimental, educational research, carefully controlling variables, constructing valid and reliable measures, and analyzing results for statistically significant differences. A splendid example of this model in operation is our own LTM, which is now conducting simultaneously a variety of experiments to measure the relative effectiveness of certain methodologies. Without preformed biases or loyalties for any particular method, the LTM will be impressed with hard, empirical results, and not fancy trappings. In the end, the LTM will indeed be able to say with some confidence that Method X is better than Y, but not as good as Z.

It is interesting that while the pragmatic model can decide which method is better-- a higher level of adequacy than the simplex model-- it is dependent on the simplex model for inputs, as it were. The pragmatic model does not generate its own methodologies, but merely weighs existing methodologies on the counter-balance. Furthermore, while the pragmatic model can show correlations of method and outcome, it does not work to posit causal relationships which predict or prescribe optimal language instruction. It should be said, however, that through a process of evaluation, LTM is producing hybrid methodologies, and thus goes beyond this model.

SYSTEMS MODEL

At the highest level of adequacy, we idealize the true prescriptive model, which I call the systems model, as it is characterized by a systematic problem solving approach. I do not claim that this model has working examples in the language world, but I would like to sketch what the model could look like, and possibly how it could be operationalized. The model bears outward similarity to the simplex model in that it begins with assumptions about language and proceeds to prescribe the instructional approach from those assumptions. The difference is that where the simplex model proceeds intuitively, or at best in an a de hoc manner, the systems model proceeds systematically, without necessarily seeing the end from the beginning. It is more like a mathematical algorithm which, proceeding step by step, leads to a result. The resulting language curriculum in this case cannot be as sure as the result of a mathematical formula, as we are dealing with causal relationships which we can only postulate, never prove; and we are dealing furthermore in a probabilistic domain, working with free agents. Nevertheless, the promise of obtaining results of some power through a methodical process seems sufficiently bright to warrant a major effort in this direction.

When we think of a system of curriculum development, we first think of the classical curriculum development model consisting of three general stages: 1) definition of behavioral objectives; 2) development of instruction; and 3) validation of that instruction (Faust, 1974). The first stage sets the goals of the instruction in terms of student outcomes, focusing on what the student is supposed to actually be able to do when he exits the instructional process. The second stage analyzes those behavioral objectives into enabling objectives and classifies them by types of behavior: psycho-motor, affective, and cognitive, the last being further broken into memory, classification, rule using, and rule finding behaviors. Once the type of behavior is known for a given objective, then rules are applied which prescribe the type of known instructional design most appropriate to that type of behavior. The last stage of the process is a cybernetic loop which evaluates the end product against the criteria of the original objectives, and indicates where changes in the system are necessary. Actually, this stage may go on simultaneously with other stages, evaluating during as well as after those stages.

Though the middle step of instructional design seems to be able to prescribe rather thoroughly what should be done, it in practice leaves open many options, such as sequencing, presentation mode, media use, etc. In fact, there are some who maintain that the model lacks prescriptive power to any interesting degree (Clark, 1975). Indeed, when we apply this model in the language learning context, we find that language behavior is sufficiently complex as not to be so handily broken into pieces to be separately dealt with. We find that we must rely on sheer creative invention of presentational devices, for which we draw on the existing options created by the simplex model.

I suspect, however, that if this sort of situation results from the developmental process, then it is finding that our system lacks significant prescriptive power to deal with this most fundamental problem of design, then it could well be that we have not used the full power of the first stage-objective setting. Too often in education, objective setting turns out to be rather superficial activity, done not by empirical study, but by a few hours of thought, pulling goals out of the subliminal and setting them down in a form acceptable to contemporary behavioralists. If this is all the first stage consists of, the second is deprived of the very basis it requires for truly prescriptive power. In the model I am suggesting, the first stage could well account for the greatest portion of activity in the entire process.

Much groundwork has to be laid before we are in a position to define behavioral objectives; that groundwork being a "task analysis." This is a thorough inspection of what the final product person actually does or is intended to do in the real job environment. In the language context, is the issue "How is language used?" After the task has been made explicit, then the questions are asked, "How much of this task can the student already do?" and "What part of the remainder is to be handled through formal instruction, and what may be handled through other means?" Only after a very thorough task analysis can we be in a position to specify those behavioral objectives to be dealt with by our instruction.

It is my opinion that there has not yet been
a sufficient task analysis done on the type of language activity that we normally target. Linguistics has addressed only a subset of the problem, and moreover, is not fully equipped to deal with the whole problem, because the whole problem is not within its domain, but mostly without. Communication science, psychology, sociology and anthropology also deal with the problem, but each in its own limited way. What we have is a multidisciplinary problem, that requires a multidisciplinary approach (Politzer, 1972). We must come out of our cloisters and start talking with each other. We must sit down and work together, rounding out a full-blown model of language use. We must get outside and find out how it is actually done in the real environment.

In addition to the task analysis, I also propose an analysis of how adults actually learn a foreign language. Our studies are typically of the form: treatment application and result measurement, and not so much finding out specifically, longitudinally, and anecdotally what our students actually do when they are learning a language. We have been doing this with child language acquisition, why not with adults?

It strikes me that once we have thoroughly addressed the two issues, "How is language used?" and "How is language learned?" and have created explanatory models which we are reasonably comfortable with, then we will be in a position to generate optimal instructional approaches.

To review this discussion, it would seem that I am strongly biased toward the systems model. In fairness to the other models, though the systems model may hold the greatest hope of prescriptive power, the more fundamental criterion for any resulting methodology is its effectiveness. It could very well be that a method of the simplex variety will come along after all, where imagination and creativity thrive. It could just as well be that as a result of a lot of experience, and hard, empirical evaluation, the pragmatic model will be the shining light. Though the systems model may be the highest valued model per se, we admit we are not in the model building business, but the language training business, and models are useful only insofar as they serve that end. All three models lead to language learning, and together, provide a synergy which argues for vigorous pursuit on all three fronts.

References


SOME PRACTICAL APPLICATIONS OF JUNCTION GRAMMAR
IN COMPUTATIONAL LINGUISTICS

Charles D. Bush
Those of you who were here yesterday may have noticed a slide presentation about our Project set up out in the hall. I saw that presentation for the first time yesterday and realized that it said in twenty minutes what I expected to take 40 minutes to present today. So, if you will watch this presentation for twenty minutes, instead of listening to me for 40 minutes, it will give me a little more time to talk about some specific test runs and cost-effectiveness statistics not covered in the presentation. To those of you who have seen it already, I apologize. You can take a twenty minute nap and then I'll be back up here.
The purpose of this presentation is to answer some of the most commonly asked questions about computer-assisted language processing at BYU.

Who works on the project?

What is the project? In 1968 a new theory of language structure, later to be known as Junction Grammar, was developed at BYU by Professor Eldon G. Lytle. It captured linguistic universals previously unnoticed, and seemed an appropriate medium for computer translation research.

In 1969 test programs written to translate Russian to English were successful, and a gift of private funds resulted in the formation of the BYU Automatic Language Processing Research Group.
The group soon thereafter received two contracts from the Department of Defense, resulting in experimental prototypes which translated from English to Spanish, French, Portuguese, German, and Japanese, as well as from Russian to English. Further research has been supported by BYU itself and by the LDS Church, with the expectation that the extensive translation needs of the Church would be at least partially served by computer-assisted systems before the end of the 1970's.

Currently, Professor Lytle heads a team of eight full-time researchers and over twenty student assistants. The team is organized into six divisions: analysis, transfer, synthesis, lexicons, systems, and special projects.

Why is the project needed? The need for economical, fast, and accurate translation is widely recognized. There are unmet translating needs in virtually every area of international affairs: business, research, politics, and religion.
Henry Fischbach, head of a New York language service firm and ex-president of the American Translators Association, has said, "There is a terrific shortage of competent technical translators."

For instance, the LDS Church, not by any means the largest of international organizations, translates yearly into 15 languages over 17,000 pages of material. Translating a basic proselyting kit of the standard works, missionary discussions, and 14 tracts takes at present six years.

Hasn't computer translation been tried elsewhere with poor results? Yes, millions were poured into automatic computer translation in the 1960's with disappointing results, and in 1967 a U.S. government report discouraged further research in this area for several years. However, due to better hardware and a better perspective of what computers can and cannot do, and better linguistic bases which now seem to hold out greater promise, projects in computer-assisted translation have been initiated all over the globe.
In other words, an intermediate code is needed which will represent the relationship of the words in a sentence regardless of the language used. The computer can then use this code to translate a given source language into any other language. Junction Grammar forms the theoretical base for the BYU translation system.

What about computerized dictionaries? A necessary and integral part of translation is the use of computerized dictionaries. These automated dictionaries allow a translator to replace his desk dictionary with a computer terminal on which he can instantly check all meanings of a word.

What is Analysis? In the Analysis phase of the translation system, the knowledge and logic of the human together with computerized dictionaries in translation facilitates both accuracy and consistency in the selection of translation equivalents.
According to Charles Bush, Lexicons Division Head, our present dictionary system consists of approximately 10,000 meanings in the English source dictionary and roughly the same amount in the corresponding target language dictionaries. A production-size dictionary would be much larger, depending upon the type of material to be processed.

We have characterized the BYU translation process as a sequence of three basic steps: 1) Analysis, 2) Transfer, and 3) Synthesis.

What is Analysis? In the Analysis phase of the translation system, the knowledge and logic of the human together with computer programs are utilized to resolve ambiguity inherent in language. The programs in the system are designed to interact with a human after detecting alternative processing paths in the input text. Where current knowledge does not permit the automatic selection of the correct alternative, . . .
What kind of interaction with the human is the computer programmed to do?

Daryl Gibb, Analysis Division Head, characterizes the two types of interaction as:
1. Referential
2. Syntactic
What is Referential Disambiguation? The first ambiguity encountered by the computer is when a word has more than one meaning. The analysis programs detect these ambiguities and request the operator to choose the meaning that best suits the word sense needed. This phase is called referential disambiguation. Suppose one wants to translate the sentence "WE SAW THE BOY IN THE CAR THAT THE GIRL LOVES." The first word the computer finds to be ambiguous is the word "saw."

Since the computer cannot distinguish between a "saw" that one cuts wood with and the action "to saw a board", or the past tense of "see" meaning to look and the past tense of "see" meaning to visit, the operator must intervene. The remaining words of the sentence that the computer will find to be ambiguous are • • •.
girl, ... 'in-car' modify?"
The operator replies with the correct answer and the computer moves to the next structural ambiguity.

and loves. Each meaning of a word entry is assigned a unique index number when the dictionaries are built by which the computer can access and match both the source and target language entries. When all referentially ambiguous words are accounted for, the system retrieves the next sentence if a text is being translated.

What is syntactic disambiguation? The last phase of Analysis is called syntactic disambiguation. A sentence such as "WE SAW THE BOY IN THE CAR THAT THE GIRL LOVES" not only has referential ambiguities but is also syntactically or structurally ambiguous. Since prepositional phrases are potentially ambiguous, the computer promptly responds:
"What does 'in-car' modify?"
The operator replies with the correct answer and the computer moves to the next structural ambiguity.

The phrase 'that the girl loves' is ambiguous and can either modify car or boy. If in context one knew that it was the boy that the girl loves, he would respond NO to the question, "Does that represent 'car'?"

and YES to the question, "Does that represent 'boy'?". When all ambiguities have been detected and answered, the sentence is represented as a completely unambiguous structure, by the rules of Junction Grammar, and it is passed on into the next phase of the system.
What is transfer? Sometimes it is necessary or convenient to adjust the Junction Grammar code prior to synthesis so that the resulting sentence will seem "native."

For example, speakers of English say "I am hungry" while those who speak Spanish say "I have hunger." The BYU translation project has developed a high-level programming language based on Junction Grammar representations which serves as a medium for manipulating sentence structure. A library of sub-routines written in this language is being prepared for each source/target language pair.

Floyd Billings, Synthesis Division Head, explains that in transfer, each sub-routine is keyed either by Junction rule, a semantic index, or a combination of these, and when executed, replaces that part of the representation which activated it with a near equivalent, compatible with the target language. This process of adjustment is referred to as Transfer.

1) They determine the order of constituents in the lexical

punctuate word strings

in a manner consistent with conventional practice. After the sentence or text has passed through the lexical rules of the Synthesis routines the result can be printed or displayed on the video screen, if necessary, for further editing.
What is Synthesis? The process which converts the Junction Grammar code back into natural language is called Synthesis. Since the special code is not ambiguous, synthesis does not require human interaction.

Floyd Billings, Synthesis Division Head, states that the synthesis system for each target language consists of language specific lexical rules which construct word sequences for Junction Grammar representations. These Lexical rules perform five functions:

1) They determine the order of constituents in the lexical string.
2) They govern hiatus or understood element phenomena.
3) They match words in the target language dictionary with semantic indices passed by the Junction Grammar code.
4) They provide for affixation patterns required by the morphological conventions of their specific language.
5) They punctuate word strings in a manner consistent with conventional practice. After the sentence or text has passed through the lexical rules of the Synthesis routines the result can be printed or displayed on the video screen, if necessary, for further editing.
What are the post-editing capabilities? If, for some reason, parts or all of the sentence or text being translated do not successfully pass through the translation system, a post-editing capability can be employed to upgrade the output.

Those parts unsuccessfully translated would be rendered word-for-word in the target language and displayed on the video screen for correction.

By using the post-editor, a human can replace words or phrases translated incorrectly; and magazine companies are using computerized editing systems for reasons of cost, accuracy, and efficiency. With these capabilities, the operator can insert a word, sentence, or whole paragraph instantly, without having to retype the whole text.
What, then, are the salient features of
insert words, phrases, or other
items that may be needed;

First, Junction Grammar facilitates the construction of an unambiguous semantic representation of any language. Second, this special representation, which is independent of the number of target languages, but also for the monitoring of analysis by persons not familiar with any of the target languages. The output produced by analysis can be synthesized into natural languages simultaneously or it can be stored for future processing.

Third, the compartmentalization of

If the output is totally unintelligible, and if it would take too long to add, delete and rearrange, the editor has the capability of inserting a new and desired translation dynamically by the use of the translate (T) command. Though it seems slow and tedious, it is interesting to note that many newspaper and magazine companies are using computerized editing systems for reasons of cost, accuracy, and efficiency. With these capabilities the operator can insert a word, sentence, or whole paragraph instantly, without having to retype the whole text.
What, then, are the salient features of the BYU project? There are five salient features of the BYU approach:

First, Junction Grammar facilitates the construction of an unambiguous semantic representation of any language. Second, this special representation, which serves as an interlingua, allows not only for a single-time computer-assisted analysis regardless of the number of target languages, but also for the monitoring of analysis by persons not familiar with any of the target languages. The output produced by analysis can be synthesized into several natural languages simultaneously or it can be stored for future processing. Third, the compartmentalization of the BYU system into discrete analysis, transfer, and synthesis components facilitates the process of change necessary to any developing system. Fourth, a special interactive design utilizes both the extensive linguistic experience of the human operator and the speed and power of the computer. Fifth, the BYU team has developed an error-recovery feature which produces at least a word-for-word translation if for any reason the more sophisticated routines falter or have not yet been programmed for a given structure or idiom. Let us discuss each of these features separately.
Junction Grammar is a generative language model positing three levels of language other than the surface level. On a level lying two levels below the surface, semantic elements are seen to be joined in several different ways including the following:

1. adjunction (the relation of subject to predicate, verb to object)
2. conjunction
3. subjunction (that is, modification)

From the focus on these and other types of junctions, the name Junction Grammar is derived. Crucial to Junction Grammar is the notion that structure and meanings are inextricably connected, contrary to the early assertions of transformational grammarians and some semanticists.

Some of the basic concepts of Junction Grammar are discussed in a recent Mouton publication of the Janua Linguarum series, *A Grammar of Subordinate Structures in English*, by Eldon Lytle. The application of Junction Grammar to machine-assisted language processing is discussed in a forthcoming issue of the Journal of the Association for Computational Linguistics. More detailed accounts are found in the yearly BYU Linguistics Symposium Reports for 1972 and 1974.
There are at least four implications of an intermediate representation:

a. For any given text to be translated, . . .

Computer-assisted analysis need be accomplished only once, while each human translator must repeat this process.

Compare the human-translation system, where each translator must repeat the analysis of sentences in the source language, with the computer-assisted translation system described.
This has obvious implications for consistency, accuracy, speed, and cost.

b. The human who monitors computer-assisted analysis need not know a foreign language. He answers questions only about the meaning of the source text.

c. The semantic code produced by the analysis program can be stored in the computer memory for transfer and synthesis into other languages at any later time. This means that other components of the system, which are fully automatic, can operate on a backlog of stored code during night hours or other periods when humans do not work.
d. It should be mentioned that the approach of the BYU group is a thorough analysis of the exact structure of each sentence. This careful approach allows synthesis from an intermediate representation into many languages simultaneously, and it will be of greater usefulness to organizations dealing simultaneously with many languages than would a one language to one language configuration.

Early translation systems tended to favor massive, undivided programs which were difficult to modify and improve.

The BYU system is compartmentalized into discrete modules of analysis, transfer, and synthesis components, and many of these are further subdivided into necessary phases, greatly facilitating the developmental change necessary to any viable system.

A fourth important feature of the BYU approach is the heavy use of interaction with a human operator during the analysis phase of the program. It is anticipated that much of this interaction can be eliminated in the future through sophisticated logic algorithms, but at present it seems a rational compromise for the following reasons:
a. lexical and structural ambiguities are extremely difficult to resolve—resolution frequently relies on factors far outside the boundaries of a given sentence or paragraph.
b. Post-editing is manifestly inefficient, as it incurs not only human time to post-edit, but also machine time to process incorrect analyses.
c. Pre-editing involves vast amounts of time, distortion of the original text, and extensive knowledge of upcoming problems.
d. Interaction is limited to the analysis phase; thus the resultant unambiguous intermediate representation can be translated into many target languages automatically with a modicum of post-editing. In addition, the interlingua can be stored for later use as synthesis routines are developed for other languages.

Thus, the BYU approach incorporates Norbert Weiner's sage advice, given over a decade ago, to "render unto the computer the things that are the computer's and unto man the things that are man's."
The BYU team has recently developed a feature which produces at very least a word-for-word translation if for any reason the sophisticated linguistic routines falter on any part of a sentence. Thus, structures, idioms, and words which have not been programmed can be handled without destroying the results of processing on other parts of the sentence. This feature also gives the team, as they attack any new language, a base program into which more sophisticated programs can be inserted as development progresses. It is at least an efficient automated dictionary and provides valuable editing capabilities.

This, then, has been an overview of the BYU computer-assisted translation project. At present, our system is still in experimental stages, but progress toward meeting the 1978 goal of implementing a prototype production model capable of translating from English to Spanish, French, German, and Portuguese, with the capability of adding additional target languages, is well under way.

The results of the evaluation are shown in the table found on the following page.

The first figure in each box is the amount of time it took the human operator to perform that step in seconds. The second figure in each box is the amount of computer CPU time that was required to perform that step in seconds and hundredths of seconds. The third figure is the amount of
I would like to continue in the vein of this slide presentation for a few minutes by asking a few questions and then giving the answers. It is always easier for me to ask the questions since I can prepare the answers ahead of time.

Perhaps the first question that might come to mind is, "Is this whole thing economically feasible?". Strictly speaking, we don't really know, but we do have some indications that it might be.

In the last few weeks (mostly in the last one week) we have run an evaluation of the system as it exists right now. The results are so fresh that those of you who came in early probably noticed that we were totalling up figures on the board up here as you walked in.

For this evaluation we ran one page of approximately 1,750 characters from each of these four source documents:

(1) "After Baptism, What?" (an LDS tract written by Mark E. Peterson),
(2) a story called "Baptism After Dark" (by Kathy Troxler, published in the LDS children's magazine The Friend),
(3) a selection from a simplified version of the Gospel of Luke (written by Stanley Morris, who was here earlier in the Symposium but has since left; it begins with the second chapter, the basic Christmas story about Mary and Joseph going to Bethlehem),
(4) "3208 Diesel Vehicular Engine Manual" by Caterpillar Tractor.

The results of the evaluation are shown in the table found on the following page.

The first figure in each box is the amount of time it took the human operator to perform that step in seconds. The second figure in each box is the amount of computer CPU time that was required to perform that step in seconds and hundredths of seconds. The third figure is the amount of
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Referential Disambiguation</strong></td>
<td>27:12</td>
<td>30:18</td>
<td>25:01</td>
<td>26:33</td>
<td>1:49:04</td>
</tr>
<tr>
<td></td>
<td>82:98</td>
<td>91:81</td>
<td>92:73</td>
<td>91:30</td>
<td>358:82</td>
</tr>
<tr>
<td><strong>Syntactic Disambiguation and Rest of Analysis</strong></td>
<td>2:24</td>
<td>4:26</td>
<td>2:51</td>
<td>6:54</td>
<td>16:35</td>
</tr>
<tr>
<td></td>
<td>32:86</td>
<td>70:80</td>
<td>30:27</td>
<td>46:04</td>
<td>179:97</td>
</tr>
<tr>
<td></td>
<td>(4:36)</td>
<td>(6.29)</td>
<td>(4:01)</td>
<td>(6:10)</td>
<td>(20:76)</td>
</tr>
<tr>
<td><strong>Transfer-Synthesis</strong></td>
<td>0:00</td>
<td>0:00</td>
<td>0:00</td>
<td>0:00</td>
<td>0:00</td>
</tr>
<tr>
<td></td>
<td>19:90</td>
<td>21:70</td>
<td>17:91</td>
<td>24:26</td>
<td>83:77</td>
</tr>
<tr>
<td></td>
<td>(2:60)</td>
<td>(2:84)</td>
<td>(2:34)</td>
<td>(3:17)</td>
<td>(10:95)</td>
</tr>
<tr>
<td><strong>Post-Editing</strong></td>
<td>1:31:25</td>
<td>1:45:20</td>
<td>1:00:00</td>
<td>42:25</td>
<td>4:59:10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2:01:01</td>
<td>2:20:04</td>
<td>1:27:52</td>
<td>1:15:52</td>
<td>7:04:49</td>
</tr>
<tr>
<td></td>
<td>174:83</td>
<td>234:06</td>
<td>178:71</td>
<td>193:09</td>
<td>780:69</td>
</tr>
<tr>
<td></td>
<td>(23:22)</td>
<td>(28:03)</td>
<td>(23.72)</td>
<td>(25.62)</td>
<td>(100:59)</td>
</tr>
</tbody>
</table>

Table 1.

Results of Evaluation, April 1975
computer charges (the amount that we were charged to run the computer) during that step. Now you may notice if you have sharp eyes that the amount of computer charge is not directly relatable to the amount of CPU time. That is because there is a charge for input-output operations and various other things, as well as for straight CPU time.

You can look at these figures and see some interesting things. Perhaps the most striking thing to notice is that the steps that take the most time are the ones that the human has to do. Transfer and Synthesis accomplish a great deal of work in the translation process, but they don't rely on any human interaction, and therefore don't take as long. Transfer-Synthesis only take an average of about 20 seconds of computer time with an average of about $2.75 in computer charges. And yet a large part of the process of actually translating the sentence takes place in this step.

On the other hand, the step where the money is really spent is the fourth section--post-editing. This step averages somewhere around an hour and a half of human time and several seconds of computer time per page. This is not an unusual amount of computer time, but it is a tremendous amount of human time. However, it should be noted that the person who did this work did it all last night. She is a native German, but she has never worked with this system before and has not had too much experience in translating. With some experience, this figure would probably go down.

Notice the difference between the four documents. The most difficult document in overall terms is probably that story from the children's magazine. And probably the easiest one is the Diesel Vehicular Engine Manual. Now that may sound a little bit strange; but part of the explanation is that in the Diesel Vehicular Manual they write in very straight-forward, although
technical, language. There are an awful lot of prepositional phrases just tacked on to each other: "the bearings of the camshaft on the right side of the cylinder block . . . ." But these are all concrete prepositions and they express concrete relationships that the person interacting on the terminal and the computer programs can handle easily. The story, however, has the problem of run-together sentences; it is a narrative. You get a sentence that says "this and this and this happened, period"; "then this and this and this happened, period"; and "So, this and this and this happened, period." While all of these sentences are logically tied together in the narrative, the computer has to try to process them as single units, as single sentences.

Now I'll go on to my next question: "How does this compare with human translators?" I am citing for my comparison data a publication by the Church Translation Department that outlines the guidelines of the amount of work that is expected of their employees--how much they are supposed to produce in an hour. To translate one page of 1,750 characters, a human translator setup is expected to take approximately 88 minutes for the whole process. This includes 30 minutes for the initial translation, 24 minutes for a review by two supervisors, 30 minutes for typing, and four minutes for proofreading. We consider that all of these operations, with the exception of one of those reviews, are included in the Computer-Assisted Translation System. Thus, it would take about 352 human work minutes to translate approximately this much material. The computer system took 425 human work minutes to do the same job plus the computer time and money. So, based on these figures, it looks like the computer system loses fairly decisively. But it is something like the old races between a horse and an automobile. In early days the horse always won, but, as they worked the bugs out of
the automobile, the results changed. Of course there isn't any contest at all when you put the two together now. We envision that the same thing will happen here as the computer system improves. Remember that this is 1975 and our goal date for a production ready system is 1978, so we have three more years to knock these statistics down.

The point has been made in the slide presentation that the Junction Grammar system differs from other systems and from human translators in that the Analysis portion of it need be performed only once. In order to translate this same material into another language, it would only require duplication of the Transfer-Synthesis and post-editing steps. That would add, under the present system, 316 minutes to the total instead of the full 425. Based on these figures, if we translate from English into three languages using the computer-assisted system, it would take 1,057 minutes. This is where we begin to get parity.

One more statistic: With some quick pencil and paper calculations, I have come up with an interesting estimate. If we can reduce the amount of post-editing time to an average of 2 1/2 minutes per sentence (where the average now is 3.8 minutes)—in other words, if we can make the output of our system of a high enough quality and also if we can train the people that work with the system to a high enough efficiency, and they can do one sentence in an average of 2.5 minutes—then we will be able to equal the human translator in a one-language-to-one-language configuration. Projected to three languages, the computer system will only take 668 human minutes, which is slightly more than 50%. This is the vision that we have—to get the system working on this level of translation.

Maybe I had better give you a chance to ask some questions before I ask any more.
QUESTION: Would you elaborate upon your "Fail-Safe" capability?

ANSWER: Sometime in December or January we came to the conclusion that in order for the translation system to be effective, it had to get something through for every sentence that was given to it. Up until that time, we were just testing the system with prepared sentences. If we would put in a sentence and it abended for some reason, we would run some documentation to see what happened to it and then go on to the next test sentence. Of course, that doesn't work very well if you are translating on a production level. So we decided to add to our system the capability that if, for some reason, the sentence could not be analyzed, synthesized, or transferred completely by the prime system, a backup system would produce the best translation possible and the human would have to fix up the rest of it in post-editing. If worst came to worst, if Analysis abended on the first word of the sentence, the post-editing would produce a word-for-word translation.

In the initial concept, the human would have to do everything from there, including putting on case endings, changing the word order, adding punctuation, and all kinds of things like that. After we had worked on the backup system for a while, we began to realize that there were a lot of very tedious things, like verb-subject agreement, adjective-modifier agreement, and things like that, that the human had to do on these output sentences that we could still assign the backup system to do. The Synthesis programs can generate endings, agreement, some word order, all kinds of things, even if they are not working on a completely analyzed sentence. By now, we have developed the system to the point where, in many cases, we can still generate the correct endings and articles, even though the sentence has bombed for some reason during the processing. The development of all
of this has taken considerably longer than was initially planned—would you believe that we guesstimated at the beginning that it would take two weeks to make the change? We have worked for three months, and the changeover is still not complete. That is one reason why post-editing is so high—-all of the recovery routines still don't work like they really should.

LYTLE: "Point out specifically how the post-editing can be reduced over the next three years."

ANSWER: The amount of post-editing will go down drastically with practically any improvement that we make in the programs. For example, the problem of having conjunctions of various kinds at the beginning of a sentence, which connects with other sentences in the text. That is one problem that Analysis has not yet handled, but will be able to handle when we program it to do so. As soon as that programming is added, then those sentences that failed in this particular test will go through correctly, and the post-editor won't have to fix them up anymore. Similarly, part of the problems that require post-editing at this time are due to the fact that many of the transfers have been identified but not written. The result is that some of the idioms and various things that needed to be transferred, weren't.

There are also problems with the lexicon, my particular area of supervision. There are lots of things wrong with the lexicons: words missing, incorrect features, things like that—-many of which we will only be able to discover as we run test material. We discovered several, for example, in this particular test: in these four pages we found 23 words that we
didn't have in our English analysis dictionary. Some of them are multiple occurrences of the same word. Several of these also were technical terms from the Diesel Vehicular Manual, like manifold, bearing, and some terms that I don't even know what they mean. Plus there were 49 meanings of words that were not satisfactory. In some cases, for example, a word had several meanings but we just hadn't put the one that we needed in yet. I think it was mentioned in the film presentation that we have about 10,000 meanings in our dictionary now. We are anticipating production dictionaries to have somewhere between 20-30,000 meanings, so we still have some developing to do.

QUESTION: How is the lexicon stored?

ANSWER: It is in a computer structure. The English Analysis lexicon, the lexicon that we use for analysis, is keyed by the word and consists of the word senses of the word, the associated features of that word sense (both language specific and language independent features), and a five-digit index number, which is used to represent that meaning throughout the processing of the sentence. In referential disambiguation, the human operator selects the appropriate word sense. From that point on, the computer uses the associated index number to represent that meaning through the rest of the processing.

QUESTION: Do you store things on tape?

ANSWER: On disc. The dictionaries must be available for direct access, so we have to use disc.

QUESTION: Is it transferred to disc in the process of Analysis, or is it
permanently on disc?

ANSWER: It is permanently on disc. We have programs that we use to build the dictionary directly on disc. The English Analysis dictionary presently takes about 40 cylinders of disc space.

QUESTION: Does this apply also to the lexicons of the target languages?

ANSWER: Target language lexicons are a little bit different in that they are keyed by the index number. They have their own set of language specific features and then the target language word for that concept, with various forms of that word that will be required in the target language. Target language lexicons are noticeably smaller and average about 15 cylinders each.

QUESTION: When a word comes up, are all of the possible meanings retrieved or only those made possible by the context at that point retrieved?

ANSWER: Right now, all of the possible meanings are retrieved. The computer looks at the character string and all of the possible variants of that character string are displayed for the human to choose from. Now, as the system develops, I would anticipate that the computer would begin to learn to recognize some things about words that would let it eliminate some. For example, I imagine that the computer could probably tell, at least in many contexts, whether the word it is looking at should be a noun or a verb so it could eliminate all of the noun meanings and then just ask for a choice between the verb meanings or vice versa. This capability will depend on some logical processing that we haven't attempted to do yet, but that we do expect to do in the future.
QUESTION: Is it not correct that in this stage it looks up the word and tries to decide the meaning before any building of structure is attempted?

ANSWER: Yes.

QUESTION: Do you use now or do you envision using more limited dictionaries? If you are translating the Diesel Vehicular Engine Manual, you know that you are not going to use certain meanings of certain words, so you can just plug into the mechanical dictionary and when 'bearing' comes up you know that it doesn't mean someone who is 'overbearing.'

ANSWER: Yes. Of course, the dictionaries that we have built now have been solely with the purpose in mind of translating Church literature. So we've got 'baptism' and all the Church terms that you certainly wouldn't encounter in a Diesel Vehicular Engine Manual. If we were to undertake translating large amounts of Caterpillar Tractor materials, for example, we would probably build a separate subset dictionary that we would use for those materials. Also, if we decided that we wanted to translate a specific document, we could build a subset dictionary of just the words and word senses in that document.

It's almost lunchtime . . . are there any other questions?

QUESTION: What is being done as far as drawing any theoretical linguistic conclusions from some of the decisions that have to be made? For example, about the lexicon, one area that I can think of off-hand is the area of semantic features versus syntactic features. It seems like it ought to be possible to discover what features (i.e. stativity or factivity), which of those are semantic and which of those are syntactic? Which ones do you have to know to make up the sentences, and which do you have to know to arrive
at the right sememe or to derive basic units of meaning. Is anyone writing up any of these things because it seems like if this system is functional, it ought to be disseminated.

ANSWER: Of course, many of the distinctions and features that you mention are dealt with in Junction Grammar theory. Brother Lytle, would you like to respond a little more authoritatively?

LYTLE: First of all, we don't draw distinctions as you do between syntactic features and semantic features. Most of what you have classified as syntactic features are actually aspects of the structure of the Junction Grammar tree--definite, indefinite; generic, specific--things of that nature, are assimilated by the structure. Other features that you would classify as being semantic generally would be referred to as referential. In other words, they are part of the reference to something--animate, inanimate, human--this sort of thing. So there are some definite conclusions; unfortunately, we have been so busy developing the system that we haven't had time to write it all up. We have very good intentions though!

LUTHY: May I break this off? Sorry, it's about time for lunch. May I say, I hope the linguistic community at large will accept the kinds of things that will come from this. I know that some attempts have been made to publish some of this and have been rejected because the work was not "in the mainstream of current linguistic thought."
APPENDIX

The following tables give detailed figures for the evaluation discussed in this paper. Some significant differences will be noted between some of these figures and the figures presented in the Symposium. The discrepancies are primarily due to errors in computation engendered by the rush to get everything compiled in time. Nevertheless, the basic conclusion that the computer-assisted approach is potentially cost-effective still seems to be justified, especially in light of the project optimization figures given in Table 11.
### Table 1 - The Test Materials

<table>
<thead>
<tr>
<th>Source</th>
<th>Sentences</th>
<th>Words</th>
<th>Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;After Baptism What?&quot; by Mark E. Peterson</td>
<td>20</td>
<td>287</td>
<td>1650</td>
</tr>
<tr>
<td>&quot;Baptism After Dark&quot; by Kathie Troxler</td>
<td>18</td>
<td>307</td>
<td>1760</td>
</tr>
<tr>
<td>&quot;3208 Diesel Vehicular Engine&quot; by Caterpillar Tractor Co.</td>
<td>17</td>
<td>318</td>
<td>1716</td>
</tr>
<tr>
<td>TOTAL</td>
<td>79</td>
<td>1247</td>
<td>6832</td>
</tr>
</tbody>
</table>
Table 2 - Referential Disambiguation

<table>
<thead>
<tr>
<th>Text Required</th>
<th>Man-Minutes</th>
<th>CPU Seconds</th>
<th>Number of Interactions</th>
<th>Instances of Inadequacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>27.20</td>
<td>82.98</td>
<td>166</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($11.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>30.30</td>
<td>91.81</td>
<td>176</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($12.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>25.02</td>
<td>92.73</td>
<td>182</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($12.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>26.55</td>
<td>91.30</td>
<td>139</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($12.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>109.07</td>
<td>358.82</td>
<td>663</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($47.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVERAGE</td>
<td>27.27</td>
<td>89.70</td>
<td>166</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($11.95)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 - Syntactic Disambiguation

<table>
<thead>
<tr>
<th>Text Required</th>
<th>Man-Minutes</th>
<th>CPU Seconds</th>
<th>Number of Interactions</th>
<th>Sentences with no Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>2.40</td>
<td>11.27</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($1.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>4.43</td>
<td>15.42</td>
<td>36</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($2.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>2.85</td>
<td>11.74</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($1.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>6.90</td>
<td>14.62</td>
<td>47</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($1.94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>16.58</td>
<td>53.05</td>
<td>123</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($7.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVERAGE</td>
<td>4.15</td>
<td>13.26</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($1.76)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 4 - Transfer and Synthesis

<table>
<thead>
<tr>
<th>Text</th>
<th>Transfer CPU Seconds</th>
<th>Transfer CPU Seconds</th>
<th>Synthesis CPU Seconds</th>
<th>Synthesis CPU Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>5.46 ($ .71)</td>
<td>14.44 ($ 1.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>5.41 ($ .70)</td>
<td>16.29 ($ 2.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>4.46 ($ .58)</td>
<td>13.45 ($ 1.76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>6.50 ($ .84)</td>
<td>17.76 ($ 2.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>21.83 ($ 2.83)</td>
<td>61.94 ($ 8.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVERAGE</td>
<td>5.46 ($ .71)</td>
<td>15.49 ($ 2.03)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5 - Post Editing

<table>
<thead>
<tr>
<th>Text</th>
<th>Man-Minutes Required</th>
<th>Average Minutes per Sentence</th>
<th>CPU Seconds Required</th>
<th>CPU Seconds Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>136.85</td>
<td>6.84</td>
<td>78.29 ($10.42)</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>113.95</td>
<td>6.33</td>
<td>49.75 ($ 6.62)</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>67.72</td>
<td>2.82</td>
<td>37.80 ($ 5.02)</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>98.88</td>
<td>5.82</td>
<td>41.12 ($ 5.46)</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>417.40</td>
<td>21.81</td>
<td>206.96 ($27.52)</td>
<td></td>
</tr>
<tr>
<td>AVERAGE</td>
<td>104.35</td>
<td>5.45</td>
<td>51.74 ($ 6.88)</td>
<td></td>
</tr>
</tbody>
</table>
### Table 6 - "After Baptism What?"

<table>
<thead>
<tr>
<th>Step</th>
<th>Man-Minutes</th>
<th>CPU Seconds</th>
<th>Required</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential</td>
<td>27.20</td>
<td>82.98</td>
<td></td>
<td>($11.05)</td>
</tr>
<tr>
<td>Disambiguation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic</td>
<td>2.40</td>
<td>11.27</td>
<td></td>
<td>($ 1.50)</td>
</tr>
<tr>
<td>Disambiguation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remainder of Analysis</td>
<td>-</td>
<td>21.59</td>
<td></td>
<td>($ 2.86)</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>-</td>
<td>5.46</td>
<td></td>
<td>($ 0.71)</td>
</tr>
<tr>
<td>Synthesis</td>
<td>-</td>
<td>14.44</td>
<td></td>
<td>($ 1.89)</td>
</tr>
<tr>
<td>Post Editing</td>
<td>136.85</td>
<td>78.29</td>
<td></td>
<td>($10.42)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>166.45</td>
<td>214.03</td>
<td></td>
<td>($28.43)</td>
</tr>
</tbody>
</table>

### Table 7 - "Baptism After Dark"

<table>
<thead>
<tr>
<th>Step</th>
<th>Man-Minutes</th>
<th>CPU Seconds</th>
<th>Required</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential</td>
<td>30.30</td>
<td>91.81</td>
<td></td>
<td>($12.23)</td>
</tr>
<tr>
<td>Disambiguation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic</td>
<td>4.43</td>
<td>15.42</td>
<td></td>
<td>($ 2.05)</td>
</tr>
<tr>
<td>Disambiguation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remainder of Analysis</td>
<td>-</td>
<td>31.90</td>
<td></td>
<td>($ 4.24)</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>-</td>
<td>5.41</td>
<td></td>
<td>($ 0.70)</td>
</tr>
<tr>
<td>Synthesis</td>
<td>-</td>
<td>16.29</td>
<td></td>
<td>($ 2.14)</td>
</tr>
<tr>
<td>Post Editing</td>
<td>113.95</td>
<td>49.75</td>
<td></td>
<td>($ 6.62)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>148.68</td>
<td>210.58</td>
<td></td>
<td>($27.98)</td>
</tr>
</tbody>
</table>
Table 8 - "The Gospel of Luke"

<table>
<thead>
<tr>
<th>Step</th>
<th>Man-Minutes</th>
<th>CPU Seconds</th>
<th>Required</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential Disambiguation</td>
<td>25.02</td>
<td>92.73</td>
<td>($12.73)</td>
<td></td>
</tr>
<tr>
<td>Syntactic Disambiguation</td>
<td>2.85</td>
<td>11.76</td>
<td>($1.56)</td>
<td></td>
</tr>
<tr>
<td>Remainder of Analysis</td>
<td>-</td>
<td>18.53</td>
<td>($2.45)</td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>-</td>
<td>4.46</td>
<td>($0.58)</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>-</td>
<td>13.45</td>
<td>($1.76)</td>
<td></td>
</tr>
<tr>
<td>Post Editing</td>
<td>67.72</td>
<td>37.80</td>
<td>($5.02)</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>95.59</td>
<td>178.73</td>
<td>($23.72)</td>
<td></td>
</tr>
</tbody>
</table>

Table 9 - "3208 Diesel Vehicular Engine"

<table>
<thead>
<tr>
<th>Step</th>
<th>Man-Minutes</th>
<th>CPU Seconds</th>
<th>Required</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential Disambiguation</td>
<td>26.55</td>
<td>91.30</td>
<td>($12.16)</td>
<td></td>
</tr>
<tr>
<td>Syntactic Disambiguation</td>
<td>6.90</td>
<td>14.62</td>
<td>($1.94)</td>
<td></td>
</tr>
<tr>
<td>Remainder of Analysis</td>
<td>-</td>
<td>31.42</td>
<td>($4.16)</td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>-</td>
<td>6.50</td>
<td>($0.84)</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>-</td>
<td>17.76</td>
<td>($2.33)</td>
<td></td>
</tr>
<tr>
<td>Post Editing</td>
<td>98.88</td>
<td>41.12</td>
<td>($5.46)</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>132.33</td>
<td>202.72</td>
<td>($26.89)</td>
<td></td>
</tr>
</tbody>
</table>
### Table 10 - Totals by Step

<table>
<thead>
<tr>
<th>Step</th>
<th>Man-Minutes Required</th>
<th>CPU Seconds Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential Disambiguation</td>
<td>109.07</td>
<td>358.82</td>
</tr>
<tr>
<td>Syntactic Disambiguation</td>
<td>16.58</td>
<td>53.05</td>
</tr>
<tr>
<td>Remainder of Analysis</td>
<td>-</td>
<td>103.44</td>
</tr>
<tr>
<td>Transfer</td>
<td>-</td>
<td>21.83</td>
</tr>
<tr>
<td>Synthesis</td>
<td>-</td>
<td>61.94</td>
</tr>
<tr>
<td>Post Editing</td>
<td>417.40</td>
<td>206.96</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>543.05</td>
<td>806.04</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td>135.76</td>
<td>201.51</td>
</tr>
</tbody>
</table>

### Table 11 - Projected Optimization and Expansion

<table>
<thead>
<tr>
<th>Step</th>
<th>Man-Minutes Required</th>
<th>CPU Seconds Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>Change</td>
<td>Projected.</td>
</tr>
<tr>
<td>Average</td>
<td>Average</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Referential Disambiguation</td>
<td>27.27 +10% 30.00</td>
<td>89.70 -30% 62.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($11.95) ($ 8.37)</td>
</tr>
<tr>
<td>Syntactic Disambiguation</td>
<td>4.15 - 4.15</td>
<td>13.26 +10% 14.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($ 1.76) ($ 1.94)</td>
</tr>
<tr>
<td>Remainder of Analysis</td>
<td>25.86 +10% 28.45</td>
<td>($ 3.43) ($ 3.77)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($ 3.43) ($ 3.77)</td>
</tr>
<tr>
<td>Transfer</td>
<td>6.15 +90% 10.37</td>
<td>5.46 +90% 10.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($ 2.03) ($ 1.35)</td>
</tr>
<tr>
<td>Synthesis</td>
<td>15.49 -40% 9.29</td>
<td>15.49 -40% 9.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($ 2.03) ($ 1.22)</td>
</tr>
<tr>
<td>Post Editing</td>
<td>104.35 -70% 31.31</td>
<td>51.74 -70% 15.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($ 6.88) ($ 2.06)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>135.77 * 65.46</td>
<td>201.51 ** 141.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($26.76) ($18.71)</td>
</tr>
</tbody>
</table>

*Overall average change -48.2%.*

**Overall average change -69.9%.*
Table 12 - Projected Costs of Present System

<table>
<thead>
<tr>
<th>Text</th>
<th>Initial</th>
<th>Additional Language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Operator</td>
<td>Computer</td>
</tr>
<tr>
<td>1</td>
<td>$16.65</td>
<td>$28.43</td>
</tr>
<tr>
<td>2</td>
<td>$14.87</td>
<td>$27.98</td>
</tr>
<tr>
<td>3</td>
<td>$ 9.56</td>
<td>$26.89</td>
</tr>
<tr>
<td>4</td>
<td>$13.23</td>
<td>$26.89</td>
</tr>
<tr>
<td>TOTAL</td>
<td>$54.31</td>
<td>$107.02</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>$13.58</td>
<td>$26.76</td>
</tr>
</tbody>
</table>

Table 13 - Projected Costs after Optimization and Expansion

<table>
<thead>
<tr>
<th>Text</th>
<th>Initial</th>
<th>Additional Language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Operator</td>
<td>Computer</td>
</tr>
<tr>
<td>1</td>
<td>$ 8.03</td>
<td>$19.87</td>
</tr>
<tr>
<td>2</td>
<td>$ 7.17</td>
<td>$19.56</td>
</tr>
<tr>
<td>3</td>
<td>$ 4.61</td>
<td>$16.58</td>
</tr>
<tr>
<td>4</td>
<td>$ 6.38</td>
<td>$18.80</td>
</tr>
<tr>
<td>TOTAL</td>
<td>$26.19</td>
<td>$74.81</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>$ 6.55</td>
<td>$18.70</td>
</tr>
</tbody>
</table>
Table 14
Cost-Per-Language Comparison