BEYOND PAUSING AND LEXICAL LOOK-UP: AN ENRICHED DESCRIPTION
OF AUDITORY SENTENCE COMPREHENSION

Anne Cutler

Massachusetts Institute of Technology

Introduction

Sentence comprehension is like riding a bicycle—a feat far
easier performed than described. Understandably, most psycho-
linguistic attempts to characterize the comprehension process
have resorted to simplification. If—the tacit approach seems
to be—we can isolate for experimental investigation the most
basic components of comprehension, then on the basis of what we
learn we can construct at least an outline model of the process;
after all, any model can always be enriched.

This paper is an attempt at enrichment of the basic model.
In fact, it is more than that, since it will argue that even the
most basic model of auditory sentence comprehension must incor-
porate more than the minimal elements to which many descriptions
have hitherto been confined. Take for example the following
enumeration of the components of a sentence comprehension model:

> The listener must recognize the appropriate set of words
> in the flow of speech directed at him. This will re-
> quire him to find a match between some internal repre-
> sentation of the way each word sounds and properties
> of the incoming information about the speech waveform.
> Once a word is recognized, its meanings must be re-
> trieved. If there are several such meanings, the one
> appropriate to the current context must be selected
> and combined with the meanings of other words in order
> to form an interpretation of the entire sentence.
> Wherever the appropriate manner of combination de-
> pends upon syntactic properties of the sentence, such
> as word order or the groupings of words into phrases,
> these syntactic properties must be determined and put
> to use. (Wanner, 1973, pp. 166-167)

Sentence comprehension consists according to this account of
three stages—identification of word boundaries, lexical look-up,
and perception of syntactic structure, or parsing. No argument
can be raised with the inclusion of these components in the
model. However, it will be suggested that these three stages do
not suffice to characterise completely the process of sentence
understanding. Take for example the sentence

(1) Cassandra is a real genius.

This sentence may be spoken in a tone of reverence and admira-
tion, in which case it expresses praise of Cassandra. But it
may also be spoken in a quite different manner, with what is
known as an ironic intonation contour (nasalized, with heavy
stress on certain words). In this case it is far from express-
ing praise of Cassandra: quite the reverse. Ironic intonation
has the effect of producing a conveyed meaning which is the con-
verse of the literal meaning (Cutler, 1974); and there is no
doubt that a listener would apprehend this conveyed meaning, not
the literal meaning, of such an utterance; that is, he would com-
prehend that the speaker intended to say that Cassandra was any-
thing but a genius. It is difficult to see how this fact about
sentence comprehension can be encompassed by a model which in-
cludes only the operations of word identification, lexical
look-up, and parsing.

Similarly, suppose (2) to be spoken in such a way that the
primary stress of the sentence falls on "above":

(2) The above sentence was ironically intended.

While the proposition expressed by the sentence can be retrieved
by the three basic operations listed, an extra dimension has
been added which they would not retrieve--the implication that
some unspecified other sentence (or sentences) was not ironically
intended. If, on the other hand, the word "intended" receives
the primary stress, the implication changes: the reader will no
doubt agree that (2) now suggests that the intention was not
realized. Such variations are called changes in the focus of a
sentence; it is surely the case that implications of the kind
they express, like intonationally signalled irony, are computed
when a sentence like (2) is heard.

The following pages contain some specific suggestions about
the manner in which the basic parsing-plus-lexical-look-up model
of sentence comprehension needs to be enriched. One suggestion
will be the inclusion of a processing stage subsequent to the
establishment of the literal meaning of a sentence in which this
meaning may be revised. This stage will be called stage B;
accordingly, "stage A comprehension" will refer to the construc-
tion of a sentence's literal interpretation. Most of stage A
comprehension is accomplished before the utterance has been com-
pleted.

Making Use of Prosody

It will not have escaped the reader's notice that both
ironic intent and focus information are intimately bound up with
the prosodic, or suprasegmental, structure of a sentence: irony
can be signalled by nasalization, exaggeratedly slow speaking
rate, very heavy stress, or all of these, while the focus of a

sentence in general corresponds to the location of the main sen-
tence stress. In order to examine how these phenomena might be
registered in sentence processing, then, let us consider the
suprasegmental aspects of speech.

These consist, it is generally agreed, of variations in
pitch, stress and timing relations. Pitch variation is usually
arbitrarily defined as variation in the fundamental frequency of
a signal; the other two dimenisons, however, are not so easily
circumscribed. Timing, for example, is certainly expressed in
the relative durations of the various segments of the speech
wave; but silent intervals that can occur between segments also
play a role in determining the rhythmic pattern that is an im-
portant aspect of a sentence's timing. Stress, again, is a per-
ceptual feature which is manifested acoustically by an extremely
complex interaction between all suprasegmental aspects of the
utterance, not to mention segmental factors as well; vowel quali-
ty (formant structure) is a segmental phenomenon, but vowel re-
duction, namely a shift in formant frequency from the sounds
which fall in the outer portions oi the vowel quadrant towards
those of the center (/ /, schwa), is a phenomenon determined by
stress level in many languages. (For a comprehensive descrip-
tion of the nature of suprasegmental phenomena see Lehiste, 1970).

Some evidence that suprasegmental factors play an important
role in stage A comprehension of a spoken sentence has recently
been collected from studies using the phoneme-monitoring tech-
nique. This is a task in which subjects are asked to understand
a sentence and at the same time to listen within it for the
occurrence of a specified word-initial target sound, and to press
a button when they hear a word beginning with this sound. Re-
action time (RT) to the target phoneme in this task, it is argued,
is directly related to the difficulty of processing the sentence
at the time when the target phoneme occurs—RTs are lengthened,
for example, by the occurrence immediately prior to the target-
bearing item of a low frequency word (Foss, 1969) or of an
ambiguous item (Foss, 1970).

It was noticed in certain phoneme-monitoring studies that
RTs were faster when the target-bearing item itself was an "open
class" item (noun or verb) than when it was a "closed class"
item (for example, a preposition or conjunction). It is general-
ly the case that open class words carry a higher level of sen-
tence stress than do closed class words. Accordingly, Cutler
and Foss (in press) measured RTs to targets on open and closed
class items while manipulating stress level of the target-
bearing item independently. They found that RTs were signifi-
cantly faster to targets on stressed items, irrespective of
word class, whereas removal of the stress differences between
open and closed class items also removed the RT difference be-
tween them.

It might be argued that this result bears no great import,

since there are notable acoustic differences between stressed
and unstressed words. Stressed words are in the main longer in
duration and higher in pitch than unstressed words, and their
amplitude is somewhat greater; unstressed words, moreover,
generally undergo vowel reduction. Thus one might wish to ex-
plain the Cutler and Foss result in terms of acoustic advantages
of stressed items: heightened intelligibility facilitates iden-
tification of the phonemes, and location of a match for the
phonemic string in the mental lexicon, thus allowing faster
identification of the target phoneme in the required word-initial
position.

That this is not the whole story, however, is demonstrated
by a further experiment (reported in Cutler, 1975). In this study
a number of sentences were recorded in two intonation versions,
with the target-bearing item of each sentence receiving high
stress in one version and low stress in the other. Thus in (3),
in which the target phoneme is /d/, the target-bearing item
"dirt" receives high stress in (3a) and reduced stress in (3b):

    (3) a.   She managed to remove the dirt from the rug,
           but not the berry stains.

       b.   She managed to remove the dirt from the rug,
           but not from their clothes.

The stress assigned to the target item was determined, as can be
seen from these examples, by varying the endings of the sen-
tences to manipulate what is commonly called contrastive stress
(Bolinger, 1961). The point at which the two versions of each
sentence vary, however, occurs sufficiently later than the tar-
get for the response button to have been pressed by the time
that part of the sentence is heard by the subject.

In addition to these two versions of each sentence, a third
version, spoken in as neutral a tone as possible, was recorded.
In this third version the stress level of the target item was
intermediate, falling between the high- and low-stress versions.
The actual target-bearing words were then removed from the highl-
and low-stress versions of each sentence by tape-splicing, and
replaced by identical copies of the same target-bearing item
taken from the intermediate-stress version of the sentence. As
a result of this manipulation, the two experimental versions of
each sentence contained acoustically identical target-bearing
items. The two versions differed, however, in the intonation
contour which preceded them, one contour being consistent with
the occurrence of a high-stress item at the location of the
target-bearing word, the other being consistent with reduced
stress at that point.

If acoustic advantages of stressed items were solely re-
sponsible for the RT advantage of stressed target words in the
Cutler and Foss study, no difference would be expected between
the high- and low-stress versions of each sentence, since the

target-bearing items themselves were acoustically identical. In fact, however, significantly faster RTs were recorded for the items which occurred in high-stress position; i.e., the stressed words maintained their RT advantage, despite the fact that they had lost the advantages of greater intelligibility.

The sole difference between the high- and low-stress versions of each sentence in this experiment lay in the intonation contour preceding the target-bearing item. It must therefore be assumed that the RT difference reflects an effect of this contour variation. We must assume, in other words, that the subjects were making use of the suprasegmental information in such a way that their processing of the sentence at the point of occurrence of the target-bearing item was affected by whether that item was expected, on the basis of the intonation contour which preceded it, to carry high or low stress.

Since the RTs were faster if the item was expected to bear high stress, processing at that point of the sentence was apparently facilitated in some way by the expectation of stress. A reasonable interpretation is that particular attention has been directed to locations of highly stressed items. Moreover, the effect was produced in this case solely by manipulation of the preceding intonation contour, indicating that an active search for the locations of high stress proceeds in the form of a tracking of the intonation contour. A model of sentence comprehension which incorporates a search for highly stressed items is obviously more complex than a basic parsing-plus-lexical-look-up model. However, the evidence of these phoneme-monitoring studies compels us to expand the model in this way.

Semantic Focus

In the introductory section of this paper the notion of semantic focus was introduced. Sentences (4) and (5), in which upper case letters represent highly stressed items, differ on this dimension, the focus of the former being "Felicity", of the latter "caviar".

(4) FELICITY eats caviar for breakfast.

(5) Felicity eats CAVIAR for breakfast.

Jackendoff (1972) defines focus as the information in a sentence which is assumed by the speaker not to be shared by him and his audience. Halliday (1967) draws a distinction between "new" and "given" information, where "what is focal is 'new' information; not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse" (p. 204).

Our current problem is the extent to which this notion

is relevant to the description of sentence processing. Can it
be said, for example, that in order to understand a sentence it
is necessary to have identified its semantic focus? If a hearer
understands (4) and (5) as identical, can he be said to have
understood them?

The position taken here will be that he cannot. That is, it
will be argued that focus information constitutes an integral
part of the semantic representation constructed by the sentence
comprehension device. Successful comprehension necessarily in-
cludes the knowledge that (4) and (5) are different; further,
comprehension of (6) will entail that the hearer realizes that

(6)   Felicity eats CAVIAR for breakfast?

acceptable responses include (7) and (8), but not (9) or (10):

(7)   Yes, she like to indulge herself.

(8)   She's on a fish-only diet.

(9)   Did you think it was Samantha who did?

(10) No, for dinner.

There is evidence from at least one psycholinguistic study
that focussed items in a sentence are differentially represented
from non-focussed items shortly after the completion of compre-
hension. Hornby (1974) presented subjects with cleft and pseudo-
cleft sentences and required them to judge whether a picture pre-
sented for a brief interval beginning one second after presenta-
tion of the sentence accurately reflected the sentence's content.
He found that subjects were more likely to make errors with re-
spect to the noun phrase in the non-focussed part of the sentence
than with respect to the focussed noun phrase. For example, a
subject who had heard (11) would be more likely to respond

(11) It is the girl who is petting the dog.

"true" to a picture of a girl petting a cat than to a picture
of a boy petting a dog.

Suppose, however, that it could be shown that whether or
not an item is focussed affects the way it is processed during
Stage A comprehension. As we have seen, the focus of a sentence
and the location of that sentence's main stress coincide; would
it therefore be beyond reason to suggest that the active search
for the main sentence stress during sentence comprehension is in
fact a search for the sentence's focus?

If this is indeed the proper explanation for the stress ef-
fect, we would expect an effect of focussing an item in a sen-
tence analogous to the effect of assigning an item high stress.
That is, we should be able to demonstrate that phoneme-monitoring
RTs are accelerated if the target word is focussed.

The main problem with such an approach lies in the fact that
sentence focus and primary sentence stress coincide; stress

produces an effect on phoneme-monitoring RTs, therefore to de-
monstrate that focus produces an analogous effect it is necessary
to remove the confounding with stress, to keep item stress con-
stant whether or not the item is focussed. Thus, focus cannot
in this case be defined as the location of the main sentence
stress; instead, an alternative means of focussing a particular
item must be used. Among possible solutions is the use of the
cleft (e.g., 12) and pseudo-cleft (e.g., 13) constructions, which
are also considered to have a focussing effect (Jackendoff,
1972; Akmajian, 1970).

(12) It was cleft sentences that Portia refused
      to utter.

(13) The construction that Doris used most was the
      pseudo-cleft.

Two investigations provide evidence for an effect of focus
in phoneme-monitoring analogous to the effect of stress. In
the first (reported in Cutler, 1975), focus was manipulated by
the use of cleft and pseudo-cleft sentences describing simple
agent-action-object situations; RT to a target-bearing item
which was clefted was compared to RT to the same item in non-
clefted position. Since cleft and pseudo-cleft sentences can-
not be spoken naturally without assigning high stress to the
clefted item, the sentences were not spoken, but instead were
generated on a speech synthesizer, with which it was possible to
hold the fundamental frequency and amplitude inputs for a sen-
tence constant, and to use identical durational specifications
for each occurrence, clefted or non-clefted, of any individual
item. Thus acoustic invariance of the target-bearing item
across its various appearances was ensured. The results indi-
cated that focussed (clefted) target-bearing items indeed
elicited faster RTs than the same items when not focussed.

However, it is unfortunately also the case that varying the
focus of an item by means of clefting makes it no longer possible
to hold the item's position in the sentence constant. In the
sentences used in this experiment focussed items occurred at the
end of a clause more often than did non-focussed items. Phoneme-
monitoring latency is known to decrease towards the end of a
sentence or clause (Foss, 1969; Shields, McHugh & Martin, 1974;
this effect will be discussed in greater detail below). In two
pairs of items which differed in focus of the target but not in
the position of the target with relation to a clause boundary,
tests showed that the focussed member of the pair still elicited
faster RTs; nonetheless, these results are suggestive rather
than convincing.

A recent investigation by Jerry Fodor and myself, however,
attacked the problem differently. In this study the means by
which focus was manipulated was extra-sentential; a question was
asked, immediately prior to the sentence, which directed the

listener's attention to one or another part of the sentence.
Thus the problem of confounding focus with position of the tar-
get item within a clause was avoided, in that the syntax of the
sentence remained constant regardless of whether or not the
target-bearing item was focussed. It was also unnecessary to
control stress by using synthesized speech; the experimental
sentences were recorded exactly once, and each was spoken without
applying particularly high stress to any item.

In each sentence two words were designated as target-bearing
items, and two questions were formulated for each sentence, one
of which directed the attention of the listener to that part of
the sentence in which the first target-bearing item was located,
while the other focussed attention on the part of the sentence
containing the second target-bearing item. Thus, (14) could be
preceded by the target specifications /b/ or /d/, and by the

(14) The woman with the bag went into the dentist's

      office.

(15) Which woman was it that went into the office?

(16) Which office was it that the woman went into?
questions (15) and (16), which focus attention on the two target-
bearing items "bag" and "dentist's", respectively.

By means of tape-splicing, four materials sets were con-
structed; each experimental sentence occurred with a different
combination of target specification and preceding question in
each set. The splicing technique enabled the same recording of
the base sentences to be used in each set, so that each subject
heard an acoustically identical version of each sentence, re-
gardless of which target-question combination preceded it.

If non-suprasegmental cues to the semantically most central
portions of the sentence can be used in the same manner as supra-
segmental cues, then we would expect that focussing a word within
a sentence, by means of asking a question to which it provides
an answer, would facilitate RT to that word's initial phoneme.
For the example sentence (14), that is, we would predict that
RTs to the first target, /b/, would be faster if the subject had
heard (15) than if he had heard (16), whereas RT to the second
target, /d/, would be faster if the subject had heard (16) than
if he had heard (15). Since each subject heard the same recor-
ding of all experimental sentences, acoustic factors of course
cannot be invoked to explain RT differences.

Exactly the predicted interaction was found. Thus we have
shown that semantic focus can exercise an effect (on phoneme-
monitoring RT) prior to the completion of Stage A processing.
Moreover, by demonstrating an effect of focussing an item
analogous to the effect of assigning high stress, we have pro-
vided support for the notion that the strategic value to the
sentence processor of an active search for the locations of high

sentential stress is that the processor is thereby enabled to
direct attention to the location of the sentence focus, to the
semantically most central portion of the utterance. The greater
attention paid to the stressed and/or focussed elements is re-
flected in shorter phoneme-monitoring latencies.

Does it seem far-fetched to envisage the sentence processing
mechanism monitoring the intonation contour of an incoming sen-
tence as part of an active search for the sentence focus? Some
recent work (Allen & O'Shaughnessy, to appear) demonstrates that
the acoustic prerequisites for this view are certainly fulfilled.
Allen and O'Shaughnessy recorded a large number of sentences in
which various devices were used to indicate the sentence's focus,
including clefting, pseudo-clefting and preposed questions, and
then measured the fundamental frequency contours of these sen-
tences. They found that all methods of focussing produced re-
liable and similar effects on the pitch contour, with fundamen-
tal frequency accent falling in each case on the element which
was focussed. In other words, speakers produce fundamental
frequency cues to the semantically central elements of an utter-
ance—is it surprising that the listener seeks to make use of
these cues?

Presuppositions, Context and the Interpretation of Irony

The notion of focus is held to be accompanied by a corres-
ponding notion of presupposition (Jackendoff, 1972), the pre-
suppositions expressed by (4) and (5) being that somebody eats
caviar for breakfast, and that Felicity eats something for break-
fast, respectively. Jackendoff defines presupposition as the
information which the speaker assumes his audience to share with
him. Does the conclusion that the identification of sentence
focus comprises part of Stage A comprehension imply that the pre-
suppositions carried by a sentence are likewise identified during
this stage?

As with focus, there is evidence that the presuppositions of
a sentence are available to the hearer at least shortly after
comprehension has been completed. Offir (1973) tested subjects'
recognition memory of a sentence which they had heard embedded
in a short paragraph. She found that changes which had been made
in the sentence were more likely to be recognized if they affec-
ted the presuppositions carried by the sentence than if they did
not—even though the sentence was often more greatly changed in
the latter case. Hornby (1971) found that subjects asked to re-
call cleft and pseudo-cleft sentences may make mistakes in sur-
face structure, but are unlikely to make mistakes about what the
sentence presupposes and what it asserts.

There exists no evidence, however, that presuppositions of
a sentence are computed during Stage A comprehension. Note that

the suggestion of an active search for, and the direction of par-
ticular attention to, the location of the focus implies that
attention is directed away from the semantically less central
parts of the utterance. It will be argued that the computation
of presuppositions borne by these parts belongs, despite the
intimate relation of presupposition with sentence focus, more
properly with the interpretation of certain extra-sentential
factors.

Although there exists no experimental evidence to buttress
this argument, some circumstantial evidence can be called upon.
A great many sentences do not carry presuppositions at all,
whereas, in the broad sense of focus that has been used in this
discussion, every sentence has a focus; in every spoken sentence
there is a point at which the relative stress level is higher
than in the rest of the sentence, and this point will always
correspond with a semantically central portion of the message.
Thus although each utterance will contain parts which are less
central, these elements will not necessarily involve a presup-
position, and the apprehension that a particular part of a sen-
tence is less important does not entail that the sentence is
thereby understood to involve a presupposition.

Note that this discussion has been in at least one respect
greatly oversimplified. As Morgan (1969) has pointed out, there
are two distinct types of presupposition, sentential and lexical.
It is the former type that can be determined by the suprasegmen-
tal contour assigned to the sentence. The latter type is carried
by a lexical item, for instance the word "stop", as in (17):

(17) Have you stopped beating your husband?

None of the points raised above apply to the lexical presupposi-
tion; there is surely a case to be made for inclusion of this
type of presuppositional import in the entry allotted the parti-
cular item in the mental lexicon. The present discussion will
continue to confine itself to sentential presupposition.

The computation of presuppositions of a sentence will be
held to be similar to the computation of contextual effects on
the interpretation of a sentence. It is surely true that many
sentences cannot be said to be understood until they are under-
stood in context. The host of a noisy party, interpreting a
neighbor's utterance of

(18) I'm trying to sleep.

as a statement of fact rather than as a request to be quiet has
not successfully comprehended it. Similarly, the ironically
spoken (1) is misunderstood if it is taken as praise.

However the effects of irony and context on the interpre-
tation of a sentence take place at a level which is certainly
beyond Stage A. Ironic intonation, as we have seen, produces a
conveyed meaning which is the converse of the literal meaning,

1.e., it negates the proposition expressed by the literal read-
ing of the sentence. But the negation of a proposition cannot
be understood without the proposition itself being understood—
in other words, successful comprehension of the ironically in-
tended message is contingent upon successful comprehension of
the literal meaning of the sentence. Likewise, the context alone
does not lead to apprehension of the request expressed by (18),
despite the fact that, under the circumstances, the appearance
of a neighbor in pyjamas amounts to a message in itself; after
all, the hearer's reaction would be quite different if the utter-
ance were (19) or (20) . That is to say, the literal meaning of

(19) May I join the party?

(20) Do you realize the house is on fire?

the utterance must again be retrieved before the contextual in-
terpretation can be applied to yield the final interpretation.

The interpretation of irony, context and presuppositional
structure does not seem to comprise part of Stage A, the estab-
lishment of a sentence's literal meaning. To account for the
obvious effects of these factors on the understanding of utter-
ances, it is therefore necessary to enrich the sentence compre-
hension model by the inclusion of a stage, subsequent to the
establishment of the literal meaning of an utterance, in which
the literal meaning is embellished or revised in the light of
extra-sentential considerations—i.e., Stage B.

It is likely that Stage B consists of a number of different
operations. The effect of ironic intonation in reversing the
literal meaning of a sentence is presumably the result of a dif-
ferent sequence of operations from those producing the effect of
context on the interpretation of (20) , (However, note that
ironic effect can also be achieved by context: when two people
walk into an empty bar, the utterance (21) will be understood as

(21) Sure is lively here tonight.

ironic regardless of the intonation used.) The identification
of the presuppositions carried by the sentence, and the checking
of these against the hearer's knowledge and beliefs, result no
doubt from a different set of procedures again. It is therefore
quite possible that Stage B is not a unitary stage, but that
there are multiple independent serial stages which a basic seman-
tic representation passes through before the sentence comprehen-
sion device is completely finished with it. However, the point
of these brief remarks is just that comprehension cannot be con-
sidered to be complete once the literal meaning of the sentence
has been established. Revisions of this representation do take
place as a result of such factors as ironic intonation and the
influence of context, although it is not necessarily the case
that every sentence undergoes such revisions. (It is worthy of
note that one of the few instances in which such effects would

be almost if not entirely absent is the processing of isolated
sentences during a typical sentence comprehension experiment.)
The stage at which these revisions take place is subsequent to
lexical look-up and the establishment of syntactic structure,
but it must nevertheless be considered an integral part of the
model of the comprehension process.

Implications of the serial position effect

One of the most reliable findings in sentence comprehension
tasks using RT methodology is that stimuli in the later portions
of a sentence produce shorter response latencies than do stimuli
in the earlier portions. Thus, Foss (1969) reported longer
phoneme-monitoring RTs to targets occurring earlier in the sen-
tence than to targets occurring later, a result also found by
Shields, McHugh and Martin (1974). Holmes and Forster (1970)
found that clicks in the second half of a sentence were detected
faster than clicks in the first half. All of these writers dis-
cussed the serial position effect in terms of facilitated pro-
cessing towards the end of the sentence; Foss successfully dis-
posed of three possible objections: that the effect merely
reflects the subjects' lower criterion for response later in the
sentence; that it reflects differential occurrence of target
items with relation to surface structure phrase boundaries; and
that it results from a reduction in the number of possible
structural continuations following target items in the later
part of the sentence.

A further possible explanation for the serial position ef-
fect arises from the phenomena discussed earlier in this paper.
In a sentence in which no particularly heavy stress is applied
to any element, the point at which the stress level will be
highest will lie at or near the end of the sentence. The expec-
ted semantic effects accompany the suprasegmental: in the un-
marked case, given information in a sentence precedes new infor-
mation (Halliday, 1967). An experimental result which supports
this explanation is that of Hornby (1972); in a task involving
subjects[1] judgements of various surface structure expressions of
simple agent-action-object sequences, Hornby found that active
sentences in which the agent was heavily stressed produced re-
sults similar to those produced by cleft-agent sentences, where-
as active sentences in which no heavy stress was applied were
treated similarly to cleft-object sentences.

It is reasonable to assume that the serial position effect
in part reflects the fact that the point of highest stress, the
semantic focus, to which the sentence processor seeks to direct
attention, lies unless otherwise determined in the last part of
the sentence. Where sentential stress was not specifically
manipulated, experiments in which the serial position effect was

reported can be presumed to have used sentences in which the
focus occurred at the end of the sentence.

It is unlikely, however, that a focus explanation can ac-
count entirely for the serial position effect. The phoneme-
monitoring experiment reported earlier, in which focus was
manipulated by means of preposed questions, exhibited a strong
serial position effect despite explicit extra-sentential cues
to the semantic focus. While focus exerted the predicted ef-
fect, RTs to the later-occurring target were overall faster than
to the earlier-occurring target. It would appear that there is
some further component of the serial position effect besides
the search for the semantic focus.

In what way might completed processing of earlier parts of
a sentence facilitate processing of later parts? There exists
a considerable body of evidence that just such facilitation does
not happen when one might on common-sense grounds expect it to
be particularly useful, namely in the resolution of lexical am-
biguity. Phoneme-monitoring RT is lengthened immediately fol-
lowing the occurrence of an ambiguous lexical item (Foss, 1970),
and a preceding context which renders only one reading of the
ambiguous item acceptable does not remove the ambiguity effect
(Foss & Jenkins, 1973; Cutler & Foss, 1974).

This effect presumably reflects retrieval from the mental
lexicon of all the readings listed for the particular item.
Although it may be impossible to use biasing context to limit
retrieval to only the relevant reading, another kind of facili-
tation from preceding context is conceivable. Imagine, for
example, that one of the operations of the sentence processor is
the construction of hypotheses about the content of the incom-
ing utterance (cf. Forster, 1975). The strategic value of such
hypotheses might lie merely in reducing uncertainty about the
incoming message, given the degraded nature of the signal upon
which the speech processor operates. After all, the most sali-
ent and at the same time most amazing fact about speech compre-
hension is its speed, even though in real speech situations
sounds, syllables, even words are missing from the spoken reali-
zation of the message, the sounds that are present may be dis-
torted or compressed, and the whole signal is, on top of this,
received often through considerable extraneous noise. The de-
termination of word boundaries in this degraded input, and hence
of the strings which are to be sought in the mental lexicon,
might reasonably be considered a highly tentative operation;
the retrieval of a reading which matched the semantic hypothesis
constructed for that part of the sentence would presumably en-
courage the processor to accept that interpretation and to re-
frain from trying alternative patterns of segmentation.

Further, the semantic hypotheses could be more specifically
useful in expediting the choice between alternative readings
retrieved from the lexicon for ambiguous items. In the

phoneme-monitoring experiments on ambiguity, biasing context exercised a facilitating effect on RT to targets in both ambiguous and unambiguous control sentences. The context effect did not remove the RT lengthening due to ambiguity, which we take to be a lexical effect, but it did slightly reduce it. On the view proposed here, this reduction would be due to an effect at the point of choice between the alternate readings.

The serial position effect can be considered to reflect the operation of semantic hypotheses in the following way: in the early part of the sentence the constructed hypotheses may often prove wrong and need to be revised, thus adding to the momentary processing load and leading to slower RTs in phoneme-monitoring and click detection tasks. As the sentence is decoded, however, the hypotheses are more likely to be correct and to need less revision, so that less of the available processing capacity is taken up with hypothesis formulation and testing, and the detection of phoneme or click targets can be accomplished more rapidly.

Strong support is provided for this view in the work of Forster on the effects of plausibility (reported elsewhere in this volume). Subjects asked to judge whether or not a string of words is an acceptable sentence of English produce longer RTs to sentences which are implausible in content—though completely grammatical and meaningful—than to sentences which are plausible. An implausible input will on the average generate more hypotheses before the correct one is hit upon than will a plausible input.

That phoneme-monitoring is sensitive to the effects of plausibility has recently been demonstrated by Morton and Long (1976), who found that target-bearing items which had a lower probability of occurrence in a particular context elicited longer RTs than items with a higher probability of occurrence in that context. If the above view of the serial position effect is correct, it might be expected that RT to phoneme (or click) targets in implausible sentences would show a lesser reduction towards the end of the sentence than RT to targets in plausible sentences. This hypothesis awaits experimental investigation.

The testing of semantic hypotheses during Stage A comprehension is independent of the similar notion of testing of syntactic hypotheses suggested for example by Bever (1970), although the two suggestions are of course compatible. Forster and Olbrei (1973) have shown that semantic variables do not simplify syntactic processing; but it is conceivable that the reverse may be the case.

Finally, a conception of the sentence comprehension process which involves the formulation and testing of hypotheses about the content of the input is highly compatible with the notion put forward above: that an active search is undertaken for the semantically most central parts of the sentence. These elements

would be of more use than any other in the construction of the correct hypothesis; thus it makes eminent sense for a hypothesizing processor to search for them.

Conclusion

It is a truism to state that sentence comprehension is an extremely complex process. Nonetheless, psycholinguists are all too prone to lose sight of this complexity. A simplified conception of the components of a comprehension model may seem to he desirable as a basis upon which research hypotheses may be formulated. But an adequate model of parsing and lexical look-up simply does not constitute an adequate model of auditory sentence comprehension.

The research reviewed here has shown that the model needs also to take into account the processing of the prosodic structure of an utterance, which takes at least in part the form of an active search for the location of the main sentence stress. We have seen that this effect appears to reflect the coincidence of main sentence stress and sentence focus; i.e., that the search for the primary stress is in fact a search for the semantic focus. It has been suggested that the comprehension of spoken sentences involves the construction and testing of hypotheses about the content of the input, and that the location of a sentence's semantically most central portion is actively sought in order to facilitate the construction of the correct hypothesis, These phenomena properly belong in even the most basic model of auditory sentence comprehension.

Further, it has been noted that the proper understanding of certain sentences must be based on a semantic representation which is not identical with the literal meaning of the sentence, and that some sentences bear presuppositions which are demonstrably available to the listener once the sentence has been comprehended. It was suggested that the sentence comprehension model be enriched by the addition of one or more stages subsequent to the identification of the literal meaning in which transformations of this meaning on the basis of various factors, some of them extra-sentential, may be accomplished. It should be noted that this latter type of enrichment of the model is not confined to the comprehension of speech, but applies also to reading.

A description of auditory sentence comprehension enriched in the directions suggested in this paper may still be a far from complete one, deficient in countless aspects. But it will certainly be closer to the truth than the unimproved model which preceded it.

References

Akamjian, A. (1970) On deriving cleft sentences from pseudo-cleft sentences. Linguistic Inquiry, 1, 149-168.

Allen, J. and O'Shaughnessy, D. (To appear) The effect of focus-shifting transformations on fundamental frequency contours.

Bever, T.G. (1970) The cognitive basis for linguistic structures. In Hayes, J.R. (Ed.), Cognition and the Development of Language, New York: Wiley.

Bolinger, D.L. (1961) Contrastive accent and contrastive stress. Language, 37, 83-96.

Cutler, A. (1974) On saying what you mean without meaning what you say. Papers from the Tenth Regional Meeting, Chicago Linguistic Society, 117-127.

Cutler, A. (1975) Sentence Stress and Sentence Comprehension. Unpublished Ph.D. dissertation, University of Texas.

Cutler, A. and Foss, D.J. (1974) Comprehension of ambiguous sentences: the locus of context effects. Paper presented to the 46th annual meeting, Midwestern Psychological Association, Chicago.

Cutler, A. and Foss, D.J. (In press) On the role of sentence stress in sentence processing. Language and Speech.

Forster, K.I. (1975) The role of semantic hypotheses in sentence processing. In Bresson, F. and Mehler, J. (Eds.), Current Approaches to Problems in Psycholinguistics, Paris, CNRS.

Forster, K.I. and Olbrei, I. (1973) Semantic heuristics and syntactic analysis. Cognition, 2, 319-347.

Foss, D.J. (1969) Decision processes during sentence comprehension: effects of lexical item difficulty and position upon decision times. Journal of Verbal Learning and Verbal Behavior, 8, 457-462.

Foss, D.J. (1970) Some effects of ambiguity upon sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 9, 699-706.

Foss, D.J. and Jenkins, C.M. (1973) Some effects of context on the comprehension of ambiguous sentences. Journal of Verbal Learning and Verbal Behavior, 12, 577-589.

Halliday, M.A.K. (1967) Notes on transitivity and theme in English, Part 2. Journal of Linguistics, 3, 199-244.

Holmes, V.M. and Forster, K.I. (1970) Detection of extraneous signals during sentence recognition. Perception & Psychophysics, 7, 297-301.

Hornby, P.A. (1971) The role of topic-comment in the recall of cleft and pseudo-cleft sentences. Papers from the Seventh Regional Meeting, Chicago Linguistic Society, 445-453.

Hornby, P.A. (1972) The psychological subject and predicate. Cognitive Psychology, 3, 632-642.

Hornby, P.A. (1974) Surface structure and presupposition. Journal of Verbal Learning and Verbal Behavior, 13, 530-538.

Jackendoff, R.S. (1972) Semantic Interpretation in Generative Grammar. Cambridge, Ma.: MIT Press.

Lehiste, I. (1970) Suprasegmentals. Cambridge, Ma.: MIT Press.

Morgan, J.L. (1969) On the treatment of presupposition in transformational grammar. Papers from the Fifth Regional Meeting, Chicago Linguistic Society, 167-177.

Morton, J. and Long, J. (1976) The effect of word transitional probability on phoneme identification. Journal of Verbal Learning and Verbal Behavior, 15, 43-51.

Offir, C. (1973) Recognition memory for presuppositions of relative clause sentences. Journal of Verbal Learning and Verbal Behavior, 12, 636-643.

Shields, J.L., McHugh, A. and Martin, J.G. (1974) Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. Journal of Experimental Psychology, 102, 250-255.

Wanner, E. (1973) Do we understand sentences from the outside-in or from the inside-out? In Haugen, E. and Bloomfield, M. (Eds.), Language as a Human Problem, New York: Norton, 165-185.