

Chapter 1

*A Survey of Studies in Sentence Perception: 1970-1976**

Willem J. M. Levelt

1. AIMS

'The fundamental problem in psycholinguistics is simple to formulate: what happens when we understand sentences?' (Johnson-Laird, 143). 'The problem of when and how a sentence is understood is, in my view, the central problem of experimental psycholinguistics' (Gough, 109). 'How do we understand the relevant meaning of sentences used in ordinary contexts? This is a central problem in psychology and a primary preoccupation of the psycholinguist' (Tanenhaus *et al.* 244). These and many similar statements in the literature place the problem of sentence understanding at the very centre of psycholinguistics. The construction of a theory of sentence understanding is apparently expected to have spin-offs in various directions: the theory of language production, of language acquisition; it might even benefit cognitive psychology as a whole.

What, then, is the aim of a theory of sentence understanding? At this point authors are much less in agreement, or at least less explicit. In the early sixties the more or less implicit aim of constructing a theory of sentence understanding was to demonstrate the 'psychological reality' of linguistic notions, especially those developed in transformational linguistics. It is understandable in that light that such a theory was expected to stimulate developments in other areas of psycholinguistics as well. It is still an explicit purpose of many studies in language understanding to find out whether linguistic units are also relevant units in language processing. It has become abundantly clear, however, that proving 'psychological reality' of linguistic units and structures is too limited an aim for a theory of sentence understanding. Fodor *et al.* (80) convincingly show that reality studies have had only limited success. There is no doubt that such linguistic units and structures as words, constituents, and clauses show up in various experimental results; but generative rules, especially transformations, do not characterize processes of understanding. A major aim for a theory of

* The reader is referred to page 74 for a table of contents of this chapter.

sentence understanding is, therefore, not so much to validate linguistic structures as to explain how such structures are created by the language user. Fodor *et al.* call these structures *internal representations*, and they identify these with what linguists call *structural descriptions* (cf. 80, p. 21). According to this view, the purpose of a theory of sentence understanding is to explain how the hearer uses his knowledge of the language to encode sentences in terms of linguistic structural descriptions. Others describe the end product of understanding in non-linguistic terms. Schank (226), for instance, puts it this way: 'a parser should associate a linguistic input with what we will call a *conceptual structure*. A conceptual structure consists of concepts and the relationship between them'. The end-term, therefore, is a non-linguistic object. The aim of the theory is to explain how a linguistic object (a text, sentence, etc.) is mapped onto a non-linguistic object (a conceptual structure).

Many shades can be found between these two extremes (see, for instance, 102), and one very common claim is that deriving a structural description is a first step in deriving a conceptual structure. Varieties of this claim will be discussed and criticized in section 4. Here, it is only important to notice that the theoretical aims are much broader in the latter framework than in the former. One is more inclined to look into the role of context, especially of non-linguistic context within the conceptual framework. Also, the role of encyclopaedic knowledge in sentence comprehension is generally recognized and given attention: one is strongly inclined to hypothesize top-down and concept-driven parsing procedures. The former, more linguistic approach, being more limited in its aims, also tends to be more local. But in this way it has led to thorough analyses of syntactic and bottom-up procedures in sentence understanding.

Apart from this range of theoretical aims, authors vary in the degree to which they stress the importance of analysing the interaction between speaker and listener in the study of sentence comprehension. Hörmann (135) defines understanding as 'making sense at a point intended by the speaker'. How the listener manages to find out what the speaker's intentions are has increasingly become a more central aim for many researchers in recent years. Speech-act theory has captured the imagination of experimental psychologists. Clark (this volume) puts it this way: 'comprehension is conceived to be the process by which people arrive at the interpretation the speaker *intended* them to grasp for that utterance in that context'. And though the inference of such intentions is hard to study experimentally, Marslen-Wilson (186) explicitly defines it as an aim for the study of sentence perception: 'The normal listener... is a participant in a social event, trying to interpret the utterances he hears in terms of the communicative intentions of the speakers in question. Even if we cannot fully reproduce this situation under laboratory conditions, we can still try to capture, in our investigations of sentence perception, its essential goal-directed and dynamic properties'.

So, in summary, the tasks which students of sentence perception have set themselves are at least threefold; let us call them *linguistic*, *conceptual*, and *communicative*. It is obvious that these aims are not mutually exclusive, and as a matter of fact it is hard to find studies that are pure cases of any of these three approaches, as will become clear in the course of this review.

What, then, is the aim of this review? Since the pioneering work of the sixties, the experimental study of sentence perception has had an explosive development in the seventies, which led to such an extensive literature that comprehensive reviews are available only for particular subfields and are mostly to be found in unpublished doctoral dissertations. It therefore seemed useful to collect a major part of the experimental literature from 1970 to 1976 and to analyse it with respect to method, results, and theory. In order not to be inundated by our material, it was necessary to place some severe restrictions on the review. Firstly, except for a few important manuscripts, only published material was taken into account. Secondly, the review ends at 31 December, 1976. Though many new findings about active research topics have been made since the end of 1976, it was not feasible to expect comprehensiveness of treatment for the first months of 1977; moreover, one has to stop somewhere. Thirdly, we have been strict on limiting ourselves to studies of sentence perception. Thus, on the one hand we have not reviewed studies of text comprehension, word recognition, word verification, or phoneme perception. On the other hand, studies of sentence memory, or sentence production have also been left out. Fourthly, we have left out the whole neurophysiological and aphasia literature. Fifthly, we have not incorporated the massive literature on reading. Sixthly, we have been forced to mention only in passing many studies—but such is the fate of all reviewers. We have been especially short on issues which are treated in other chapters of this volume. Finally, some omissions have occurred in spite of the effort to be comprehensive: the meshes of our net have apparently been too large here and there. But it may also be noted in this respect, that one had to be realistic and stop somewhere, apologizing to the fish that escaped.

The review is organized in the following way. Section 2 deals with experimental procedures in the study of sentence perception, because it became more and more clear to us that thorough analyses of experimental tasks and dependent measures are a necessary requirement for evaluating experimental results and theories. Section 3 deals with the independent variables, the linguistic and non-linguistic factors which have been manipulated to study their effects on sentence perception. It is especially here that variation in theoretical aims becomes apparent. Section 4, finally, is fully devoted to these theoretical issues.

2. TASKS AND DEPENDENT VARIABLES

Psycholinguists have been highly inventive in the development of experimental procedures (for a recent review see Olson and Clark, 2005). The variety of experimental tasks which subjects have had to perform in studies of language perception is remarkable: click localization, dichotic switch localization, recall, completion, verification, paraphrasing, translation, judging comprehensibility or grammaticality, phoneme-monitoring, shadowing, etc. This productivity has its drawbacks, however. One cannot rid oneself of the impression that techniques are trademarks for research groups. This is especially clear in cases where different techniques are used to study basically the same theoretical issue. The issue of perceptual segmentation, for instance, has led one group (around Bever) to use click-localization tasks, another group (around Foss) to use phoneme-monitoring, a third group (around Wanner) to use a transient memory-load procedure, etc. Studies in which different methods are applied to the same material are very limited and tend to be rather inconclusive. As a consequence results obtained by different techniques have until now been less than cumulative. In this section procedures will be categorized into two major groups. The first group involves procedures where measurement takes place during reception of the stimulus, i.e. during the presentation to the subject of a clause, sentence, etc. These techniques will be called *simultaneous measurement*. One example is measuring pupil diameter while the subject listens to a sentence. The second group consists of procedures where measurement takes place after presentation of the stimulus. This will be denoted as *successive measurement*. An example is immediate recall of the stimulus. This distinction is important, since only in the first case can one be sure that results arise as a consequence of the input process, i.e. are strictly perceptual. In the latter case reconstruction activities, and therefore factors such as response bias, may codetermine the results. The results are perceptual only in the broad sense that they concern the way the subject reacts to a perceptual stimulus, and it is much harder, although not necessarily impossible, to determine in which phase of stimulus processing the experimental variable is effective.

It is quite simple to partition most techniques into these two classes, but a further qualification has to be made for a particular set of methods, those which can be called '*subsidiary task methods*'. Here, the subject has as a main task to interpret the verbal material, and as a subsidiary task to attend to something else in the mean time. Examples are click-localization and phoneme-monitoring experiments. Though the subsidiary stimulus is given during input of the test-stimulus, measurement can be either simultaneous or successive. Phoneme-monitoring is simultaneous since an immediate reaction to the subsidiary stimulus (the phoneme) is required. Click localization is successive, since the subject's response is given only after the full sentence has been presented. Subsidiary methods, therefore, fall within both main groups.

2.1 Simultaneous measurement

Three major ways of simultaneous measurement are in use: subsidiary task methods (which in simultaneous measurement are always monitoring methods), shadowing, and psychophysiological measurement.

2.1.1 Subsidiary task methods

During presentation of a word list, sentence, or text, the subject has to attend to the stimulus and at the same time to monitor for a particular test stimulus. (It is often only assumed that the subject attends to the sentence or text, without explicitly instructing him to do so.) This test stimulus can either be part of the verbal material, such as a particular syllable (191,222), word (1,249), or phoneme, or it may be an additional stimulus such as a light flash (272) or a click. The subject's task is always to react immediately to the test stimulus; reaction time is the dependent variable. Here we will give some attention to phoneme-monitoring and click-monitoring.

(i) *Phoneme-monitoring.* This technique was introduced by Foss and Lynch (92) in a study where subjects had to listen to a self-embedding or right-branching sentence which was followed by a comprehension test. Apart from understanding the sentence, subjects had the subsidiary task of pressing a reaction key as soon as they heard a word with /b/ as the initial phoneme. The authors hypothesized the existence of a fixed-capacity decision mechanism (something like a working memory) where different decision-making processes have to share time. Difficulty of decision with respect to syntactic processing will therefore delay reaction to the test phoneme. Such difficulty was predicted for certain positions in self-embedded sentences. The authors found longer RTs for self-embedding than for right-branching sentences, provided the target phoneme appeared rather late in the sentence. Several phoneme-monitoring studies by Foss and his colleagues have followed this initial study (36,37,84,87, 88, 91, 117, 118, 119, 120, see also 157). At this place we shall only discuss some issues related to the technique itself; the theoretical content of these studies will be treated in sections 3 and 4. A major issue is whether phoneme detection can be considered as an independent process, only competing for limited capacity. This was the original assumption, but after Savin and Beyer's (222) finding that monitoring for a syllable is quicker than monitoring for its initial phoneme, Foss and Swinney (93) shifted to the assumption that the phoneme is identified after first analysing the word, i.e. word or syllable recognition is a condition for phoneme detection. Mortan and Long (201) show that this is not a necessary conclusion: phoneme detection can be relatively slow because of greater response inertia. But still, phoneme-monitoring reaction times are also sensitive to factors affecting recognition of the word to which the phoneme belongs. One such factor is transitional probability between the preceding word and the target

word. Morton and Long showed that this factor alone could induce a 70 msec variation in phoneme-detection time. This finding in itself necessitates caution in interpreting phoneme-monitoring results in all cases where this factor has not been controlled. It also sheds new light on experiments where the frequency of the preceding word affected monitoring reaction times (36,87). If word recognition is a condition for phoneme detection, the target word's frequency itself should also be an important factor. The fact that several studies use different target words (with the same target phoneme) at different test positions in the sentence, makes this a major uncontrolled factor. There is not a single study known to this author which measured the effect of variations in target-word frequency.

A further, little understood, factor, which is probably a context variable that indirectly affects phoneme detection (through target-word identification) is the position of the target word in the sentence. Three studies (37, 87, 117) show relatively short RTs where the target word occurs relatively late in the sentence as opposed to early; two other studies (92, 120) show the inverse effect. Various factors can be involved here: transition probability, syntactic category of preceding word (see 37), increasing expectation, increasing processing load, response bias, intonational level, etc. (see especially 67), and rhythmic structure of the sentence (190, 233). Some of these factors are studied in a word-monitoring study by Marslen-Wilson and Tyler (187).

A final variable which has not been taken into account is the acoustical shape and environment of the phoneme. It is certain that in nearly all studies the target phoneme is not acoustically identical in different test positions. Moreover, its acoustical environment is also never controlled. Thus, variations in pitch, coarticulations, loudness, and so forth make it impossible to maintain that phoneme-monitoring studies provide the listener with physically constant test stimuli. The present dilemma for the phoneme-monitoring technique can be stated briefly as follows: either phoneme detection is an independent process competing for limited processing capacity, or phoneme detection is dependent on target-word identification and analysis. In the case of the former (and rather unlikely) alternative, it is essential to keep the physical stimulus and its immediate environment constant. This is hardly possible with phonemes as stimuli. In the second (more likely) case, one is essentially measuring word recognition. It is not only important, then, to control such nuisance variables as word frequency and transitional probability, but also to decide whether it would not make more sense to replace the indirect phoneme-monitoring by the more direct word-monitoring. The answer to that question will depend on the effectiveness with which one can control semantic and syntactic expectations in word-monitoring paradigms.

(ii) *Click-monitoring*. The subject presses or releases a key after detecting a click during presentation of a sentence. Abrams and Bever (3) introduced this

method with an argument similar to the one used by Foss and Lynch: detection of click and processing of the sentence are in competition for limited attentional capacity. If understanding is heavily loading the system, then detection will be relatively impaired. In retrospect we see that this initial study was merely exploratory; it was beset with experimental shortcomings (see 241), some of which also affected Holmes and Forster's study (130). (This latter paper seemed to show that RTs to clicks are relatively long in the midst of a clause.) The study by Streeter and Bever (241) controlled for relative frequency, transitional probability, acoustic shape, etc. A major finding was quicker RTs to clicks at the beginning of a clause than at the end, in correspondence with the attentional theory. An equivalent word-monitoring task tended to give the inverse effect: reactions to end-of-clause words were relatively quick.

The authors' interpretation is that the word, as an integrated part of the clause, is receiving attention contrary to the click. This conclusion makes it urgent to use the Streeter and Bever sentences in a phoneme-monitoring task. If phoneme detection is dependent on target-word identification and analysis (which as we have seen is rather likely), phoneme-monitoring should follow the word-monitoring results and not the click-monitoring results. Such a finding would be in good agreement with the close parallelism which Marslen-Wilson and Tyler (187) found between word-monitoring and rhyme-monitoring.

Bever and Hurtig (24) compared the Streeter and Bever (241) RT findings for supraliminal clicks, with detection for clicks at threshold. This is an essentially successive measurement situation, and response bias resulting from processing the whole sentence could affect the subject's yes/no decision which is given after the sentence. Though the results are compatible with those of Streeter and Bever, they should be interpreted with caution: inverted speech was used to control for acoustic masking by the sentence, a procedure based on the erroneous experimental assumption that forward-masking equals backward-masking. Click-monitoring is simultaneous measurement, as opposed to click localization (see section 2.2.1) and liminal click detection. Moreover, it is not beleaguered with most of the phoneme-monitoring problems. It is therefore surprising that click-monitoring has not been more widely used. Flores d'Arcais' (77) contribution to this volume gives further evidence of its usefulness.

2.1.2 Shadowing

Shadowing, i.e. the subject's repeating an acoustically presented sentence or text, insofar as possible in an on-line fashion, has been highly popular in former years. However, only a few studies since 1970 have made use of the technique. Rosenberg and Jarvella (216, 217) showed an effect of semantic integration of sentences on shadowing latency under noise conditions. Shadowing of word lists is used by Treisman, Squire, and Green (250) in order to test findings by Lewis (176), Garrett (101), and MacKay (182), regarding the semantic effects of

words heard in the unattended ear on the interpretation of words heard in the attended ears. Treisman *et al.* (250) cast doubt on models postulating full lexical analysis of unattended words. Darwin (69) extends this work to sentences in order to study the effect of prosody. Marslen-Wilson (184,185) used shadowing in order to demonstrate that so-called 'close-shadowers', who performed with a delay of about one syllable, were highly sensitive to syntactic and semantic aspects of the text. This demonstrates the immediacy of the syntactic and semantic processes during sentence processing. The origin of delays or errors in shadowing is not always self-evident. They can be due to a strictly perceptual failure or to complex difficulties in response selection.

2.1.3 Psychophysiological measurement

Psychophysiological studies of sentence processing are very much at an exploratory stage. Wright and Kahneman (273) studied pupil size as a measure of 'mental load' in tasks where the subject listened to a sentence, recalled it, or answered questions about it. Though results clearly indicated that effort was positively correlated with pupil size, no effects of phrase structure were found. It should be noted, however, that the experimental sentences, though complex, were essentially of a one-clause type. Two-clause sentences were used in a study by Abrams and Bever (3) in which the listener received slight electric shocks at critical points in the sentence. Galvanic skin responses were measured and turned out to be relatively strong for shocks occurring at the end of a clause. Unlike the pupil-size method, which does not interfere with the normal understanding process, the shock procedure is likely to divert the subject's attention from the syntactic material; it is probably the size of this orientation reaction which is reflected in GSR. This reviewer is unfamiliar with any follow-up study of these interesting beginnings.

2.2 Successive measurement

Studies on language perception using successive measurement outnumber, by a factor of about six, those using simultaneous measurement. It is impossible to review all the technical variations developed by inventive researchers. We will limit ourselves to the major types: subsidiary tasks, recall, recognition, paraphrasing, verification and question answering.

2.2.1 Subsidiary task methods

Apart from attending to the sentence, the subject has to attend to other simultaneously occurring events, but he is required to report on these only after the verbal stimulus has disappeared. The major method here is click localization; two other techniques are dichotic switch localization, and transient memory-load measurement.

(i) *Click localization.* Ladefoged and Broadbent (166) made subjects listen to a spoken message. The subjects had difficulty in locating the position of a short burst of noise (a 'click') which was presented during the message. The authors used the technique to determine perceptual units in the message. The subject's localization judgment would depend on what for him would have been a perceptual unit. In the psycholinguistic experimentation of the sixties, there was a real hunt for methods which could establish the 'psychological reality' of linguistic entities. The click procedure was discovered (79,100) as an almost god-given measure to prove the perceptual reality of constituents, and/or clauses. Everybody's imagination was caught when it was observed that subjects were better at locating clicks presented in major constituent boundaries (79,104) than at locating those presented just before or after such boundaries. Moreover, mislocations in the latter cases were more often than not in the direction of the boundaries. In short, boundaries 'attracted' clicks. At this point we will limit ourselves to a discussion of some of the technical questions which arose after the first euphoria subsided.

Let us recapitulate the original purpose of the click studies. They should show (a) the 'psychological reality' of certain linguistic units and (b) that this reality is strictly perceptual, i.e. that these units have their effect during input of the sentence.

With respect to (a), the main experimental problem was to disentangle the linguistic variable (i.e. the clause-boundary position) from other, often covarying, factors. The following covarying factors have been studied for their possible effects on click mislocation: serial position of the click in the sentence (21, 209), intonational pattern (210), and transitional probability (26). All these factors affected click localization, but for each of these factors it could be shown that the clause-boundary effect can occur if the factor is neutralized (see 21, and 25 for serial position; 104 for intonational pattern, and 26 and 104 for transitional probability). What has not been tried so far is to control for all these factors at the same time. One might still argue that even if one factor is neutralized, the click displacement is caused by one of the other factors, not by a 'pure' clause boundary effect. Thus, for instance, in a very careful study on response bias (23), which will be discussed later, the displacement effect could have been fully caused by intonation instead of by the clause boundary. The only way to control for all factors at the same time is to use the splicing method (as in 241). The study by Garrett *et al.* (104) comes the closest to this ideal.

Before treating the second issue of the perceptual origin of the effect, we should mention various additional factors which have been shown to affect click position. These factors, however, do not normally covary with clause-boundary position. Bertelson and Tisseyre (16) showed that prior knowledge of the sentence reduces preplacement of the click, and that separation of the click and sentence in acoustic space affected the degree of preplacement of the click (17). This effect apparently depends on reading direction, since in Hebrew it is inverse

(15). This spatial separation effect is not due to the response mode (written or oral report of click position, 18); it can explain the fact that negative displacement of the click is strongest if the speech is in the right ear and the click in the left ear (25, 26). Finally, click localization is more accurate (i.e. there is less of a clause-boundary effect) if the subject is free to ignore the contents of the sentence (see 21, 23, 229), especially if he is not required to recall, but receives a written version of the sentence on which to mark the click position (as in 16,209, 210). This could not be verified in 131.

With respect to the second issue, regarding the perceptual origin of the click-displacement phenomenon, the click-localization technique shares the disadvantage of all successive measurement tasks: the effect arises somewhere between the initiation of the sentence and the subject's response, and it is hard to find out exactly where. Since the localization response usually occurs several seconds after presentation of the sentence, there is good reason to consider whether the mechanism of response selection is partly or wholly responsible for the click shifts. We will in the following pages go into some detail in order to demonstrate that even with sophisticated experimentation it is extremely troublesome to argue for the (strictly) perceptual origin of a phenomenon on the basis of a successive measurement technique, be it click localization or anything else. The obvious way to check for response selection factors is to compare real click localization with 'click localization' when no click is presented. Ladefoged (165) was the first to try this out in an exploratory way; it was followed by an equally exploratory study by Reber and Anderson (210).

The first experiment systematically measuring the distribution of guessing bias was Reber's (209). It revealed a rather dramatic guessing effect: subjects clearly preferred to locate the non-existent click in or around the major syntactic break. This is the case in spite of the fact that subjects are not requested to recall the sentence, so that one would expect such results *a fortiori* if the subject were also set to actually recall the sentence. When this was tested by Bever *et al.* (23), the results were negative: they found only a small effect of clause break for the 'subliminal' click condition. The graph of their results shows that there is a bias to locate the non-existent click either a half-syllable before the break or in the break. But the attraction of breaks is significantly stronger in the real-click situation, although there is also in this case a bias towards preposing the click. It is certainly correct to conclude from these results that the clause-boundary effect is stronger for real clicks than for non-existent clicks, and therefore that since guessing cannot fully explain the click-displacement effect, something more strictly perceptual is also involved. It cannot be concluded, however, that this perceptual factor is the clause boundary *per se*. In this experiment, clause break was contaminated with intonation, and it therefore could have been a perceptual effect of intonation alone (which has been shown to be an important variable; see 210). The fact that Bever *et al.* (23) found a lesser effect for guessing than Reber did (209) needs explanation. The recall set in Bever's study as opposed to

Reber's is difficult to interpret as a possible cause since Reber found strong guessing-bias effects in a situation where no recall was required. It is not clear why this bias should disappear if the subject is forced to attend to the contents of the sentence. There are additional differences between Reber's and Bever's experiments which should be explored in more detail. The first possible cause of this difference is that Reber used 'normal intonation', whereas Bever *et al.* used 'monotone intonation'. In view of the strong effect of intonation on click localization (210), this experimental difference is worth a further check. The second possible cause is to be sought in the number of response alternatives. Reber's subjects could localize the non-existent click at any position in the sentence, i.e. at any of 25 possible locations (within and between syllables). Bever *et al.* gave their subjects a 'window'; that is, a limited region on the typed-out sentence within which the click should be located. This region contained nine possible positions. If one now looks at the clausal break and the two immediately adjacent positions, one will see that they make up 12% of the possible positions in Reber's case, and 33% in Bever's. Reber found 33.8% (subliminal) click localizations in this region and Bever 52%—in both cases substantially more than could be expected on a non-bias basis. The fact that Bever's curves are less 'peaked' than Reber's can therefore be partly due to the window technique: the subjects have to 'pile up' their responses in a limited region. A final difference between the two experiments may have worked the other way. Reber used a control group for his 'subliminal' click condition, whereas Bever *et al.* used click and non-click conditions mixed within subjects. In their experiment the response window for the subject was always centred around the real click position; the subject might therefore conclude that this was always the case, and apply this knowledge in the non-click condition; i.e. he might position the click in the middle of the window on the answer sheet. This would have favoured the break ± 1 positions.

In summary, it has been shown that a great many factors can influence click localization, some of them presumably affecting early stages of the decision-making process (especially intonation). The experiments showing an early origin of the clause-break factor contaminate this factor with other factors; on the other hand, clause break also seems to have an (early or late) effect of its own if it is uncontaminated.

(ii) *Dichotic switch localization.* This method was introduced by Wingfield and Klein (267). The subject listens to the sentence dichotically. The sentence starts out in one ear, and at some point switches to the other ear. It is the subject's task to localize the switching point after he has heard the sentence. The method is very much akin to the click-localization technique, and in fact it gives very similar results. Wingfield (266) and Wingfield and Klein (267) used the tape-splicing method of Garrett *et al.* (104), which is an effective control for prosody

and transitional probability. An example is given in the following sentence pair (1) and (2):

- (1) Besides *commercial uses of colour movies*, they are simply enjoyable;
- (2) Among the *commercial uses of colour*, *movies* are most typical,

where the middle part of the sentence can be spliced into the other sentence.

Results showed that switch localization was more accurate if the switch coincided with the clause boundary. However, this was the case only if prosody 'agreed' with the clause structure (i.e. in unspliced sentences). If clause structure conflicted with the prosodic pattern (i.e. in the spliced versions), there was no significant clause-boundary effect, only a significant effect of prosody. Wingfield also found, in agreement with the click studies (21, 23, 229), that localization is more accurate if the subject is not required to recall the sentence. In conclusion, there is also here at most a very slight effect of clause boundary, but a strong effect if it coincides with prosodic pause. Darwin (69) used sentences similar to Wingfield and Klein's (i.e. normal and cross-spliced sentences) in a dichotic switch situation where the subject did not have to indicate the switch position but simply to write down the sentence. Darwin analysed errors as a function of switch location in the sentence. Cross-spliced sentences led to substantially more syntactic errors, but not to more lexical errors. Errors near the intonational boundary had highest rates if an intonational boundary preceded (rather than followed) a major syntactic boundary. We will return to this finding in section 3.1.

Can one control for response bias in the switch-localization technique? Flores d'Arcais (see his paper in this volume) has managed to do so by presenting both parts of the sentence to both ears while manipulating the loudness balance between the ears: a switch is then perceived as a small 'shift in the head' of the message. In this manner fake shifts can be interspersed (by not affecting the balance at all) in order to measure response bias effects.

(iii) *Transient memory-load.* Savin and Perchonock (223) measured the mental load of incoming syntactical material by additionally giving the subject (either just before or just after the sentence) a short word list to remember. The number of recalled words was supposed to be an indication of the spare STM capacity during the processing of the sentence. Foss and Cairns (90) used this technique to show that recall of complex sentences was worse than recall of simple sentences if the subject's first task was to remember the word list. Wanner and Shiner (261; see also 259, 260, 262) extended this method in the following way: a sentence is presented visually word by word. The sequence is interrupted at some (test) position for presentation of a short list of words to be remembered; after this the sentential sequence continues. The subject's task is (i) to comprehend the sentence and (ii) to remember the list of words. Wanner

and Shiner (261) showed the validity of the technique on a visually presented arithmetic task. Wanner *et al.* (259, 260, 262) used it in psycholinguistic experiments, especially to test predictions from an Augmented Transition Network (ATN) model (see section 4.3.2) of sentence parsing. In these experiments the authors use a combined comprehension plus recall score as their load measure, a concession which expresses the feeling that the situation is not really one of a main plus an additional task but one of two main tasks. The combined measure, however, is quite sensitive. A major disadvantage for study of sentence perception is the unnatural character of the presentation: a slow word-by-word presentation combined with a major interruption. It is likely that the perceptual process itself is also slowed down, and is quite different from the process which occurs in normal listening.

2.2.2 *Recall*

Recall is an obvious procedure to be used in the study of memory. Psycholinguistic studies of memory of sentences are very numerous (see 172), and various recall methods have been used for this purpose. Here we will limit ourselves to studies where the researcher's interest was basically in sentence perception but where some form of recall was used in the experiments. Many of these studies have been designed to show that a certain variable affects the subject's processing of verbal material, without much concern for the particular stage between stimulus presentation and recall at which the variable accomplished this. These are often loosely referred to as studies in comprehension. Examples are Bransford and Johnson (32) and Frederikson (96), who used free recall to study the effect of context on the understanding of text. Both studies claim that the effect of context is located somewhere in the encoding process, not in the recall phase. Only the first study systematically varies context before or after the prose passage. Results show that context placed before is much more effective, which certainly makes it likely that context is helpful during reading in these cases. Kintsch *et al.* (158) varied the number of word concepts in a text and found an effect on recall; the variable also affected reading time, indicating that it might operate during encoding. In other studies more specific perceptual claims are made, and they invariably lead to more difficulties in interpretation. A typical example is Levelt's (169) experiment where subjects were presented with sentences embedded in white noise; after each sentence they had to write down what they had heard. The data were analysed in terms of conditional probabilities that word i would be reproduced when given that earlier word j had been reproduced, for all words i and j in the sentence. In this way it was possible to show that the data structure was strongly hierarchical, and it was concluded that 'hierarchical left-to-right chunking will often be an adequate model for sentence processing'. It is not at all certain, however, that the hierarchical structure reflects the way in which the sentence is

chunked during presentation. Loosen (179) repeated this experiment with one change. He presented the sentences with their word order scrambled, but asked the subjects to reproduce these word lists in the form of a sentence (if possible). The data were very similar to Levelt's, even though the subjects could not have chunked the words in the way hypothesized in that study. The hierarchical structure must have resulted during a complicated storage or retrieval process. (Kempen, 155, could show that it is in fact a retrieval effect.) Of course this does not prove that the original conclusion was wrong, but it certainly shows that it was insufficiently motivated. Dooling (72) used a similar technique to Levelt's in a study where subjects were set for a particular rhythmic and/or surface structure. Only change of rhythm had a (major) effect on subject's sentence reproduction, and it appeared quite likely that rhythmic expectations enhance perceptual efficiency. Still, it is not impossible that the subject's response selection is facilitated by a given rhythmic structure. If a perceived element is not quickly read out of STM it may be lost, and reading out may be hampered by a change in the response rhythm. An experiment of Forster (84), using immediate sentence reproduction in combination with a rapid visual presentation technique, did control for the possibility of losing a perceived element from STM. (The technique of rapid serial visual presentation, RSVP, involves a quick word-by-word presentation of the sentence in such a way that successive words are centred at the same visual localization.) Here, subjects were requested to monitor for a word beginning with a particular pair of letters. Subjects performed better at monitoring than at recall. This indicates that in the recall of rapidly presented sentences, perceived elements may get lost during retrieval from STM. A recall technique alone is insufficient for deciding on the stage where the effect arises. Carpenter (43) combines the technique with a verification reaction-time procedure.

Immediate recall is also used in the experiments by Jarvella *et al.* (140, 141, 142), but here the purpose is rather different. Jarvella gave a spoken (142) or written (141) text to his subjects; it was interrupted at some point and the subjects' task was to recall immediately as much of the text as they could. A major finding was that verbatim recall was at a high level for the last clause and/or sentence only. In this way Jarvella could distinguish between stages of comprehension: the earlier material had already been interpreted and only the last clause or sentence was still available in verbatim, uninterpreted form. Here also, control was necessary to exclude retrieval as the sole explanation for the findings. It could be that the subject started out by reproducing the last clause and only thereafter went back to earlier material. This time-lag might explain the results, but Jarvella controlled for such a possibility in an experiment in which subjects were presented with prompt words and were requested to recall from that point on. The results corroborated the earlier ones: a prompt word was most effective if it had been selected from the final clause. Jarvella's experiments have been extended by Marslen-Wilson and Tyler (188), to whom we will return (section 4.1.2). For a similar method, see 211.

2.2.3 Recognition

Retrieval processes which are always at work in recall tasks are less important in recognition tasks. A good example is Caplan's (38) study where a word-recognition test followed the presentation of a sentence. The subjects' task was to judge whether the word had occurred in the sentence, and reaction time was measured. The results confirm very well Jarvella's recall findings (141,142): the last clause has a substantially higher verbatim availability. See also Kornfeld (161) for replication and extension of these findings.

Larger units, such as complete sentences, have also been used in recognition tests. The classic study here is by Sachs (212), who gave her subjects a story in which a critical active or passive sentence had been embedded. The story was followed by a recognition test for either the same sentence or its transformed version. If the sentence was the last sentence in the story, recognition was close to perfect; but if the critical sentence had occurred earlier in the story, recognition went down to about 60%. This shows, again in agreement with Jarvella's findings, that the verbatim version of a sentence is quickly lost. Anderson and Bower (5, 7) have used and extended Sachs' technique with essentially similar findings. A signal-detection analysis of Sachs-type data can be found in 235. Although reconstructive retrieval processes are probably eliminated in recognition experiments, they share the disadvantage of all successive measurement techniques that it is impossible to differentiate input explanations from storage explanations. How was the sentence registered in memory, and how well was it kept there? It is likely that in all the above experiments recognition failure resulted from forgetting, not from failure to detect in the first place. But only independent detection tests, such as Forster's (84), can be decisive in this respect.

Finally, caution is needed with the often implicit assumption that recognition equals recall minus reconstruction. Carey and Lockhart (41) show for word lists that subjects employ different encoding operations, dependent on whether they expect a recognition task or a recall task (see 251 for a similar result). Nothing is known about whether the same task dependency is operative in sentence or text comprehension.

2.2.4 Paraphrasing

Though paraphrasing tasks are often used for the sole purpose of forcing the subject to interpret the sentence, there are some studies where paraphrasing, especially paraphrasing reaction time and accuracy, is used to determine the perceptual complexity of a sentence. Examples are the studies by Fodor *et al.* (80,83), where among other variables self-embedding and verb complexity were shown to affect paraphrasing accuracy and reaction time. Just as with the recall studies, it is ambiguous at which stage between stimulus and response the variable is effective. Hakes (117) writes 'Paraphrasing seems to require that the S

Table 1.1 Paraphrase and phoneme-monitoring results in relative pronoun deletion, self-embedding, and complement-that deletion

Authors	Independent variable	Dependent variable					
		Ph.mon.RT	Paraphrase				compl.
			Acc.	RT	Acc/RT	Sent.	
Fodor and Garrett (1967)	Rel. pronoun deletion	0	0	0	+	0	
Fodor, Garrett and Bever(1968)	Verb structure	0	+	-	0	0	
Foss and Lynch (1969)	Right branching vs self-embedd.	0	+	0	0	+	
	Rel. pronoun deletion	0	-	0	0	-	
Freedle and Craun (1970)	Degree of self-emb	0	+	0	0	0	
Hakes and Cairns (1970)	Rel. pronoun deletion	+	+	-	0	0	
Hakes (1971)	Verb structure						
	Expt 1	-	+	0	0	0	
	Expt 2	-	-	0	0	0	
Hakes (1972)	That-deletion in complement:						
	Expt 1, Easy sentences	+	-	0	0	0	
	Expt 2, Diff. sentences	+	+	0	0	0	
Hakes and Foss (1970)	Relative pronoun deletion	+	-	0	0	0	

0 dependent variable not tested.
 + dependent variable tested and significantly affected.
 - dependent variable tested and not significantly affected.

at least: (1) comprehend the presented sentence; and (2) by modifying the sentence's lexical content and/or structure, construct another sentence that is roughly synonymous. Thus, while a paraphrasing error might reflect a comprehension failure, it might also reflect a failure occurring during the construction of the paraphrase'. Hakes *et al.* in a series of papers (117,118,119, 120) tested some of the original Fodor *et al.* variables comparatively in paraphrasing and phoneme-monitoring tasks. The results of these and some similar studies are summarized in Table 1.1. These results are at best confusing. Relative pronoun deletion has been studied in three experiments. Fodor and Garrett (82) measured both paraphrasing accuracy (Acc) and paraphrase delay (RT). They expressed their data in Acc/ RT. If one reads their text carefully it appears that in at least one experiment delay itself is significantly affected by the variables. Foss and Lynch (92) used two other dependent measures, phoneme-monitoring and sentence completion. Neither showed an effect of relative pronoun deletion. Hakes and Cairns (119), finally, found significant effects on phoneme-monitoring RT and paraphrasing accuracy, but not on paraphrase delay. Thus, each dependent variable showed an effect in just one study. Verb structure was manipulated in two studies. Fodor *et al.* (83) found an effect on paraphrasing accuracy, but not on paraphrase delay. Hakes (117) found an effect on paraphrasing accuracy in one experiment, but not in the other, and no effects on phoneme-monitoring. Degree of self-embedding was manipulated in two studies, using different dependent variables: phoneme-monitoring and sentence completion in Foss and Lynch (92) and paraphrasing accuracy in Freedle and Craun (97). All showed significant effects. These and the other results on Table 1.1 show that paraphrase delay, paraphrase accuracy, and phoneme-monitoring behave quite independently and differently for different variables. Thus, there is no basis for the strong generalizations that can be found in some of these papers.

2.2.5 *Verification*

In a typical verification experiment, the subject is presented with a sentence and is asked to verify its truth with respect to some other source of information (a picture, another sentence, or pre-existing knowledge of the world). There are two dependent variables, reaction time and error rate. Mostly (but not always; see 137), reaction time is the critical variable. It is used to discover or test processing stages mediating between stimulus and response. The basic assumption is Donders' (71) notion that a response is the result of a linear sequence of operations, with each operation taking a certain amount of time, with these times being additive. Donders' so-called subtraction method consisted of adding or removing an operation in the sequence and registering the resulting increase or decrease in RT. The difference would amount to the characteristic time for that operation. A much-used application of the subtraction method in sentence

verification is the following. It is supposed that coping with a negative involves an independent operation. In order to measure its effect, negative sentences are compared with affirmative sentences and differences in RT are determined. More important than determining characteristic operation times is demonstrating that a certain operational stage is really independent. It has to be shown that the stage contributes to the RT independently of what happened in earlier or later stages.

In a very fundamental paper, Sternberg (238) extends Donders' method in order to show the additivity and independence of processing stages; this is called the *additive factor method*. This method does not require the elimination of a processing stage but only the manipulation of the processing durations of each of the hypothesized stages. If two stages are supposed to be independent and successive, then it should be possible to find an experimental variable which specifically affects the duration of one stage but not the other, and inversely. Experimentally this should show up in the absence of an interaction between the factors in an analysis of variance (on the raw RT data). This confirms additivity but not necessarily independence. For independence it is, moreover, required that the higher cumulants of the RT distribution (i.e. variance, skewness, etc.) are also additively affected by the factors. Donders' method is a limiting case of Sternberg's in the sense that eliminating a stage is equal to reducing its processing duration to zero. Verification RT measurement, while successive in character, does not share all the disadvantages of other successive measurement techniques, but only by making additional assumptions. More specifically, one has to make assumptions about the successiveness, nature, and order of stages. If, for instance, a perceptual stage is hypothesized, then one can define an experimental variable which, by the nature of the stage, will affect its processing time. By manipulating the variable one can find out whether the stage is really independent and what its temporal characteristics are. Thus although neither Donders' nor Sternberg's method can decide on the *order* of stages, it is possible to hypothesize such an order on independent grounds and then to verify the model by choosing appropriate experimental variables in a factorial design. To take just one example, Banks, Clark, and Lucy (13) presented subjects with such questions as 'Which balloon is higher/ lower?' and 'Which yoyo is higher/ lower?' followed by a picture of two balloons tied up at different heights or of two yoyos hanging down at different heights. The subject had to press the left or right reaction key in correspondence with the highest/lowest balloon or yoyo. The model is a two-stage model." a perceptual followed by a linguistic stage. At the perceptual stage the picture is registered. Various factors may specifically influence the duration of this stage. The authors selected as their experimental factor discriminability, i.e. difference in height of balloons (or yoyos). They used three levels for this factor. The linguistic stage is supposed to compare the perceptual output with the linguistic representation of the question sentence. A factor which might specifically affect the duration of this is the congruity

between perceptual and linguistic code. If a balloon is perceptually coded in terms of 'high', and a yoyo in terms of 'low' (the authors' hypothesis), the congruent question ('Which balloon is higher', 'Which yoyo is lower?') should be easier than the incongruent question ('Which balloon is lower?', 'Which yoyo is higher?'). It turns out that the two factors have additive effects on mean RT; there is no statistical interaction, testifying to the successiveness of the perceptual and the linguistic stage. The authors do not test the additivity of the higher moments of the RT-distribution; strictly speaking, therefore, the independence of the stages has not been shown. There is, however, not a single sentence verification study where the additivity of higher cumulants is analysed—a very general technical omission in these experiments. Otherwise, this study is an exemplary application of Sternberg's method. If the two test factors do show statistical interaction, then one must reject the possibility that there are successive independent stages. The interpretation of the way in which the two processes interrelate, then, depends on whether the joint effects of the two experimental factors is super- or subadditive. Sternberg points out that superadditivity may indicate limited capacity sharing of two parallel processes, whereas subadditivity may indicate parallel processing where RT is determined by the slowest of the two operations. Krueger (162) did an experiment where subjects matched a short phrase (such as 'is north', 'isn't east') with the position of a small circle with respect to the typed phrase (above, below, left, or right). He tested whether negation coding ('isn't' versus 'is') occurs at a different stage from feature matching (e.g. comparison of 'north' and position above). He found an interaction, and the interaction was superadditive. His conclusion was that the two processes occur in parallel but have to share a limited-capacity central processor.

These examples suffice to show some of the purely technical possibilities and difficulties attached to the verification method. The major issue in this review article should, of course, be the theoretical 'stuffing' of the different stages. We will return to this point in section 3.5. Here we will conclude by distinguishing three types of verification tasks. The main set of verification studies is concerned with sentence/ picture comparisons. Coming back to our trademark metaphor, we see that this is certainly the trademark of Clark and his colleagues (see 13,44, 45,49, 50,54, 56, 57, 58,60, 148, 149). But it has also been widely used by others as well (65, 75, 98, 108, 160, 162, 204, 215, 247, 248, 256, 257) especially for the study of ambiguity (40, 89, 205). Table 1.2 (see section 3.5) summarizes most of these studies. (Word-verification studies are not mentioned.)

The second variety of verification studies has the subject test the truth or falsity of a sentence presented in isolation on the basis of his common knowledge of the world (e.g. 'Lions are more ferocious than sheep'—true or false?). Most studies in this category are basically interested in the organization of semantic memory and employ test sentences having the form of quantified statements ('All canaries are birds') (see 63, 106, 107, 134, 192, 213, 214). Other studies

concern comparatives (43, 75) or variation of imagery value (147). Another set of studies is concerned with the relationship between clauses of a sentence; e.g. verifying the truth (151) or oddness (231) of 'if. . . then' statements. These latter studies are similar to the third variety of verification studies.

This last, and smallest, group of studies involves some sort of sentence/sentence verification. A test sentence has to be verified on the basis of some earlier presented sentence. Here one can either vary the delay between the sentences (5, 42) or present the two sentences simultaneously, as in Stillings' (240) experiment. Since we have categorized the Banks *et al.* study (13) in the picture-verification group, even though it is essentially a question-answering task, we should also mention here some question-answering RT experiments using sentences as test material. These are mostly inference experiments (e.g. 'A is better than B, B is better than C. Who is best?'; see 146, and also 52, 139), and are concerned with an imagery versus congruence explanation of the latency data. For a review of this Clark/Huttenlocher discussion, see Johnson-Laird (144). Amnon *et al.* (4) studied latencies for answering questions about self-embedded constructions and compared their results with probe-latency data. Garrod and Trabasso (105) studied latencies for answering questions about earlier presented text, very much as Anderson (5) has done.

2.2.6 *Question answering*

Question answering (QA) is the trademark of researchers in semantic memory. With respect to sentence comprehension we have already mentioned the QA latency studies in the framework of the preceding section. QA accuracy has been used in just a few studies. Among them are Blaubergs and Brain (27), who investigated self-embedding constructions as opposed to right-branching ones. Wright (271) measured QA accuracy as a function of syntactic correspondence between question and answer. Finally, three studies (Smith and McMahon, 234, Bever, 19, and Locatelli, 178) deal with the comprehension of sentences expressing two-event temporal orders. The paradigm is to ask the subject, 'Which event occurred first? Neither Smith and McMahon nor Locatelli found an effect of order of mention on QA reaction time. Bever relates this finding to data on 2-4 year-olds where order of mention does matter.

2.2.7 *Miscellanea*

Various other tasks have been given to subjects in order to study sentence comprehension. In *comprehension latency tasks*, subjects are requested to push a button as soon as they grasp the meaning of the sentence. Examples are 59, 126, 159 and 207. Sometimes a subject is merely asked to make a *comprehensibility judgment* (123, 133, 196, 228, 254, 255). *Grammaticality*, *acceptability*, or *meaningfulness* judgments were collected in various studies (55, 127, 197, 199). In *choice tasks* the subject is requested to choose between two

meanings (175) or two implications (160) of a sentence. In some experiments subjects had to *judge compatibility* of sentences (35, 74, 111, 112), and various other more idiosyncratic techniques have also been used to study sentence comprehension.

2.2.8 *Some conclusions*

In the preceding sections some of the major experimental procedures used in studies of sentence comprehension have been reviewed. Among others, the following observations can be made:

(i) The research group is often a better predictor of a research method than is the theoretical problem.

(ii) Comparisons between methods, if made at all, are mostly inconclusive since apart from changing the dependent variable, one nearly always changes the independent variable as well.

(iii) If the research objective is merely to show the 'psychological reality' or 'validity' of some linguistic notion, then most of the procedures we have mentioned can be appropriate.

(iv) If moreover, one wants to determine the stage at which a certain variable is effective, it is preferable to use either a simultaneous measurement procedure or a (successive) reaction-time method. The former is especially adequate if one is interested in the strictly perceptual stage, i.e. the processes operative during input of the sentence. The latter in principle allows for a Sternberg-type additive-factor method, but one has to make assumptions about nature, successiveness, and order of stages.

(v) Contrary to the original interpretation of phoneme-monitoring results in terms of the sharing of a limited capacity, detection and analysis of the target word seems to be a condition for phoneme detection. It is therefore necessary to control for all variables which might affect detection of the target word—which is rarely done.

(vi) Psychophysiological measurement during sentence perception is still an underutilized procedure.

(vii) Click localization can be affected by a great variety of factors, some of which clearly affect the very early stages of processing. It appears to be very troublesome, however, to vary just one of these factors, while keeping all others constant.

(viii) A major problem for recall tasks is to control for reconstructive activities which take place between the stimulus presentation and the response.

3. LINGUISTIC AND EXPERIMENTAL VARIABLES

The experimental measurement procedures discussed in the former section have all been developed to study linguistic and other factors which could be expected to affect language perception. In many cases these independent variables had

been derived from more encompassing theories, in other cases they seem to appear out of the blue, so to say, derived from interesting observations, well-known phenomena, and so forth. Whatever their origin, we will summarize the most important experimental variables that have figured in studies of sentence perception. In section 4, we will go deeper into the more general theoretical considerations which have been in the background of some of these studies.

3.1 Prosody and rhythm

Though the importance of suprasegmental features in the perception of sentences has never been seriously denied, variables such as pause structure, intonation contours, and speech rate have often been treated as nuisance variables which had to be controlled in order to study the effect of more important factors such as syntactic complexity and constituent structure. The result is that relatively little is known about the role of suprasegmental features in sentence perception, and especially little about how these features interact with other factors in the perceptual process.

Martin and his associates (189, 190, 233, 242) have done some basic work on the role of rhythm in sentence perception. They varied the structure of sentences by splicing tapes, inserting pauses or stretches of white noise at different places in naturally spoken sentences, or removing short stretches of speech. Their main measurement procedure has been phoneme-monitoring. They found a strong increase in reaction time when the rhythmic structure of the sentence was mutilated, even if the mutilation took place long before the target phoneme. Martin (190) concludes that 'the main outlines of syntactical structure often are communicated relatively early in an utterance', and 'suprasegmental cues enable the listener to expect or anticipate rough outlines of speech not yet heard'. Dooling (72) confirmed this latter claim by showing that if a subject was set to recognize sentences (in noise) with a certain rhythmic pattern, he would tend to fail if a sentence with a different rhythmic pattern was suddenly presented.

Dooling counterbalanced syntactic changes that might also be involved and found that only the rhythmic changes were effective. In view of the above-mentioned 'nuisance' status of suprasegmental features, it is no surprise to find various studies in which such features are 'traded against' syntactic structure, as in Dooling's article. As we have already noticed in section 2, some effort has gone into showing that suprasegmental features, and especially intonation, do affect phoneme-monitoring results (e.g. 67, 233), click localization (210), dichotic switch localization (266, 267), and bias in ambiguous sentences and phrases (173, 274). A similar effort was put into showing that these features were not the sole ones and that syntactic variables alone would also be effective (80, 104, 241).

A more integrative approach would be to study exactly how the listener uses prosodic information to make early guesses about the clausal structure of the

sentence. Here we should mention the pioneer work on intonation contours by Cohen and his associates at the Institute for Perception Research (61, 62, 124, 125, 154). They found that in order to render a syllable or word more prominent the speaker could use either a rise or a fall in intonation, or both. In the latter case, such a rise/fall has little predictive value. However, if only a rise is used, then at some place before the end of the sentence the intonation should go back to normal. The speaker has two options here: he could either go back to normal at another prominent word, using the fall as a marker there, or—and this is the critical issue—maintain his high intonation until the end of the clause and then begin the next clause at the normal level. This latter way of speaking does not lend prominence to the first word of the new clause. Thus, if a high intonation is maintained by the speaker after a rise, the listener can predict that either a second prominent word or a clause break will follow. The experiments show that listeners do have such expectations, and that they mostly expect clause breaks. Also the experiments by Darwin (69) mentioned earlier (sections 2.1.2 and 2.2.1) show that intonational cues are used to listen selectively and to delimit higher-order syntactic units. The question remains whether speakers normally provide sufficient prosodic information to enable the listener to use it predictively. Levelt *et al.* (175) found that disambiguating prosodic information is nearly absent if the ambiguous sentence is spoken in a disambiguating context. Here, clearly, the listener is supposed to be able to derive the correct reading of the sentence by using other sources of information.

Finally, variations in rate of presentation have also been studied, but almost exclusively in connection with visually presented sentences. High presentation rates have been used to study the effects of linguistic complexity under borderline conditions (84, 86, 129, 132, 163). Schwartz *et al.* (228) found that under high-rate conditions subjects failed to understand sentences but were still able to judge reliably the comprehensibility of the verbal material. This is in correspondence with Mistler-Lachman's results on levels of processing (197, 198, 199) which will be discussed in section 4.1.2. Miron and Brown (196) studied the effects on intelligibility of speeding up acoustic material. It was found that speech-pause times especially could be reduced without much effect on intelligibility.

3.2 Syntactic complexity

Studies of syntactic complexity originated from the aim to test the 'psychological reality' of linguistic notions. As was discussed in section 1, this aim was twofold: to demonstrate the psychological validity of linguistic units and structures, and to show a one-to-one relationship between linguistic (transformational) rules and perceptual processes. The former goal was achieved more successfully than the latter. Characteristic of studies of the latter sort were efforts to vary the transformational complexity of sentences and to

show that comprehensibility would covary. Although Fodor and Garrett (82) unmasked and disproved the 'Derivational Theory of Complexity' (DTC) underlying these studies, active/passive and affirmative/negative variables have continued to be used in studies of sentence perception, although usually not for the purpose of proving DTC.

3.2.1 *Voice*

Although there is a general tendency for passive sentences to be harder to comprehend than active sentences—in correspondence with DTC (see e.g. 177)—a closer look at the situations of use has shown that passives may lose this additional difficulty if they structurally match other aspects of the task situation. Olson and Filby (204) found that verifying a passive sentence with respect to a picture is easier if the picture evokes a 'passive' encoding—for example, if the largest or the focal object is the recipient of the action. A similar finding was reported by Flores d'Arcais (76) in an experiment where appropriateness of the descriptive sentence was judged by the subject. Wright (271) showed that answering questions about nouns in a sentence is relatively easy if the question corresponds in voice to the sentence. Green (112) found that voice correspondence between two sentences facilitated judgments of sameness and difference in meaning. Garrod and Trabasso (105) gave subjects short paragraphs of four sentences followed by a question about one of them. Apart from generally longer latencies for passive than for active questions, they found an interaction between voice of question and voice of sentence: if the two correspond, latencies are shorter, at least for the first and the last short sentence in the paragraph. A similar correspondence effect is reported by Anderson and Bower (7), especially for the last sentence in a relatively long text. These experiments were all inspired by the original study of Sachs (221), who showed that active test sentences at the end of a text are not recognized as passive (and vice versa), but that such confusions are the rule if the test sentence appeared earlier in the text.

The basic question, given these ubiquitous correspondence effects, concerns the availability of the surface and the underlying form of the sentence to the subject. On the one hand the correspondence effects can only be explained if the subject is able to work from a surface representation, while on the other hand there is evidence (see, for example, Clark, 54, Fodor *et al.*, 80) that subjects can operate on the basis of underlying structure. One solution is to assume that subjects have the surface form, and therefore the voice of the sentence, available for a rather short period, but that they will lose it as soon as the sentence has been interpreted.

This assumption is compatible with the theory that the sentence clause is available verbatim in STM but that only its interpretation is transferred to

LTM. For the voice experiments this theory is contradicted by the following evidence: (i) Wright's (271) finding that the correspondence effect still occurs if the subject is forced to count backwards between test sentence and question—this activity presumably erasing STM content; (ii) Garrod and Trabasso's (105) finding that not only the last sentence in their list of four, but also the first (which should have been overwritten in STM) shows the correspondence effect; and (iii) the finding of Sachs (221) and of Anderson and Bower (7) that some correspondence effect remains even if the test and the recognition or question sentence are far apart.

The other solution is that both surface and underlying form are available to the subject for a relatively long period, as already suggested in Garrod and Trabasso's article. Further evidence is given by Anderson (5). This solution is compatible not only with the idea that the surface form is available both in STM and LTM (and Baddeley, 9, 10, gives convincing evidence that the recency effect can very well occur for information in LTM), but also with the more radical view that memory is not multi-store (see Craik and Lockhart, 66) but that perceptual (surface) and propositional (underlying) information have different rates of decay. This issue will be further discussed in section 4.1.2. That traces of such surface information can remain available in memory may be due to the communicative function of active versus passive sentences.

A final set of studies is concerned with the function of passives in discourse. Anisfeld and Klenbort (8,160) showed that subjects preferred implications from passive sentences in which the logical subject is also the focal information point (comment). Passives are more markedly topic-creating in this respect than are actives. Grieve and Wales (115) and Hupet and Le Bouedec (138) show that voice can be confounded with definiteness of subject and object noun-phrase, especially in full passives, and that definiteness had a topicalizing effect of its own. Hupet and Le Bouedec show that these variables are not independent in normal communicative usage: subjects clearly prefer the grammatical subject to be definite in both passive and active sentences, indicating that the passive voice is used if the logical object is presupposed, and that the logical subject carries the assertional information. Further evidence for this presupposition-creating function of passives can be found in Hornby's (137) study. Presumably, listeners assume this functional interaction between surface form and presupposition when they interpret passive forms.

3.2.2 *Negation*

'What is so difficult about negation?' is the title of one paper (Wales and Grieve, 253). One answer is that negative sentences are transformationally more complex than affirmative ones and are therefore harder to understand. This answer, however, is a variant of DTC and clearly not satisfying. Wales and

Grieve conjecture that a possibly important factor is 'confusability', i.e. the degree to which the negative statement's interpretation resembles the affirmative's interpretation (e.g. 'seven plus nine is not fifteen' should be harder than 'seven plus nine is not seventy seven'). Though this factor seems to work, Greene and Wason (113) show convincingly that confusability does not interact with negation. The two factors are additive. One could say that they involve independent stages (in Sternberg's sense) in the comprehension process. The stage-approach to negation has been popular in many of the verification studies (see section 2). In particular, many of the studies by Clark and his associates (45, 57, 58, 149) use negation as one of the independent variables (see Table 1.2); and the classical finding is that true affirmative (TA) sentences are the easiest and true negatives (TN) are the hardest to verify. Somewhere in between are false affirmative (FA) and false negative (FN), usually in this order. The initial processing models developed for the explanation of these findings (see in particular Clark, 54, and Trabasso *et al.*, 248) have in common an independent stage where the negative is recoded as a positive proposition. For instance, if the ball in the picture is either red or green, then the sentence 'The ball is not red' gets internally represented as 'GREEN (BALL)'. But at the same time, Clark (54) calls this model 'cheating', since normally there will be more than two alternatives, thereby excluding the possibility of this easy transformation. The two-alternative case does occur naturally, however, in the case of marked versus unmarked adjectives. Negating an adjective may lead to recoding (see, e.g., 150), and the use of a marked adjective badly complicates the understanding of double-negative sentences (232). Not only marked adjectives but also connectives such as *or* with a single alternative (*and*) behave as if they are psycholinguistic negatives (see 237). Clark (54), as well as Trabasso (247), handles the multiple-alternative case by assuming that the negative is encoded as a positive proposition embedded in a negative (e.g., Neg(RED BALL)). Since the picture is assumed to be positively coded, the Neg-element must lead to a change in the truth value when comparing sentence and picture codes. This requires additional time. This model can therefore handle the longer latencies for negative sentences, since an (additional) change of truth value is required if the picture coding is assumed to be positive. Tanenhaus *et al.* (244) object to this solution, especially to Carpenter and Just's (45) version of it, and we will return to the issue in section 3.5.

It should be noted that all these laboratory studies concern the cases TA, FA, FN, and TN, and that in such experiments negatives are usually harder to process than affirmatives. One might wonder whether these cases reflect the normal communicative uses of negation. Why, indeed, use negatives if they are harder to understand? Johnson-Laird (145) gave a first experimental example where negatives are in fact easier: One of the test sentences 'John is not rich' or 'John is poor' could follow the statement 'Either John is intelligent or he is rich'. The subjects found the negative test sentence easier to use when inferring that

John is intelligent. Johnson-Laird made the point that negation has a natural function. Wason and Johnson-Laird (263) present a thorough review of experimental studies of negation—to which the reader is referred for further details—and they analyse what this natural function might be. They firstly cite Russell (220): 'When I say truly "this is not blue", there is on the subjective side, consideration of "this is blue", followed by rejection, while on the objective side there is some colour, "different from blue" to suggest that the natural function of a negative is to deny a certain preconception'. This hypothesis was first tested and confirmed in two inventive studies by Greene (111, 112), who found, basically, that if a subject had to judge the agreement of meaning for a pair of sentences, one affirmative and the other negative, the task was easier if the sentences disagreed than if they agreed. The natural function of negation is to signal a change with respect to an existing belief or expectation. Wason and Johnson-Laird point out that contrary to most laboratory situations, in normal life listeners do not have to 'transform' negatives into affirmatives, since the affirmative is already there as a preconception—otherwise the negative would not have been used. Or to state it in the terms of the original laboratory experiments: negatives are normally used only as false negatives. The listener can be sure that the negative sentence is used to falsify an existing expectancy or belief. True negatives are unnatural laboratory constructions and they are harder because of their unnaturalness. This philosophy fits nicely in Clark and Haviland's (59) 'given/new' theory. One could say that the negative should only be used if the corresponding affirmative is given. Levelt and Noordman (174, 202) add to this a so-called principle of minimal change: the given affirmative is not fully rejected; it is instead maximally maintained: only a single component (argument or predicate for sentences, semantic component for lexical items) is affected by negation. This is at least so for 'normal' communicative settings. Language does provide means, however, for more radical negation, namely by negating sentences containing positive polarity items (e.g. 'John indeed didn't answer the letter'). The borderline grammaticality of such constructions will lead to an 'echo' interpretation, that is, it will negate a verbatim repetition of what the former speaker said. In such a case the whole positive preconception may be dropped (cf. Baker, 12). This, however, has never been studied experimentally. A thorough and comprehensive review of negation studies can be further found in Clark (54).

3.2.3 Relatives and complements

Section 2.2.4 described various studies which used paraphrasing and phoneme-monitoring tasks. Many of these studies used variations in relative clause and complement structure as dependent variables, and we refer to Table 1.1 for a summary of the major results. It was mentioned there that deleting the relative pronoun gave varying and in some cases conflicting results; it led to longer

phoneme-monitoring RTs in 119 and 120, but not in 92; it decreased paraphrasing accuracy in 82, 119, had no effect in 120, and increased accuracy in 92; and it had no effect on paraphrasing reaction time (119). The major rationale for these studies of relative-pronoun deletion was the so-called deep-structure clue theory: the listener makes direct inferences from the surface properties of the sentence as to its underlying structural relations. The relative pronoun could, for instance, function in a decoding strategy (see section 4.1.1) that might run as follows: For the sequence NP-Rel-NP₂, interpret the constituents NP₁, and NP₂ as object and subject of the same verb. Deleting the Rel might interfere with the strategy. The strategy is not error-proof, as Foss and Lynch (92) have shown, and the empirical evidence for this instance of the deep-structure clue theory is dubious.

Another reason for experimenting with relative constructions is the old question of why centre-embedded sentences are so difficult to understand. (The issue of self-embedding in psycholinguistics stems from Chomsky's proof that English is not a regular language; see Levelt, 172, volume 2.) Nearly all studies show centre-embedded constructions to be harder to understand than corresponding right-branching constructions, as long as the degree of embedding is two, or in some cases more than two (see Blaubergs and Brain, 27, with a recall task, Foss and Lynch, 92, with a phoneme-monitoring task, Freedle and Craun, 97, and Hakes and Foss, 120, with paraphrasing accuracy, Hamilton and Deese, 123, with comprehensibility rating, and various other studies). Note, however, that Hakes and Foss (120) did not find the predicted self-embedding effect in a phoneme-monitoring task and that Hakes *et al.* (121) criticized earlier studies because they had confounded self-embedding versus right-branching with object versus subject relatives. In a phoneme-monitoring experiment, and without this confounding, the author could not find a difference between self-embedding and right-branching constructions. However, some effect of self-embedding did occur for a paraphrase task. Blaubergs and Brain (27) were able to show that some learning did take place in the handling of self-embedding constructions, when subjects had been trained by performing tasks involving lesser degrees of embedding (but not inversely). Baird and Koslick (11) showed that verb-object relations were better recalled if the relative clause was of the subject-focus type ('The boy who kicked the ball chased the girl') than if it was of the object-focus type ('The boy whom the girl chased ran home'). Moreover, subject-verb relations were equally well recalled in nested and non-nested constructions, thus casting doubt on an interruption theory of self-embedding complexity.

Finally, we should mention three experiments with complements. Forster(84) showed that, with the exception of a few complement constructions, two-clause sentences were generally more complex than one-clause sentences (method: rapid serial visual presentation). Holmes and Forster (132) were able to extend this finding to include *that*- and *for to*-complements. They conclude that

sentence complexity is not related simply to the number of clauses. Hakes (118) found an effect of *that*-deletion in *that*-complement constructions, namely that in phoneme-monitoring RTs were longer in the deleted constructions. The effect of deletion on paraphrasing accuracy, however, was dubious.

3.2.4 *Verb structure*

The deep-structure clue theory also assigned a major role to the main verb of the sentence. Verbs have characteristic underlying structures in which they can occur, so that direct structural inference should be possible on the basis of the verb-token in the sentence. Such inference should be relatively difficult to make if the verb has several underlying structures into which it would fit. Fodor *et al.* (83) tested this prediction for 'complex' verbs such as *know*, which can take either a *that*-complement or an object-NP, as compared with 'simple' transitive verbs such as *meet*, which do not allow for a *that*-complement. The results were partly positive: paraphrasing accuracy and anagram solution accuracy were higher for simple verbs than for complex verbs. However, this variable affected neither paraphrasing RT nor anagram solution RT. Hakes (117) confirmed the paraphrasing accuracy result, but failed to find an effect on phoneme-monitoring. Holmes and Forster (132) used their rapid visual presentation method and showed that sentences with simple verbs were recalled better than sentences with complex verbs. These different results can be explained if the effect of verb complexity is merely reconstructive; that is, the creation of a complex verb sentence (in the paraphrasing recall and anagram tasks) may be harder, but this does not necessarily imply additional perceptual complexity. Levelt *et al.* (174) showed that sentences containing either simple or complex verbs of motion (like *move* vs *rise*) could be verified more quickly for complex verbs than for simple verbs, in perceptual situations where both sentences were true. In such situations, for example where one observes a quickly rising dot, the complex verb is more appropriate, in accordance with Grice's (114) maxim of quantity. A very different approach to verb structure is exemplified by Stillings (240). The subject was typically presented with a sentence such as 'Mary just loaned a book to John', and then had to verify the further sentence 'John didn't have the book'. The verification model predicting the reaction time is based on a system of meaning rules, which are written as programs for converting one sentence into the other. Stillings compared *borrow* and *loan* in various verification tasks and was able to substantiate his model.

3.3. **Ambiguity**

The effects of ambiguity on the comprehension of sentences have been studied intensively since the papers of MacKay and Bever appeared (180, 183). Initially a fruitful distinction was made between lexical, surface, and deep-structure

ambiguities. The first type involves ambiguous lexical items, as with 'the soldier put the gasoline into the tank'; the second involves word-grouping ambiguities as in: 'the three masted British ships were sailing south', and the last involves underlying relations without surface-grouping effects: 'Italians like opera as much as Germans'. Only the first and last variety are genuine homophones. Surface ambiguities are at best homographs; they should be, and in fact are, perceptually non-homophonic with explicit pronunciation (175). Experimental results with lexical ambiguities have sometimes been rather different from those with underlying ambiguities. Since there are theoretical reasons why they should be different (see 22, 91), and since almost all studies are on lexical ambiguity, we will begin with a review of lexical ambiguity and conclude this section with a few remarks on deep-structure ambiguities.

3.3.1 *Lexical ambiguity*

The main issues in the study of lexical ambiguity have been the following: (i) Does lexical ambiguity complicate the processing of a sentence? (ii) Are both meanings of the lexical item computed during sentence perception, or only one? (iii) How does context affect ambiguity resolution? (iv) When is ambiguity resolved: immediately or with a short or long delay? These questions are not independent; moreover, each of them needs further qualification. Let us consider them in turn.

(i) *Does lexical ambiguity complicate sentence processing?* The following studies have compared sentence processing with and without lexical ambiguity: 22,35,37,64,88,89,91,101,128,180,183,197,206,230. These studies involved a large variety of tasks, and in almost all cases an effect of ambiguity was found. In particular, one finds longer reaction times in the ambiguous case if the subject has to judge compatibility between the test sentence and a second sentence (35) or a probe word (230), or even if one has to judge the colour of the visually presented probe word (64). Longer RTs are also found for phoneme-monitoring if the test phoneme shortly follows the ambiguous word (37, 88, 91). Verifying the truth of a sentence (205), deciding on the ambiguity of the last word of a sentence (128), and verifying the sentence vis-a-vis a picture (89), all lead to longer reaction times in the ambiguous case. Furthermore, clicks are less accurately localized (101) in lexically ambiguous sentences. On the basis of a depth of processing argument (see section 4.1.2), Mistler-Lachman (197) predicted and found no effect of lexical ambiguity on RT in a meaningfulness judgment task.

The only tasks for which the results are conflicting are completion tasks. If the subject is asked to add an appropriate continuation sentence to the test sentence one finds either no effect of lexical ambiguity on RTs (22), or an effect with unspecified statistical significance (197), which, moreover, seems to appear if the

test sentence is preceded by a disambiguating context sentence. Also, if the subject has to complete an ambiguous sentence fragment, then the effect of lexical ambiguity is unclear when the fragment is an incomplete clause. MacKay (180) found 0.5 sec longer completion times (measured by stopwatch) for lexically ambiguous fragments; the effect is of unspecified significance. Olson and MacKay (206) found a significant effect on completion RT, but Bever *et al.* (22) and Cairns and Kamerman (37) did not. These differences, however, maybe more apparent than real: in the latter two studies, (insignificant) differences in RT of 240 and 200 msec were found respectively, whereas Olson and MacKay found a (significant) RT-effect of 200 msec, which is of the same order of magnitude. It should be added, moreover, that the statistics used in all of these studies are subject to Clark's (53) critique. At any rate, production tasks are not very sensitive to lexical ambiguity. All other tasks, however, showed the predicted effects of ambiguity, so that it makes sense to pose the next question:

(ii) *Are both meanings of the lexical item computed during sentence perception, or only one?* The additional effort of computing both meanings may explain several of the foregoing results. However, the question (which is asked in several of the reviewed papers) is highly confusing if it is not further qualified. What does it mean to 'compute' both meanings of a lexical item? It is certainly insufficient to state that 'both readings are in some sense available' to the listener (164), or that 'both meanings are activated in memory' (64). One might use Morton's (200) logogen-model to distinguish two senses of 'available'. If a homophone activates both logogens in LTM, they both become more available since both have an increased probability of reaching activation-threshold. If, moreover, both logogens do reach threshold they both fire (simultaneously, or in quick succession), thereby delivering the responses to working memory. This is a second and stronger sense of availability: the subject is conscious of both meanings. Foss and Jenkins (91) explicitly proposed this latter strong sense of availability. In the literature strong availability is often determined by asking the subject whether he noticed the ambiguity. If one trusts this test, as Foss and Jenkins apparently do, and if in the majority of cases the subjects turn out not to have noticed the ambiguity (a very general finding), it is in my view contradictory to maintain the strong version of double availability. In my opinion the test is not fully satisfying, and we are therefore left with three possibilities.

(a) During a lexically ambiguous sentence the homophone initially activates one logogen only. At a certain moment it fires, and that particular meaning is in working memory for structural integration.

(b) The homophone activates both logogens until one fires: one meaning becomes conscious, the other logogen gradually turns back to rest state. We call this the *weak theory of double availability*.

(c) The homophone activates both logogens up to threshold level: both meanings become conscious. This is the *strong theory of double availability*.

Alternative (a) should be dismissed. It contradicts substantial experimental evidence (see Morton, 200), and it is unclear how the homophone would 'know' which logogen to activate. However, there are naive versions of the so-called *one-meaning* or *unitary perception* theories that come close to (a). Both (a) and (b) are one-meaning theories in the sense that only one meaning of the word becomes available to working memory. Originally, many authors did not consider the possibility of activating a meaning without making it conscious. In Lashley's (167) *garden-path* paper, but also in Foss (88), Carey *et al.* (40), Cairns (35), and Hogaboam and Perfetti (128), we find the theory that only the most likely meaning of an ambiguous word becomes conscious and that as long as the most likely meaning is also the correct meaning (given the context and the task) there is no difference whatsoever from the unambiguous case. Problems only arise if the secondary meaning is intended. The subject should then retrace the blind alley and find the other meaning, this leading to longer reaction times. There is evidence, however, that even though only one meaning becomes conscious, the secondary meaning gets activated during comprehension; this is exactly the difference between models (a) and (b). The evidence mainly derives from the dichotic studies of Lackner and Garrett (164) and the extensive replication thereof by MacKay (183). Basically, these authors found that one can bias the interpretation of a lexically ambiguous sentence, presented to one ear, by simultaneously presenting a word to the other ear, which is related to one or the other reading of the ambiguous word. The stimulus conditions are such that the subjects are unable to reproduce the biasing word. If only the dominant meaning were retrieved in the way of model (a), it would be impossible to bias one way or the other (see 164). It seems necessary to allow for the possibility that both logogens become activated, as in model (b). In that case the biasing word might add to the activation of the corresponding logogen, increasing the chance that it will fire first. The other bit of evidence against model (a) is Olson and MacKay's (206) result that the completion of an ambiguous sentence fragment takes longer than the completion of a non-ambiguous control sentence. If the subject completes according to the first meaning which comes to mind, there will be no retracing, and model (a) would not predict the RT difference. Would model (b)?

To answer this question, one has to look into the relevant results in more detail. The authors find an effect of lexical ambiguity on completion RT only for those cases where the two possible interpretations are about equally likely (bias range of 40-60%). They propose the so-called *perceptual suppression theory*: the two sets of features (*logogens* in Morton's terminology) are supposed to interact: in order to activate one logogen to reach threshold level, the other one has to be de-activated. It is not hard to suppress the recessive logogen activity in the case of a 90/10 bias; the reaction time will not be noticeably different from the unambiguous case. Suppression, however, is difficult in a 50/50 bias situation since both logogens have a high level of activation. It should be remarked that in Morton's model logogens are independent; and one should ask whether

that simple principle should be abandoned in the light of these data. Not yet, I believe. In the same way as it is agreed that it takes relatively much effort to suppress a likely meaning (as in the 50/50 case), one could state that it takes relatively much effort to activate the other logogen up to threshold. This might be very easy for the dominant meaning in the 90/10 case — one does not need a suppression model here. Olson and MacKay give two further arguments for their suppression theory, but these concern the case where one meaning is already conscious. Suppression here means erasure from working memory. This has nothing to do with interaction between logogens. And finally, all other existing evidence for logogen interaction with homographs points to facilitation instead of inhibition, i.e. in order to activate one logogen, the activation of a related one may be helpful. This is the homograph effect reported by Rubenstein and his associates (218,219), and Schvaneveldt and Meyer (227). (Whether this effect is due to logogen-interaction, however, is doubtful, in view of the long reaction times in these experiments.) So far model (b), the weak theory of double availability without subthreshold suppression, seems to suffice. What about model (c), the strong theory of double availability? It seems to me that it is in general not necessary to suppose that both interpretations are simultaneously present in working memory. Except for the relatively infrequent cases where subjects notice the ambiguity, there are no compelling data in the literature suggesting that the strong model (c) should be adopted. There is substantial confusion here about the requirements of the weak and the strong theory, especially with respect to the role of context. Let us, therefore, turn to the next question.

(iii) *How does context affect ambiguity resolution?* We have already mentioned the findings by Lackner and Garrett (164) and MacKay (182) in which biasing information in one ear affects the interpretation of the material presented to the other ear. Normally, disambiguating context either precedes or follows the ambiguous item; some authors have studied the role of such context in the comprehension of ambiguous sentences. Foss and Jenkins (91) distinguish two ways in which prior context can influence the interpretation of an ambiguous word in a sentence. In the so-called Prior Decision Model, the prior context activates (to use Morton's terms) all related logogens: the logogen corresponding to the contextually appropriate reading of the ambiguous word thus gets an activation advantage before the word itself has appeared. This is the weak model (b) above. According to the Choice Point Decision Model both interpretations of the ambiguous word are activated and transferred to working memory: context asserts its effect only later, i.e. it helps to select from among items already present in working memory. This is clearly the strong model (c), which predicts an effect on working memory-load regardless of whether disambiguating context has been presented. With a phoneme-monitoring task, Foss and Jenkins found strong evidence for the latter model: biasing prior

context does not affect monitoring RTs, whereas ambiguity itself quite clearly does so. This supports the strong model. Quite similar results were obtained by Conrad (64), who also found that biasing prior context had little or no effect on her colour-naming task: after the (to be remembered) sentence had been presented, the subject was visually presented with the ambiguous target word (a category name, which was either appropriate or inappropriate to the correct interpretation of the ambiguous word) or a control word. The subject had to name the colour of the printed word. RTs were longer for ambiguous words and their category names than for control words. But again, the effect of biasing context was negligible. Conrad concludes that apparently an independent stage of lexical search exists. But though this finding fully agrees with the results of Foss and Jenkins, Conrad favours 'a theory in which contextual information increases the strength of one or more of the activated meanings of a word to the point that it reaches threshold and becomes conscious'. This, however, is the weak model (b)!—so essentially the same results are interpreted as confirming the strong theory in one case and the weak theory in another. Putting this confusion aside, there are at least four reasons for not adopting the strong theory: Firstly, Foss and Jenkins found that most subjects had been unaware of the ambiguity, and that both the aware and the unaware subjects gave essentially the same results with respect to the critical variables. This is unlikely to have happened if the two interpretations had been in working memory, i.e. conscious (see above). Secondly, Foss and Jenkins correctly state that an effect of context might be critically time-bound, i.e. a disambiguating context appearing either too early or too late might not be effective at the moment of entrance of the ambiguous word. This may also be the case for Conrad's results, who is measuring a fairly long time after the lexical search process (the visually presented test word follows the acoustically presented sentence), and is therefore using a successive measurement technique (see section 2.2) that does not allow direct inferences as to the precise moment of the lexical search. Thus, further measurement might show a context effect, such as found by Morton and Long (201) in a phoneme-monitoring task with non-ambiguous material. Thirdly, there are the general problems with the phoneme-monitoring technique, discussed in section 2.2.1. Finally, as Hogaboam and Perfetti (128) state, one should consider these experimental results from the point of view of the primary and secondary meanings of the ambiguous word. Even if a word has a 50/50 bias, an individual subject will (given the weak theory) allow one (his primary) interpretation to occupy working memory. Measurable effects may occur in just those cases where this (primary) interpretation conflicts with the context which requires the secondary meaning. Summing over subjects will then show an overall effect of ambiguity but no context effect, as found by Foss and Jenkins and by Conrad. Although this fully fits the weak theory (b), it should be noted that the effect of context is supposed to take place in working memory, and not directly through activation of logogens. Hogaboam and Perfetti's experiment is

itself, however, somewhat inconclusive, since asking the subject whether the last word of a sentence was ambiguous is highly unnatural, just as in Olson and Mackay's (206) study, where he subject is asked to find a second meaning after one has already been given. Such conscious search is not typical for sentences containing ambiguous items.

(iv) *When is lexical ambiguity resolved?* If the weak theory of double availability (b) is correct, the effect of lexical ambiguity should come to an end as soon as the one meaning is transferred to working memory. Only in cases where later context requires the alternative meaning will reprocessing become necessary; this is the well-known *garden-path* effect. In all other cases, however, one would expect the effect of lexical ambiguity to be very shortlived. Note that in the alternative strong theory (c), where both interpretations are available in working memory, longer lasting effects could be expected since rejection of one of the two meanings will depend either on later disambiguating information or on the arrival of the end of the clause (as Bever *et al.* (22) propose). The available evidence is limited but in agreement with the weak theory; there is not the slightest evidence for long-lasting effects of lexical ambiguity (apart from garden-path effects). In order to test the duration of the ambiguity effect one can only use simultaneous measurements (see section 2.1). The Bever *et al.* (22) sentence-completion task (on visually presented materials) is a successive measurement; moreover, in spite of their theory, lexical ambiguity had no effect on completion latencies. The only simultaneous measurement experiments are those of Foss *et al.* (88, 91) and of Cairns and Kamerman (37), who used the phoneme-monitoring technique (with all its disadvantages; see section 2.1.1). Foss (88) found a significant 40 msec effect of ambiguity on monitoring latency. However, this was a mixture of lexical and underlying ambiguities. How large the non-significant differences between the two types was is not stated in the paper. Also, the target phoneme was given at different delays after the ambiguous element, but this was not a systematic variable. Foss and Jenkins (91) found a significant effect of lexical ambiguity of 38 msec. The target delay never exceeded two words in this experiment, and it was mostly shorter. The only experiment where target delay was systematically varied was conducted by Cairns and Kamerman (37). At zero delay, where the target phoneme immediately followed the ambiguous lexical item, a significant 20 msec delay was found, but at a two-word delay the ambiguity effect disappeared. The authors conclude that lexical decisions are taken immediately before, and not after, transfer to working memory. In summary, although lexical ambiguity can clearly complicate sentence comprehension, there is little evidence that working memory is involved in this additional load (except for garden-path phenomena or any other task where later recomputation is required for the subject). Rather, a weak theory of double availability can explain all available data: the ambiguous element—as well as earlier context—can activate both logogens, but

normally only one reaches threshold, so that one interpretation becomes available to working memory; this happens very rapidly after the appearance of the lexical item. This interpretation is in good agreement with the general notion that semantic and syntactic processing is very much 'on-line' activity, (see section 4.1.1, and especially the work of Marslen-Wilson).

3.3.2 *Deep-structure ambiguities*

Deep or underlying-structure ambiguities have been the subject of various experimental studies (see 22, 28, 39,49, 88, 89,164, 180, 181, 182, 183,197,262). Here we will limit ourselves to a few remarks about the question of whether in comprehension tasks underlying ambiguities behave differently from lexical ambiguities. Evidence shows that in most experiments this is not so. Apart from cases where subjects are led up a garden path, i.e. spontaneously produce one interpretation when another interpretation is required (see for instance 89), or cases where subjects are asked to find the other meaning (as in 183), underlying ambiguities are not more difficult to handle than lexical ambiguities (see 88,164, 197).

There is, however, some evidence that underlying ambiguities can be distinguished from lexical ambiguities in critical tasks. Bever *et al.* (22) found that sentence completion may be relatively quick if an underlying ambiguous clause (in contrast to a non-ambiguous clause) is presented to the subject. This is not the case for lexical ambiguities. If the presented clause is incomplete, however, deep ambiguity tends to slow down completion. Again, this is not so in the case of lexical ambiguities. This different behaviour is explained by Bever *et al.* by assuming that the hearer 'is carrying out two distinct perceptual operations during presentation of a sentence fragment with an underlying structure ambiguity'. If completion is required before the end of the clause, the subject must choose between two incomplete and independent interpretations. In lexical ambiguity there is less independence, which supposedly makes the choice easier and less forced. If completion is required at the end of a clause, the fact that there are two independent interpretations available increases the chance of finding (at least) one completion. Apart from a need for confirmation of these results (the effects are around borderline significance and no Min F'-test (see Clark, 53) has been used), there is a need for clarification of 'carrying out two distinct perceptual operations'. This would require parallel processing in working memory, and one would like to see a more explicit model of the operations involved. The best available model here is Wanner's (259, 262). Based on an ATN analysis of relative-clause parsing (see section 4.3.2), Wanner and Shiner (262) predict a preference for a direct object interpretation of ambiguous sentences of the sort 'The patient that the nurse brought the doctor, hated rainy days'. In an experiment with control of semantic bias, strong evidence was obtained for this prediction. At the heart of the Wanner model is a

mechanism where syntactic function assignment is postponed until the relevant information arrives. Therefore, contrary to the Bever *et al.* notion, Wanner and Shiner suppose that neither are both interpretations computed nor a single one: rather none is computed; the decision is postponed, and since postponement requires the storage of information, it is a load-increasing process.

3.4. Constituent and clause structure

In this section we will limit ourselves to presenting some of the major opinions about perceptual units of segmentation to be found in the literature. Since many of the studies on segmentation employed the click-localization technique, we should keep in mind the methodological difficulties proceeding from this technique (see section 2.2.1). Also, the recall techniques (see section 2.2.2) that seem to support the psychological reality of clause structure do not necessarily show that this reality is perceptual. Relatively little simultaneous measurement has been used for segmentation units (see, however, 3,24,77,186). In the search for the psychological reality of linguistically defined segments, major constituents were among the first candidates of investigation. Fodor and Bever (79) claimed perceptual reality of major constituents on the basis of click-location results. Bever, Lackner, and Kirk (25) modified this view in the sense that only those constituent boundaries which related to deep-structure sentoid boundaries were able to attract clicks. Opposition to this position came from Chapin *et al.* (48) and Toppino (246), who tried to show that surface boundaries not relating to sentoids could attract clicks. Fodor *et al.* (80) challenged Chapin *et al.*'s critique on the basis of an analysis of the linguistic material which they had used for their experiment. After a review of the major literature on clause structure, Fodor *et al.* (80) conclude that 'surface constituent boundaries which correspond to junctures between sentoids define the potential points of perceptual segmentation of sentences; whether any such point is in fact taken as the boundary of a perceptual unit may depend on a variety of other structural features' (p. 339)—a rather cautious statement. An important addition to the search for reality of linguistic segments was an effort to determine the functional role of such segments in the perceptual event. Fodor *et al.* (80) propose the theory that 'As the sentence is received, it is assigned to a short-term store where the fragments that constitute each of its sentoids are collected together. Material is dismissed from this storage as soon as it can be assigned to a completed sentoid. It is because each sentoid is dismissed from this store en bloc that the clause functions as a unit of speech perception' (p. 342-343). Thus, at the end of the clause the materials are assembled and assigned relations according to the sentoid structure. There have been reactions against both the sentoid interpretation of segmentation units and the clause-by-clause processing theory. In spite of the fact that Fodor *et al.*'s most recent statement about segmentation

units was quite modest, Tanenhaus and Carroll (243) attacked the position that the segmentation unit could be *structurally* defined; in its place they propose the notion of 'functional clause', which expresses, among other things, so-called 'functional completeness': A clause can be a candidate for a segmentation unit only if it expresses a complete set of grammatical relations (as, for example, in '*After Mary finished the cake, she took an apple*', and not in '*After finishing, Mary took an apple*'). It was mainly Marslen-Wilson (184, 185, 186, 187, 188) who attacked what he called the 'staggered serial model', i.e., the notion that syntactic analysis would only take place after a whole syntactic unit, such as a phrase or clause, had been gathered. His shadowing and word-monitoring experiments clearly show that syntactic analysis takes place right from the beginning of the clause, inter-actively with phonetic and semantic analysis. A more formal model of such on-line processing has been proposed by Wanner *et al.* (259, 260).

What is lacking at the present moment is a restatement of the relationship between segmentation units and working memory. If clausal 'reality' is not in fact caused by end-of-clause interpretation plus release from working memory, are we then bound to dismiss clausal reality after all, or should we seek its cause among other memorial functions? The contributions of Marslen-Wilson *et al.* and Carroll *et al.* to the present volume shed a wholly new light on this issue.

3.5. Pictorial context

Many studies in sentence perception have used pictorial variables in order to study how sentence comprehension interacts with non-linguistic perceptual context. Almost all of these studies were of the verification type (see section 2.2.5). Table 1.2 gives a summary of the linguistic and pictorial variables used in these sentence-verification studies, as well as of the major results obtained.

What are the major principles which govern most of these results? All theoretical accounts (45, 54, 56, 58) are based on the assumption that the sentence is internally represented in abstract propositional format. The same is assumed about the internal representation of the picture. Verification proceeds through a serial process of comparison between these two internal representations, with the order of comparison being determined by the propositional hierarchy. Reaction times are determined by the number of operations to be performed in the serial comparison. The number of operations critically depends on the number of mismatches between the two propositional representations: this is known as the *congruence principle* (Clark, 54).

In general, the serial nature of sentence/picture verification processes finds substantial support in the literature; see, however, our earlier remarks in section 2.2.5. Though the congruence principle seems to find similarly strong support, it should be noted that this support depends on (a) the representation one chooses for sentence and picture and (b) the order of comparison one assumes. With

Table 1.2 Sentence-picture verification experiments






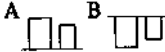
Authors	Linguistic variables	Pictorial variables	Order	Main results
Banks, Clark and Lucy (1975)	1 Which balloon/yo-yo is higher/lower? Which string is longer/shorter?	A  or B  and varying string difference	S->P	Balloon/yo-yo x A/B interaction (congruency) No congruency x length difference effect
	2 Same	A  or B  constant string-length (1 cm)	S->P	Congruency effect. No congruency x height diff. effect
Cary, Mehler and Bever (1970)	1 Set-list of four Adjectival (A) (they are incoming signals) or Progressive (P) (they are unearthing diamonds) sentences, followed by ambiguous test sentence (they are lecturing doctors) Ear of entry	Pictures for which both interpretations are true (TT) only one (TF or FT) or both false (FF)	P->S	For ambiguous sentences A/P x ear of entry interaction for reaction times. Interpretation bias according to set. No true/false effect
	2 Same	Same	P->S	No interpretation bias
Carpenter and Just (1972)	Type of quantifier: minority/majority, few/many (of the dots are red/black)	A 2 black, 14 red dots B 2 red, 14 black dots	S->P	Interaction between quantifier and locus of fixation (i.e. red or black subset)
Carpenter and Just (1974)	Negation and colour name: It is (n't) true that the dots are (n't) red/green	16 dots of 1 colour: red, green or black	Simultaneous	No negation x colour congruence interaction. RT increases with number of constituent comparisons
	2 Same	Same	S->P	Same

Table 1.2 (continued)

Authors	Linguistic variables	Pictorial variables	Order	Main results
Clark, Carpenter and Just (1973)	1 Which is taller/shorter Which is deeper/shallower		Simultaneous	RT (tall/short) < RT (deep/shallow). RT (tall, deep) < RT (short, shallow). No interaction simultaneous/S -> P with any variable
	2 Same		S->P	RT (tall/short) < RT (deep/shallow). Interaction A/B x (tall, short)/(deep, shallow)
	3 Same	Five boxes 'looked into' from different angles	S->P	Interaction box-type x (tall, short)/(deep, shallow)
	4 Dimensionality of comparative: which is taller/shorter/bigger/smaller? which is wider/narrower/bigger/smaller?	Two rectangles, one being a square differing only in width. Two rectangles, one being a square, differing only in height	S->P	Two-dimensional terms quicker, one-dimensional terms slower for square
	5 Same	Same, but no square rectangles	S -> P	One-dimensional terms quicker, two-dimensional terms slower for the more rectangular figure
Clark and Chase (1972)	1 Star (plus) is (n't) above (below) plus (star)	Two pictures: star above or below plus	Simultaneous	RT (above) < RT (below) Aff < Neg True < False. No interactions
	2 Same	Same	S -> P P -> S	Same, plus interaction pos/neg x true/false S -> P quicker than P -> S

	3	Same	Same, instructions <i>to attend bottom/top of figure, or whole figure</i>	P->S	RT (whole) < RT (bottom/top). No interactions
Clark and Chase (1974)	1	Dependent linguistic variable (spontaneous descriptions)	Relative prominence of objects, i.e. star above (below) line (prominent) or circle (non-prominent)	P->S	Preference for above. Above-bias for non-prominent picture independent of orientation; for the prominent picture dependent on orientation
	2	Star (line) is (n't) above (below) line (star), or star (circle) is (n't) above (below) circle (star)	Same	S->P	No above/below x star/position interaction
Clark and Lucy (1975)		Ten different types of request to either or not ('polarity') colour circle blue or pink	Blue or pink circle	S->P	Truth x polarity interaction
Cormish(1971)		The circle is not all blue (red, green, yellow)	Circle with different differently large sectors of red (etc.)	S->P	Non-monotonic sector size effect on verification RT
Flores d'Arcais(1974)	1	Comparatives with 'marked' vs 'unmarked' adjectives (<i>tall vs short, etc.</i>). Orally <i>presented</i>	Pictures with two persons, animals, etc. differing on the antonym's dimension	S->P	Longer RT for marked than for unmarked. True quicker than false
	2	Comparatives with marked vs unmarked adjectives. Orally presented	Two long test lines among small ones, <i>or</i> two short test lines among long ones. Test lines differed in size. Similarly for	S->P	Longer RT for marked than for unmarked. True quicker than false. No effect of size of test lines (etc.)

Table 1.2 (continued)

Authors	Linguistic variables	Pictorial variables	Order	Main results																								
		positions of weighing balance, more or less filled glasses, smaller or larger blocks																										
	3 Same as 2	Same as 2	P->S	Same as 2																								
	Marked or unmarked cue-adjective presented before picture. Sentences with larger vs smaller, or less large vs less small	Large and small circle	P->S	Same as 1, plus cue-adjective x sentence adjective interaction																								
Foss, Bever and Silver (1968)	Ambiguous or non-ambiguous sentences	Pictures for which sentence is either true or false	S->P	Longer RTs for ambiguous sentence only if picture presents unexpected meaning																								
Glucksberg, Trabasso and Wald (1973)	<table border="0"> <tr> <td>{</td> <td>Car</td> <td>{</td> <td>hit</td> <td>{</td> <td>fence</td> </tr> <tr> <td>}</td> <td>truck</td> <td>}</td> <td>passed</td> <td>}</td> <td>pole</td> </tr> <tr> <td>{</td> <td>bus</td> <td>{</td> <td>pulled</td> <td>}</td> <td>tree</td> </tr> <tr> <td>}</td> <td>train</td> <td></td> <td></td> <td></td> <td></td> </tr> </table> <p>In either active or passive (A/P) form. Thus, some reversible, some irreversible (R/I)</p>	{	Car	{	hit	{	fence	}	truck	}	passed	}	pole	{	bus	{	pulled	}	tree	}	train					Corresponding pictures	P->S S->P	For true sentences: A/P x R/I interaction for S->P only. For false sentences: A/P interaction with locus of mismatch
{	Car	{	hit	{	fence																							
}	truck	}	passed	}	pole																							
{	bus	{	pulled	}	tree																							
}	train																											
Hornby (1974)	1 Cleft/pseudo-cleft sentences	Picture true, or picture false with respect to resp. focus or proposition	S->P	More errors for false/pre-supposed than for false/focused																								

		Standard/cleft/pseudo-cleft sentences (S/C/P) Sentence voice (A/P)	Same	S->P	More errors for false/pre-supposed. A/P-effect (for errors) S/C/P-effect
Just and Carpenter (1971)		Three types of negatives: syn-not (not, none, no) syn (few, scarcely, hardly) sem (a minority, a small proposition etc.)	4 x 4 array of dots. All 16 same colour or two with different colour	Simultaneous (sentence left)	RT(Neg)>RT(Aff) Aff/Neg x True/false interaction for syn-not and syn. Type of neg x Aff/Neg x True/False interaction
	2	Same	Same	P->S	Aff/Neg x T/F interaction only for syn-not
	3	Two types of negatives: syn and sem	Same, but instruction to code colour of larger or smaller subset	P->S	For 'code larger subset' same interactions as in expt. 1; for syn and sem. For 'code smaller subset' inverse interaction
Krueger(1972)	1-4	Is (n't) north east south west	Circle above, below, left of, right of sentence	Simultaneous	Interaction Neg/Aff x location of circle
Olson and Filby (1972)	1	The $\left\{ \begin{matrix} \text{car} \\ \text{truck} \end{matrix} \right\}$ hit the $\left\{ \begin{matrix} \text{truck} \\ \text{car} \end{matrix} \right\}$ Active/Passive	Picture for which sentence true or false. Two directions of movement, two instructions: attend subject/object of action	P->S	RT (Act) < RT (Pass) Instruction x A/P interaction
	2	Same	Same, but other instructions: attend car/truck	P->S	RT(Act)<RT(Pass) RT (True) < RT (False) Instruction x position (car/truck) interaction

Table 1.2 (continued)

Authors	Linguistic variables	Pictorial variables	Order	Main results
	3 Same	Presentation of one vehicle, followed by presentation of the other higher or lower on the slope (focus)	P->S	RT (Act) < RT (Pass) RT (True) < RT (False) Focus x Act/Pass interaction
Roncato and Sonino (1976)	'A $\left\{ \begin{array}{l} \text{inside} \\ \text{outside} \end{array} \right\}$ B', where A, B are different figure names. Visual presentation	Embedded or juxtaposed figures (circles, squares, triangles), which corresponded or did not correspond to figure names in sentence	S->P P->S	S->P faster than P->S Faster RT if smallest object is mentioned first. Faster RT if figure does not correspond to name. True faster than false
Wannemacher (1974)	1 Active/Passive (A/P) Reversible/Irreversible (R/I)	Pictures such that they create mismatch for subject, verb, or object of sentence or combinations thereof ('treatments')	P->S	Different RTs for 'treatments' (object mismatch slowest). A/P-effect, R/I-effect Interactions: A/P x R/I A/P x treatment A/P x R/I x treatment
	2 Same	Same	Simultaneous	Same
Wannemacher (1974)	1 As in Wannemacher (1974)	As in Wannemacher (1974)	S->P	Interaction R/I x treatments
	2 Same	Same, but only the three one-component mismatch types	S->P	Same i.e. serial self-terminating comparison: logical subject -> verb -> logical object

respect to the representation of sentence and of picture, different authors give different solutions. These are partly notational variants, but for another part important empirical issues are involved. The question is whether the same sentence (resp. picture) can be represented in different formats, dependent on task and practice, and on the prior presentation of the picture (resp. sentence). It is indeed the case that these factors can affect the way in which the subject codes the information. Examples of task and practice effects have already been mentioned in section 3.2.2, especially Trabasso *et al.*'s (248) finding that subjects apparently represent the sentence 'The ball is not red' by 'GREEN (BALL)'. But this occurs only if there is no more than one (colour) alternative, and after the subject has had some practice. This is called the *recoding strategy* (45, see also 150); it is assumed to have taken place if the true negative (TN) is judged faster than the false negative (FN). One wonders whether recoding is not a mere laboratory artifact. Interdependence between sentence and picture representations has been shown to exist by Clark, Carpenter, and Just (56); it has especially been demonstrated that the representation of the picture can be affected by prior presentation of the sentence.

Recoding and interdependence give the theorist quite a lot of leeway in modelling his data. One would like to have independent evidence that a particular situation leads the subject to recode (i.e. independent of the finding that $TN < FN$); and similarly for interdependence. But otherwise, there is nothing against invoking such principles. Tanenhaus *et al.* (244) in a sharp critique of Carpenter and Just's (45) paper, to which we will return shortly, criticize the arbitrariness in the use of a once-chosen format of representation. They refer especially to inconsistencies in the use of an affirmative embedding predicate (AFF) in the internal representations, inconsistencies which do indeed affect predicted reaction times. More agreement, however, seems to exist with respect to the order of comparison one must assume. The principle seems to be 'inside out': the embedded propositions are verified before the embedding propositions.

Let us, finally, return to the Tanenhaus *et al.*'s (244) critique of verification studies, which was mostly addressed to Carpenter and Just (who reacted to it in 46). Apart from technicalities relating to the question of whether the choice of internal representations is ad hoc, the exchange of views is primarily demonstrative of a clash in aims. The Tanenhaus *et al.* paper demonstrates what we have called the linguistic aim in section 1: to explain how the hearer uses his knowledge of the language to encode sentences in terms of linguistic structural descriptions. It is argued that the verification approach does not tell us how the verification representations are derived but only how the comparison is done. The latter is said to be of marginal interest for a theory of sentence understanding. Carpenter and Just, on the other hand, place themselves much more in the communicative framework: how does the listener use linguistic and non-linguistic information in order to answer questions? For this, it is highly

important for the listener to find out the referents of what is said, and this process is studied in the verification literature. So far, the discussion has little theoretical import: if one is mainly interested in studying the derivation of internal representations, one should not join the verification club. If, however, one's aim is to find out how linguistic information is used vis-a-vis perceptual or encyclopaedic knowledge, then verification studies can be quite useful. An important theoretical issue does arise, however, if one wants to realize both aims at the same time. Tanenhaus *et al.* assume that the verification representation which is used in the verification process, and which is task-dependent, is in its turn derived from a prior, and much more general-purpose representation. This is, of course, the structural description mentioned earlier. Here is an interesting empirical assumption: deriving a structural description is a first step in deriving anything else (a conceptual structure, a verification representation). This issue will be taken up again in section 4.1.1.

3.6. Non-pictorial context: Given and new information

In section 2.2.5 Various studies were listed in which a sentence had to be understood (verified), given another sentence or certain encyclopaedic knowledge. In most respects these studies were concerned with the same theoretical issues as the sentence/picture verification studies (i.e. congruence, markedness, etc.) Here we will limit ourselves to mentioning some studies of a different theoretical flavour, namely those deriving from Clark's notion of the given/new contract. Clark and Haviland (59) starting from Grice's (114) notion of a 'cooperative principle' in conversation, propose the existence of 'cooperative' information-exchanging strategies used by speaker and listener. At any moment the speaker presupposes certain items of information to be available to the hearer. These are the 'givens' to which the speaker can then add new information. The speaker will use special linguistic devices to express what he supposes to be given, and what he intends to be new. The listener, in his turn, will match this strategy by using what is linguistically marked as 'given' in order to locate the appropriate information in memory. He can then proceed to add the information linguistically marked as 'new' to the existing data structure. Clark and coworkers did various studies to show that violations of this given/new contract lead to complications in the comprehension process (59, 126, 236). A typical paradigm is to switch the linguistic marking of given and new, as in the following examples:

Context: Michelle ignored Sam, who was waiting at the bar. Turning her back, she began talking to a group of strangers. Sam was wild with jealousy.'

Sentence (a)r. 'The dancer *who disregarded Sam*, motioned to Jack.'

Sentence (b). 'The dancer who motioned to Jack, *disregarded Sam.*' If the context is followed by sentence (a), everything is according to the contract. The relative clause expresses (as it should) information already given in the context.

In (b), however, the relative clause (linguistically marking information as given) in fact expresses new information. Clark found longer reading and comprehension times required for (b) than for (a). In (b) the listener is forced to switch the roles of given and new in order to integrate the sentence with the earlier context.

Though the given/new notion has a long tradition in linguistics, these are the first studies where its role in sentence processing is studied. It appears from two other studies (Hornby, 137; Just and Clark, 151) that given (presupposed) information is harder to verify than new; this is in accord with the general theory. (See especially Clark's contribution to this volume.)

3.7. Some conclusions

In the preceding review of linguistic and other independent variables used in sentence-comprehension studies, some of the major observations were:

(i) There is a gradual shift from considering prosody as a nuisance variable to a genuine interest in the comprehension of spoken (as opposed to written) language.

(ii) The effects of linguistic complexity on processing are not uniform. Not only are some complexity variables, such as verb structure, only marginally effective, but in many cases (voice, negation) the effect depends on the communicative setting in which the sentence is used.

(iii) Lexical ambiguity almost always increases the processing load of a sentence. However, in all but garden-path and similar cases, working memory seems not to be involved in this.

(iv) It is still undetermined which perceptual mechanisms cause clause-like entities to function as processing units.

(v) The controversy about the usefulness of verification studies for a theory of sentence understanding signals a clash of aims in the psycholinguistic literature.

4. THEORETICAL CONSIDERATIONS

In section 1 it was pointed out that there is little disagreement about the central importance of a theory of sentence understanding. There is also little controversy about the general aim of a theory of sentence comprehension, which should be able to explain the processes involved in understanding sentences used in ordinary contexts. But from this point on every single step is controversial. There is a double tension to be noted in the literature. Firstly, there is a disagreement concerning the definition of understanding. Fodor *et al.* (80), Tanenhaus *et al.* (244), and others reserve the term for what I will call *immediate linguistic awareness* (ILA): the almost instantaneous awareness of the underlying prepositional structure of the sentence. Others (Schank, 226, Bransford, 30,34) use the

term to denote the process mediating between syntactic input and the completion of a linguistic or non-linguistic task (memory task, paraphrase task, verification task, etc.). This difference is not a matter of the mere restriction of empirical domain. That would be the case if the latter type of processes would necessarily contain the former type as proper subpart. This is, rather, an empirical issue. The question is whether in all sentence comprehension there is an initial stage during which the underlying linguistic structure of the sentence is derived. We will call this the *immediate linguistic awareness (ILA) hypothesis*. Closely connected to this issue is the question of whether this initial stage, if appearing at all, corresponds to the existence of a separate memory mechanism, a short-term store, or whether one has to assume a continuum of depth of processing. This is the subject of section 4.1.

The second tension concerns the measure in which the theory should be formalized and the type of formalization to be chosen. There is the full range, from completely unformalized theories such as the one proposed by Bransford and his associates (e.g. 31, 34) to completely formalized ones such as Wanner's ATN models (259, 260, 261). Usually the formalized theories either concern a very limited aspect of comprehension or, if more generally applicable, are still very limited in the range of phenomena for which they have been worked out and empirically tested. They are discussed in section 4.3. That section is preceded by some remarks on the role of context in comprehension (section 4.2).

4.1. Immediate understanding and depth of processing

4.1.1 Immediate awareness and task dependency

All experimental studies on sentence perception involve some explicitly stated or implicit task for the subject. This is not entirely artificial: in ordinary language perception also the hearer listens for something. Dependent on the situation, this can be almost anything: whether the voice is male or female, whether there is an implicit request, whether the sentence answers an earlier question, etc. If there is any truth to the claim that laboratory studies of sentence perception are *ipso facto* unnatural and thus meaningless, it cannot be ascertained simply because of the fact that specific tasks are set in the laboratory. Subjects participating in laboratory experiments never tell us that this is not the real world. The only argument for the claim could be that the task dependency of linguistic processing is not sufficiently studied in itself and that too general conclusions are drawn from too limited task environments. Various studies have show that sentence processing is critically dependent on the task requirements (4, 41, 95, 96, 110, 197, 199, 252), and one could conjecture that sentence understanding is idiosyncratic to the task and that generalizations from laboratory studies are not allowed. Although some authors come close to this position, there is a very general tendency in the literature to gravitate towards another solution. We called this the 'immediate linguistic awareness (ILA)

hypothesis', and a general formulation of it could be as follows: *The initial part of any sentence comprehension consists of deriving a complete underlying representation of the sentence.* The ILA hypothesis involves two strong empirical claims. First, the initial stage is identical for any comprehension task, and thus task-independent. Second, the output of the initial stage is a complete underlying representation. Since the derived representation should be general purpose (i.e. suitable for any task), it should be an almost complete linguistic analysis of the sentence, containing all semantically relevant information. Authors who adhere to the ILA hypothesis differ substantially in the characterization of this initial internal representation, as will appear from the following enumeration:

The immediate perception hypothesis has its roots in what in the 1960s was called the *coding hypothesis*: a sentence was supposed to be coded in memory in the form of its deep structure (see Johnson-Laird, 143; Levelt, 168; and Fodor *et al.*, 80 for detailed statements and analyses of the coding hypothesis). The hypothesis was proved to be false (see especially Bransford *et al.*, 30, and Fodor *et al.*, 80), but the idea that the underlying structure of the sentence becomes available at some time during sentence processing is still rather generally endorsed. Clark (51) stresses the importance of 'the base strings': 'they constitute the essential part of the interpretation of the sentence and should therefore play an important part whenever the interpretation is needed at a later time'. Clark (54) proposes a 'deep-structure hypothesis' to characterize the internal representation of the sentence. Even if the subject's task is such that he should ignore the literal interpretation of the sentence—as is the case in coping with conveyed requests—Clark and Lucy (60) have shown that this literal interpretation is nevertheless derived first, thus testifying to the empirical validity of the task-independency of the immediate linguistic awareness. A similar claim is made by Cutler (68).

Garrod and Trabasso (105) also express the deep-structure hypothesis: 'the prior sentence is held long enough for a deep-structure representation to be derived and stored in long-term memory'. They add the claim that surface information may be stored in LTM as well, which is further supported by a study by Anderson (5). Fodor *et al.* (80) argue extensively that the internal representation of a perceived sentence coincides with the linguistic structural description. If empirical data show that the structural description proposed by transformational theory is not available to the subject, Fodor *et al.* (81) would prefer that the linguists change their theory rather than have the psychologists forfeit this principle of identifying internal representation and linguistic description. It should be noted, however, that Fodor (78) argues for a much shallower immediate processing: 'The representation of a sentence that must be recovered in understanding it is relatively *unabstractly* related to the surface form of the sentence' (p. 154). Another example of the ILA hypothesis can be found in Tanenhaus *et al.* (244). As discussed in section 3.5, these authors severely criticized Carpenter and Just's (45) processing models derived from

sentence/picture verification studies mainly because the sentence encodings (i.e. internal representations) which are proposed are different for every study, rather arbitrary in nature, and extremely task-dependent. In section 3.5 we mentioned that the theoretically important part of the criticism is that a model of sentence understanding should adhere to the ILA-hypothesis. The authors propose a two-stage model: 'First, they (the subjects) understand the sentence, developing a representation for it. Then they extract a *verification representation* from the comprehension representation. The form of the verification representation depends on the particular verification task'. The first stage is what we called 'immediate linguistic awareness', and it is task-independent. The exact nature of the comprehension representation, however, remains unspecified in this paper; it is even left open whether it can be linguistically defined. Forster and Olbrei (85) take the risk of making themselves extremely vulnerable at this point: 'when syntactic analysis is required by the task conditions, it is executed without regard for the meaning of the sentence. This conclusion tends to suggest that there must be a psychologically real level of description which is purely syntactic, and quite independent of the semantic representation'. The statement does not imply the two-stage model, but Forster and Olbrei's experiments (see also Gamlin, 99) would make a strong case for Tanenhaus *et al.*'s comprehension representation to be syntactic. Miller (193), inspired by the work of Davies and Isard (70), proposes a stage of immediate awareness, by disconnecting 'the understanding of a sentence from any actions it might entail'. One can refuse to *obey* a command, but not to *understand* a command. In computer terms the first stage ('understanding') consists of compiling the program, with natural language functioning as a higher-level programming language. The output, then, is a set of lower-level sub-routines which may or may not be executed, dependent on the task-environment. If this output is linguistically characterized by means of grammatical rules, then these rules are abstractions from the computer (resp. language understander); they are not components of it.

Thus, although all these authors assume the existence of a task-independent initial phase, the character of the initial internal representation differs widely: a syntactic deep structure (semantically uninterpreted), a propositional hierarchy, a much shallower surface-related representation, and a set of (semantic, conceptual?) subroutines have all been used as representations.

The different variants of the ILA hypothesis are empirically interesting only if they are explicit about the structure of this internal representation. The existence of a task-independent representation can only be demonstrated if one is fairly precise about the nature of that representation. I am not familiar with any empirical evidence (either in favour or against) which is of this strong nature. The advantage of these theories is, however, that they are stronger than alternative theories in which understanding is completely task-dependent. Stronger theories are more stimulating, even if they are quite probably false.

Another major obligation of these theories is to describe the processes by which this initial internal representation is derived. In this respect the situation is plainly poor. Clark (54) does not propose a parser for this deep-structure hypothesis; he is clearly more interested in later task-dependent aspects of comprehension. Fodor *et al.* (80) propose and empirically test some so-called strategies. One of them is the *canonical sentoid strategy*: whenever the hearer encounters the surface sequence NP-V(-NP), he or she assumes that these items are, respectively, subject, verb and object of a deep sentoid. Although there is empirical evidence for the correctness of some of these strategies, or parsing principles (see also Kimball, 156), there is a major lack of clarity about how these strategies interconnect, whether they are hierarchically organized, what happens if two strategies lead to opposite results, at which point in the sentence a strategy is called for—in short what sort of control structure is involved in sentence perception.

One way to explore these questions more systematically is to cast strategies in the form of augmented transition-network (ATN) models; we will return to this in section 4.3.2.

As far as derivational processes have been proposed, one is left with the impression that the hearer accumulates surface information during presentation of the clause, and that by the very end of the clause there is a flurry of internal processing, leading more or less immediately to determination of the underlying structure. As we have seen in section 2, such theories are mostly based on successive measurement, which supplies little information about the precise structure of the knowledge which is accumulated during the input of the sentence.

There are, moreover, theoretical and empirical arguments against this view of processing. If one looks into Riesbeck's (212) parser, which is intended to be psychologically realistic, it is clear that processing at *all* levels (phonological, syntactic, semantic, conceptual) takes place right from the beginning. Parsing is not hierarchical in the sense that syntactic operations precede semantic and conceptual operations; it is heterarchical: a process at one level is able to call a process at any other level without a hierarchical transfer of control (see Haggard, 116; Winograd, 268).

One of the most challenging of empirical issues is to demonstrate the psychological reality of such distributed control. Apart from the ATN-work (see section 4.2), the shadowing experiments by Marslen-Wilson (184,185,186,187) should be mentioned here. Phonological, syntactic and semantic effects on shadowing performance could be demonstrated to be taking place already at the very beginning of the sentence, even if shadowing delay were not more than one syllable. A similar simultaneous availability of phonological and semantic information could be shown by measuring monitoring latencies: reaction times for rhyme-monitoring were the same as RTs for (semantic) category-monitoring. Also, Mistler-Lachman's experimental results (see section 4.1.2)

suggest that a full syntactic analysis, if performed at all, may follow semantic analysis. If sentence processing is of this heterarchical nature, there is reason to doubt an immediate derivation of deep structure (as in Clark, 54; Garrod and Trabasso, 105; Fodor *et al.* 80, Tanenhaus *et al.*, 244). If conceptual parsing is taking place from the first word on, it seems to be rather awkward, as Marslen-Wilson (186) remarks, 'that the initial target of the processing system is a deep-structure representation of the input'. This doubt is also expressed by Bransford and McCarrell (34): 'It seems reasonable to assume that the surface structurally similar sentences *The house squeaked* and *The window squeaked* are understood differently *not* because one first discovers a deep structure and then interprets the meaning of the lexical items, but because one's knowledge of the entities and events in the sentence forces different semantic interpretations to be made'. Other statements of this so-called constructive view of sentence comprehension can be found in 14, 30, 94, and 136. However, there is certainly not enough empirical evidence now to reject all versions of the ILA hypothesis, especially those which are more semantic in nature. Much more precise experimentation is necessary to refute the existence of a task-independent first processing stage and of a general-purpose internal representation.

4.1.2 Stores versus levels of processing

An attractive though not essential processing model for variants of the ILA hypothesis is a multi-store model, consisting of at least a short-term and a long-term store (STS, LTS). The STS would contain a verbatim representation of the sentence or clause, which is maintained until the parsing operations involved in deriving a deep structure have taken place. The resulting deep representation is put into LTS, to be available for further processing. This is essentially the model proposed by Jarvella (141) and Fodor *et al.* (80). Garrod and Trabasso (105) and Anderson (5) have argued that surface information may as well be sent to LTS; they maintain, however, that STS will not contain deep information. Since decay of the memory trace is quicker in STS than in LTS, variants of this successive-stores model can explain the longer persistence of propositional over verbatim information.

An alternative view has been expressed by Craik and Lockhart (66). These authors reject the multi-store model but try to explain the longer persistence of semantic information from the deeper level of processing necessary for deriving this information. The store model is initially replaced by a stagemodel. If extracting meaning presupposes the recognition of words, an early stage of processing would involve word recognition, whereas meaning extraction would take place at a later stage. These stages cannot be identified by pointing to successive stores with their own preferred internal code and their own decay time. The character of these stages is only stimulus- and task-dependent. Later in the same paper the authors qualify the stage terminology; they do not mean

that logically prior analyses necessarily take place at earlier stages, and they refer to Savin and Bever (222) who showed that the syllable is usually more quickly recognized than is its initial phoneme. 'Spread' of encoding would be a better term than 'depth' of encoding, and the stages in fact form a continuum of increasing spread. This notion of spread of encoding agrees with the earlier mentioned idea of heterarchy: there is no hierarchical ordering of processing, neither in terms of stores, nor in terms of logical order.

Experimentally, depth of processing has been manipulated in two ways: by task and by stimulus. Most studies manipulate by task. The best examples are found in the studies by Mistler-Lachman (197, 198, 199). In 197 she presented sentences to subjects with or without ambiguity. Two of her tasks were: judgment of meaningfulness and constructing a good continuation of the sentence. She measured reaction times in the two situations. Presumably judgment of meaningfulness does not require ambiguity resolution: processing can be relatively shallow. Deeper processing is required for the other task, and presumably ambiguity should be resolved. RTs on sentence continuation indeed turned out to be longer for ambiguous sentences, whereas ambiguity did not affect the time required for making a meaningfulness judgment. In 198 Mistler-Lachman showed that the thus demonstrated shallower processing in meaningfulness judgments leads to a quicker decay of the memory trace, as predicted by Craik and Lockhart. Gamlin (99) also used tasks as a variable: question-answering probe latency with similar results.

One can also affect the level of processing by manipulation of the stimulus material. Marslen-Wilson and Tyler (188) repeated Jarvella's experiments (see section 2.2.2) but varied the quality of the text: it could either be normal text, so-called 'syntactic prose', text which is semantically uninterpretable but syntactically correct, and random word-order text. The Jarvella-effect, i.e. less accurate recall of the penultimate clause, was substantially more marked for syntactic prose than for normal prose. Syntactic prose prevents deep (semantic) processing, and since recall is a function of depth (Craik and Lockhart), the theory correctly predicts this substantial forgetting of syntactic prose. It is not obvious how a two-store model would make the same prediction. Although task-dependency is naturally handled in the depth of processing model, the existence of an initial task-independent phase in sentence comprehension is not contradictory to this view: the dependability of multi-store models remains an empirical issue. For a critical discussion of the sort of information that might discriminate between single- and multi-store models, see Wickelgren (264).

4.2. Comprehension in context

Language understanding is largely a matter of finding out the speaker's intentions. This review was originally to have included a section on the various means and processes by which the listener can determine intentions on the basis

of literal meaning of what is said plus the context in which it is said. I would have discussed empirical work on given and new information, conversational postulates, anaphoric and definite reference, and major attention would have given to the work of Clark and coworkers. In view of Clark's own very full treatment of these issues in this book, my own review would be superfluous; it is therefore omitted. It is only for referential completeness that the bibliography contains numbers 14, 28, 31, 32, 33, 59, 60, 73, 94, 126, 135, 137, and 151, which are relevant to the issues under concern.

4.3. Formal theories of sentence understanding

4.3.1 Transformational models

Transformational models of sentence perception emerged during the sixties as a psychological answer to the newly developed transformational linguistics. Extensive reviews of these models and their empirical tenability can be found in the literature; the reader is especially referred to Levelt (170), and to Fodor *et al.* (80). Both reviews show that initially transformational grammar was adopted as an isomorphistic model; this isomorphism had two related but separable characteristics: the first is the *derivational theory of complexity* (DTC), the second is the so-called *coding hypothesis*. DTC means that the processing of a sentence simulates the transformational derivation of the sentence on a micro-scale: each rule in the grammar has a one-to-one relationship with a particular psychological operation. This can still be done in very different ways, such as inverting transformations (Petrick, 208) or analysis-by-synthesis (Halle and Stevens 122).

If rule-for-rule isomorphism exists, one would also expect input-output isomorphism, i.e. at some early stage in comprehension the hearer is supposed to have available a surface-structure-like representation of the sentence, whereas after applying the transformation-related rules, this has been replaced by a deep-structure-like representation. This latter view is called the coding hypothesis (for different senses of the coding hypothesis, see Levelt, 170). Of course the latter can be valid, without DTC being valid, and that is the position taken by Fodor *et al.* (80). Carroll and Bever (47) put it this way: '(1) There is a variety of evidence for the psychological reality of linguistically defined surface and deep-structure representations of sentences. (2) There is no consistent evidence for the perceptual reality of transformations as perceptual processes'. This conclusion is not drawn in Levelt's (170) review, where both DTC and deep-structure output theories are characterized as empirically untenable. The point is sharply made by Wanner (258) in his review of the Fodor *et al.* book: 'Why should we believe that deep-structure phrase markers are determined during comprehension, when the characteristics of deep structure are partially selected just in order to simplify the operation of the transformational rules, which themselves

lack psychological reality? So far as I can see, there is no compelling reason to adhere to such a belief. One can search in vain through the deep-structure experiments which FBG review without finding any which test the psychological reality of the transformationally motivated aspect of deep structure.'

Miller and Johnson-Laird (194) make a similar point by stating that 'Fodor, Bever, and Garrett assume, however, that the decoding processes compute the same structural description that a grammar does (p. 369), as if parsing were an end in itself, rather than an abstraction from (or trace of) the computations required to translate a sentence into an executable routine'. We will return to this point at the end of the next section.

4.3.2 Perceptual strategies and Augmented Transition Networks

Given the theoretical framework of Fodor *et al.*, the important psychological problem is to define and test psychological operations which can derive underlying structures from surface forms. These operations need not be linguistically justifiable (see 20), and they are not necessarily foolproof. One example is the canonical sentoid strategy, mentioned above in section 4.4.1. There is no doubt that this approach has partly freed psycholinguists from their linguistic captivity: one could resume the search for interesting perceptual operating principles without being hampered by existing linguistic rule systems. This openness has led to a variety of highly interesting empirical observations, for which the reader is referred to the Fodor *et al.* book.

At the same time, however, this approach has created a theoretical vacuum. As already noted in section 4.1.1, the control structure of the strategies approach remains unspecified; there is no theoretical basis for distinguishing possible from impossible strategies or for limiting the number of strategies to an 'optimal', or canonical set. This vacuum is not easily filled so long as the aim of parsing is defined in deep-structural terms: formal objects which have no relevance for understanding will have to be created by the parsing process, whereas linguistic constraints are ignored in defining the perceptual operations. Augmented Transition Networks (ATNs) may provide the formal means of leaving the vicious circle. Some important arguments are:

(i) Transformational grammars can always be translated in the ATN-formalism, since ATNs have Turing machine power. By doing so, properties of TG-generated structures which are only there for the sake of *TG-formalism* can disappear in an ATN representation. There is no further need to generate them by the parsing process.

(ii) The ATN is a *linguistic* theory; but as I have written elsewhere (170), 'It is no longer the psychological theory which is adapted to the grammar, but rather the grammar which is written for the representation of psychological processing operations. If such a network at the same time provides all input sentences with

their correct grammatical parsing, this new isomorphism is of a more acceptable kind than the naive isomorphism discussed (above)'. Therefore, the ATN does behave properly according to linguistic constraints.

(iii) The output of an ATN need not be a deep structure (even stripped from its incidental formal characteristics). Underlying grammatical relations will be operative at various stages of the parsing process, without necessarily being represented in the end-product of the parsing process.

(iv) ATN networks allow for semantic and pragmatic conditions on transitions and for building operations which create semantic and pragmatic output (see especially Miller and Johnson-Laird, 194).

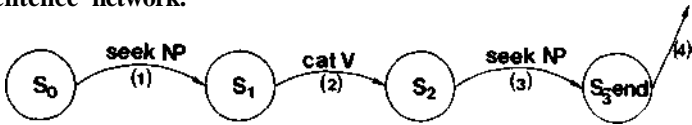
Though syntactic and non-syntactic aspects of ATNs are clearly separable, ATNs give a natural means for expressing distributed syntactic—semantic control. Specific operating characteristics of ATNs which make them especially 'human' (such as left-to-right processing) have been enumerated by Wanner and Maratsos (260).

Rather than enumerate these advantages, we shall present an example of an ATN network, taken from Wanner *et al.* (259), and designed as a demonstration example.

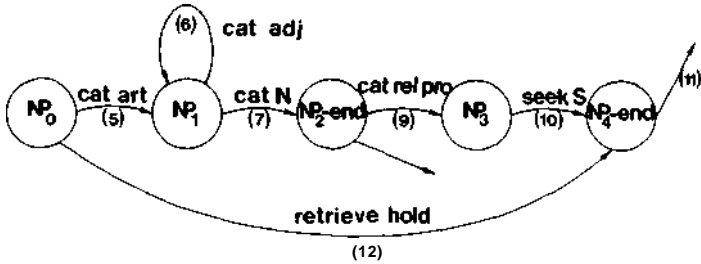
In the ATN of Figure 1.1 circles denote states, and arrows transitions. The initial state is S_0 (S for 'sentence') and the upper network is the sentence network. A set of conditions and a set of actions is assigned to each transition. Conditions are given over the arrows in Figure 1.1, actions are listed separately. If the ATN is in a given state, a transition from there to a next state can be made if the particular condition is fulfilled. For instance, the transition from S_0 to S_1 requires that an NP has been detected (i.e. one has to 'seek' an NP in order to make a transition). To go from S_1 to S_2 requires the input of an item (word) of category V, i.e. a verb, and so on. In the case of SEEK conditions, control has to be shifted to the appropriate network. In order to satisfy the SEEK NP condition from S_0 to S_1 , the state S_0 is pushed down and control shifts to the NP network, specifically to NP_0 . If this network is successfully passed, i.e. reaches the NP_4 -end state, control is popped to S_0 and the transition to S_1 can be made. An ATN may consist of several networks, between which control can shift back and forth. If there is more than one possible transition from a given state, as in NP_1 it is customary to arrange them clockwise in such a way that the first arc after 12 is tried first, and so forth. The final state is reached when (a) the ATN is in the end state of the sentence network, and (b) there are no states left in the push-down store.

Actions can be of several kinds. If one wants to use an ATN as a syntactic parser which gives phrase structures as output, actions have to assign category symbols (such as NP) and/or grammatical functions (such as subject), and to assemble types of structures. In other words, a phrase (e.g. an NP), when built and categorized, is stored in a functional register (a subject-register, an object-

sentence network.



noun phrase network:



Arc	Action
1	ASSIGN SUBJECT to current phrase
2	ASSIGN ACTION to current word
3	ASSIGN OBJECT to current phrase
4	ASSEMBLE CLAUSE SEND current clause
5	ASSIGN DET to current word
6	ASSIGN MOD to current word
7	ASSIGN HEAD to current word
8	ASSEMBLE NOUN PHRASE SEND current phrase
9	HOLD
10	CHECK HOLD ASSIGN MOD to current clause
11	ASSEMBLE NOUN PHRASE SEND current phrase
12	(no action)

Figure 1.1. A simple ATN grammar for sentences with relative clauses (from Wanner, Kaplan, and Shiner)

register, etc.). But actions can equally well be of certain semantic types. Miller and Johnson-Laird (194) call such actions *executions*. These may be very task-specific activities which the listener chooses to perform. One example is that at a certain transition the referent of the particular word or phrase, specified in the condition, is looked up in long-term memory. (Such examples are not given in Figure 1.1.)

Wanner *et al.* (259) demonstrate the working of the example ATN by parsing the sentence 'The old man that the boy loved caught the fish'. Roughly, the parsing proceeds as follows. From S_0 , control shifts to NP_0 in order to find an NP. 'The old man' leads via transitions 5, 6, and 7 to state NP_2 -end. In order to make transition 9, one needs a relative pronoun, which is in fact present ('that'). The action taken here is important. The categorized information (i.e. DET, MOD, HEAD; see actions 5, 6, 7) are put in a special register, the so-called HOLD register; that is, no grammatical function (subject, object) is assigned. This is postponed to a later stage. There is a good reason for doing this, since the grammatical function of the first NP in a relative-clause sentence depends on later information. In order to make the last transition within the NP network, control has to be transferred to the sentence network (i.e. at the moment S_0 and NP_3 are on the push-down store). For the first transition control has to go again to the NP network, which parses 'the boy' as an NP. From S_1 the verb 'loved' leads to S_2 , after which the NP network takes over again. Since transition 5 cannot be made, transition 12 can be made by retrieving the information from the HOLD register (The_{DET} old_{MOD} man_{HEAD}). Transition 11 leads to assembling this as an NP; and at popping back to S_2 , this NP is assigned to OBJECT-function. In this way, the ATN takes care that 'the old man' becomes the object of 'loved'. The further steps are obvious: After assembling the relative clause (4), control returns to S_0 , transition 1 can be made, and the verb 'caught', followed by the NP 'the fish', leads to S_3 -end, plus an empty push-down store.

Returning now to the ATN-literature, it should be remarked that the notion stems from Thorne *et al.*'s (245) parsing system for English, and was further developed by Bobrow and Fraser (29) and Woods (270). It found its first psychological application in Kaplan (152), who showed how some of Bever's 'perceptual strategies' could be cast in the ATN formalism. Further psychological applications were made by Kaplan (153), Stevens and Rumelhart (239) on a reading task, by Wanner *et al.* on relative clause comprehension (259, 260), and by Wanner and Shiner on structural ambiguities (262). Wanner and Maratsos (260) give evidence for the memory-loading effect of using the HOLD register (see above). The elegant theoretical analysis deserves further empirical verification by means of acoustical simultaneous measurement techniques.

But the theoretical concepts also need further psychological interpretation. It is normal for an ATN to have various stores and registers: a push-down store for keeping track of control shifts, registers for assigning functions to built subtrees, and a HOLD register for keeping uninterpreted structures. How do these stores relate to working memory, short- and long-term storage, levels of processing, and similar concepts described in the memory literature? In this connection, Anderson's (6) critique of the ATN approach should be especially mentioned. Anderson clearly recognizes the various attractions of ATNs but doubts whether they will ever become psychological models. Because of their Turing machine power, they have no intrinsic limitation. Human cognitive limitations, therefore, should be explicitly imposed on the models. The question is whether

there are natural ways to do so, or whether one is forced into an arbitrary set of limiting assumptions. Anderson shows, for instance, that HOLD cannot be the only reason for the increased processing load of centre-embedded sentences. The recursive shifts of control (SEEK operations) should be a cause as well, but then right-branching structures should also be difficult, since they involve the same SEEK operations. Moreover, there is no natural limitation on the size of a network, nor therefore on the size of the structure to be assembled after reaching the end-state. Would a large assembled structure use up so much memory capacity that it would interfere with the assembling of further structures? In addition, ATNs are of little help for syntactically scrambled or incomplete sentences which may still be interpretable. And in actuality, human listeners often skilfully interpret such utterances. Clearly, therefore, ATNs cannot be the answer to all problems; and in fact Anderson replaces it by his so-called ACT model, which retains some of the ATNs advantages but is less powerful and claimed to be psychologically more realistic.

What is important is to recognize that we have a class of models here which is at the 'right' level of theorizing. They are more local and experimentally more testable than are AI-models of language understanding (see the following section), but at the same time they are more structured than mere lists of perceptual strategies.

If we now reconsider the 'ILA hypothesis' (section 4.1.1), we see that ATN models may be able to give a refined version of it. Miller and Johnson-Laird (194) use ATNs to 'translate' the sentence into routines, as in a compiler. These (semantic) routines may or may not be executed, depending on the task and the motivation of the hearer; but the compiling process itself is automatic, and in general outside voluntary control. 'Immediate awareness' here means compiling, it is not task-specific, and it is the first phase of understanding. However, it is the first phase only in a logical sense; this does not mean that one has to wait until the end of the clause or sentence in order to execute the routines. Routines appear at successive stages of the ATN parsing process, and execution can be done as soon as the routine appears. Thus, syntactic and interpretative activity may go on in parallel, in correspondence with Marslen-Wilson's empirical findings (184, 186, 188), whereas at the same time a distinction is maintained between automatic-general and voluntary task-specific operations in language understanding.

One thing, however, is not taken care of so long as one maintains the logical priority of 'translation'. There is some evidence that the results of executing routines can feed back to the translation process. Marslen-Wilson and Tyler's (187) finding that word-, rhyme-, and category-monitoring are slower for syntactic (i.e. meaningless) prose than for normal prose points in this direction. From these and other results Marslen-Wilson concludes that syntactic and semantic analysis proceed in parallel and interactively. In order to describe such an inverse (top-down) flow of information more formally, one has to consider what is called *goal-oriented parsing* in artificial intelligence.

4.3.3 *The artificial intelligence approach*

Riesbeck and Schank's contribution to the present volume contains a thorough statement of the principles and structure of goal-oriented, natural language parsing. This makes it somewhat superfluous to review these issues here. We will limit ourselves to a few general remarks. The parallel-interactive features of sentence comprehension are a natural consequence of AI-parsing principles, at least those of Schank (224, 225, 226) and Wilks (265): 'The common linguistic model with a syntactic analysis phase followed by a semantic interpretation is a good example of what we are against. The modules in the understanding process affect each other in both directions' (Riesbeck and Schank, this volume). If the ultimate goal of parsing is to extract meaning, i.e. to derive a conceptual representation, semantic considerations should have priority over syntactic ones. The latter become important if the interpretation is highly unexpected. 'I think that syntax exists so that we can say improbable things' (Garrett, 103). Parsing in these systems goes from left to right, without much backing up. It is essentially expectation-based, where expectation is due to knowledge of the world as well as to prior (con)text. It is encouraging that such systems can be built.

One should keep in mind, however, that even the most promising systems still have serious limitations. Like ATNs, they all have Turing machine power, and the question which has to be answered is what sort of restrictive principles have to be imposed in order to make these programs human-like. Also, in spite of their power, all present systems show time characteristics which make them unrealistically slow if their knowledge of the world, or data base, is expanded to a more realistic size. Efforts to partition the data base into 'frames' (195) or 'scripts' (2) are very much in a beginning phase, whereas the whole notion is badly undefined. Levelt (171) discusses how existing systems are unable to account for language acquisition. Anderson (6) makes a similar point and enumerates several other weaknesses of computer language systems. Nevertheless, there can be no doubt that AI (and AI-related research) has become increasingly influential in the study of language understanding. Some major sources in this literature are by Winograd (268, 269), Anderson and Bower (7), Anderson (6), Norman and Rumelhart (203), Schank (226), and Miller and Johnson-Laird (194). The recent convergence of insights from AI and experimental psycholinguistics gives one the hope that essential AI principles (as opposed to accidental properties of parsing programs—and authors are often not very clear about this distinction) will lend themselves to empirical tests in the near future.

4.4. Conclusions

Section 1 cited three broad types of aim in the study of sentence perception: linguistic, conceptual, and communicative. During the first few years of the

reviewed period, the linguistic aim was the dominant one. Two causes may have contributed to this state of affairs. Firstly, there was still an echo sounding from the linguistic revolution of the nineteen-sixties. Secondly, the range of research methods developed during the nineteen-sixties belonged mostly to the class of successive measurement. These methods are particularly inapt for disproving the stage-like character of the models which had been developed within the linguistic framework. More specifically, it is hard to reject the notion that syntactic analysis of the clause precedes semantic analysis, that semantic analysis precedes the search for referents in long-term memory, and so on. Therefore, one could always maintain that in comprehending a sentence a listener first takes a general-purpose linguistic (and probably syntactic) step, and only then resorts to specific inferential activities which are tailored to the conceptual and communicative requirements of the particular task. This, moreover, would make the different types of aim compatible, and in fact complementary.

Two other causes, however, have badly undermined this picture during the second half of the reviewed period. Firstly, the development of simultaneous-measurement techniques has made it possible to demonstrate the highly interactive nature of sentence-understanding processes: semantic decisions which can take place quite early in the sentence can affect syntactic and phonetic decisions, and it is not quite clear any more what occurs prior to what. Secondly, developments in artificial intelligence and related fields became increasingly influential in psycholinguistic theory construction. Goal-oriented parsing is most effectively realized in a heterarchical system where control is distributed.

It would be wrong, however, to replace one pet theory by another. Not undermined, for instance, is the notion that syntax plays an important role in sentence understanding. It will take much and thorough experimentation to determine how syntactic operations interact with semantic and conceptual decisions as the listener infers the intentions of the speaker. This type of research might profit from theories which are more local than full-size computer models of language understanding but more inclusive and more structured than a mere list of understanding strategies. Augmented transition networks form one example of such a level of theorizing, but other types of models should not be excluded.

ACKNOWLEDGEMENTS

The author is very grateful to Trudy Las, who assisted in searching the library, and in categorizing and summarizing materials. Gerard Kempen and William Marslen-Wilson were helpful in giving comments on earlier drafts of this paper. The author's gratitude also goes to the Institute for Perception Research, Eindhoven, where he wrote, as a guest, a major part of this chapter.

5. REFERENCES

1. Aaronson, D. Stimulus factors and listening strategies in auditory memory: An experimental demonstration. *Cognitive Psychology*, 1974, 6, 133-158.
2. Abelson, R. P. The structure of belief systems. In R. C. Schank and K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973.
3. Abrams, K. and Bever, T. G. Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology*, 1969, 21, 280-290.
4. Amnon, P. R., Ostrowski, B., and Alward, K. Effects of task on the perceptual organization of sentences. *Perception and Psychophysics*, 1971, 10, 361-363.
5. Anderson, J. R., Verbatim and propositional representation of sentences in immediate and long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 149-162.
6. Anderson, J. R. *Language, memory and thought*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976.
7. Anderson, J. R. and Bower, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.
8. Anisfeld, M. and Klenbort, I. On the functions of structural paraphrase: the view from the passive voice. *Psychological Bulletin*, 1973, 79, 117-126.
9. Baddeley, A. D. *The psychology of memory*. New York: Harper & Row, 1976.
10. Baddeley, A. D. and Hitch, G. Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press, 1974.
11. Baird, R. and Koslick, J. D. Recall of grammatical relations within clause-containing sentences. *Journal of Psycholinguistic Research*, 1974, 3, 165-171.
12. Baker, C. L. Double negatives. *Linguistic Inquiry*, 1970, 1, 169-186.
13. Banks, W. P., Clark, H. H., and Lucy, P. The locus of the semantic congruity effect in comparative judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, 104, 35-47.
14. Barclay, J. R. The role of context in remembering sentences. *Cognitive Psychology*, 1973, 4, 229-254.
15. Bertelson, P. Listening from left to right versus right to left. *Perception*, 1972, 1, 161-165.
16. Bertelson, P. and Tisseyre, F. Perceiving the sequence of speech and non-speech stimuli. *Quarterly Journal of Experimental Psychology*, 1970, 22, 653-662.
17. Bertelson, P. and Tisseyre, F. Lateral asymmetry in the perceived sequence of speech and non-speech stimuli. *Perception and Psychophysics*, 1972, 11, 356-362.
18. Bertelson, P. and Tisseyre, F. Lateral asymmetry in judgments of click location: Not an artifact of reporting mode. *Perceptual and Motor Skills*, 1973, 36, 849-850.
19. Bever, T. G. The comprehension and memory of sentences with temporal relations. In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
20. Bever, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley, 1970.
21. Bever, T. G. Serial position and response biases do not account for the effect of syntactic structure on the location of brief noises during sentences. *Journal of Psycholinguistic Research*, 1973, 2, 287-288.
22. Bever, T. G., Garrett, M. F., and Hurtig, R. The interaction of perceptual processes and ambiguous sentences. *Memory and Cognition*, 1973, 1, 277-286.
23. Bever, T. G., Hurtig, R. R., and Handel, A. B. Response biases do not account for the effect of clause structure on the perception of non-linguistic stimuli. *Research Bulletin Educational Testing Service*, RB 75-30, 1975.

24. Bever, T. G. and Hurtig, R. R. Detection of a non linguistic stimulus is poorest at the end of a clause. *Journal of Psycholinguistic Research*, 1975, 4, 1-7.
25. Bever, T. G., Lackner, J. R., and Kirk, R. The underlying structures of sentences are the primary units of immediate speech processing. *Perception and Psychophysics*, 1969, 5, 225-234.
26. Bever, T. G., Lackner, J. R., and Stolz, W. Transitional probability is not a general mechanism for the segmentation of speech. *Journal of Experimental Psychology*, 1969, 79, 387-394.
27. Blaubergs, M. S. and Braine, M. D. S. Short-term memory limitations on decoding self-embedded sentences. *Journal of Experimental Psychology*, 1974, 102, 745-748.
28. Bobrow, S. A. and Bell, S. M. On catching on to idiomatic expressions. *Memory and Cognition*, 1973, 1, 343-346.
29. Bobrow, D. and Fraser, B. An augmented state transition network analysis procedure. In D. Walker and L. Norton (Eds.), *Proceedings of the International Joint Conference on Artificial Intelligence*. Washington D.C., 1969.
30. Bransford, J. D., Barclay, J. R., and Franks, J. J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 1972, 3, 193-209.
31. Bransford, J. D. and Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, 2, 331-350.
32. Bransford, J. D. and Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 717-726.
33. Bransford, J. D. and Johnson, M. K. Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
34. Bransford, J. D. and McCarrell, N. S. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weimer and D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
35. Cairns, H. S. Effects of bias on processing and reprocessing of lexically ambiguous sentences. *Journal of Experimental Psychology*, 1973, 97, 337-343.
36. Cairns, H. S. and Foss, J. Falsification of the hypothesis that word frequency is a unified variable in sentence processing. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 41-43.
37. Cairns, H. S. and Kamerman, J. Lexical information processing during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 170-179.
38. Caplan, D. Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics*, 1972, 12, 73-76.
39. Carey, P. W., Mehler, J., and Bever, T. G. When do we compute all the interpretations of an ambiguous sentence? In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
40. Carey, P. W., Mehler, J., and Bever, T. G. Judging the veracity of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 243-254.
41. Carey, S. T. and Lockhart, R. S. Encoding differences in recognition and recall. *Memory and Cognition*, 1973, 1, 297-300.
42. Carpenter, P. A. Extracting information from counterfactual clauses. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 512-521.
43. Carpenter, P. A. On the comprehension, storage, and retrieval of comparative sentences. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 401-411.

44. Carpenter, P. A. and Just, M. A. Semantic control of eye movements in picture scanning during sentence-picture verification. *Perception and Psychophysics*, 1972, 12, 61-64.
45. Carpenter, P. A. and Just, M. A. Sentence comprehension: A psycholinguistic model of sentence verification. *Psychological Review*, 1975, 82, 45-73.
46. Carpenter, P. A. and Just M. A. Models of sentence verification and linguistic comprehension. *Psychological Review*, 1976, 83, 318-322.
47. Carroll, J. M. and Bever, T. G. Sentence comprehension: case study in the relation of knowledge and perception. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception. Vol. 7: Language and speech*. New York: Academic Press, 1976.
48. Chapin, P. G., Smith, T. S. and Abrahamson, A. A. Two factors in perceptual segmentation of speech. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 164-173.
49. Chase, W. G. and Clark, H. H. Semantics in the perception of verticality. *British Journal of Psychology*, 1971, 62, 311-326.
50. Chase, W. G. and Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
51. Clark, H. H. Linguistic processes in deductive reasoning. *Psychological Review*, 1969, 76, 387-404.
52. Clark, H. H. Difficulties people have in answering the question 'Where is it?' *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 265-277.
53. Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359.
54. Clark, H. H. Semantics and comprehension. In T. A. Sebeok (Ed.), *Current trends in linguistics. Vol. 12: Linguistics and adjacent arts and sciences*. The Hague: Mouton, 1974.
55. Clark, H. H. and Begun, J. S. The semantics of sentence subjects. *Language and Speech*, 1971, 14, 34-46.
56. Clark, H. H., Carpenter, A., and Just, M. A. On the meeting of semantics and perception. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
57. Clark, H. H. and Chase, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, 3, 472-517.
58. Clark, H. H. and Chase, W. G. Perceptual coding strategies in the formation and verification of descriptions. *Memory and Cognition*, 1974, 2, 101-111.
59. Clark, H. H. and Haviland, S. Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse comprehension and production*. Norwood, N.J.: Ablex Publishing, 1977.
60. Clark, H. H. and Lucy, P. Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 56-72.
61. Cohen, A. and 't Hart, J. On the anatomy of intonation. *Lingua*, 1967, 19, 177-192.
62. Collier, R. and 't Hart, J. The role of intonation in speech perception. In A. Cohen and S. G. Nooteboom (Eds.), *Structure and process in speech perception*. Berlin, Heidelberg, New York: Springer Verlag, 1975.
63. Collins, A. M. and Quillian, M. R. Experiments on semantic memory and language comprehension. In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
64. Conrad, C. Context effects in sentence comprehension: A study of the subjective lexicon. *Memory and Cognition*, 1974, 2, 130-138.

65. Cornish, E. R. Pragmatic aspects of negation in sentence evaluation and completion tasks. *British Journal of Psychology*, 1971, 62, 505-511.
66. Craik, F. I. and Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
67. Cutler, A. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics*, 1976, 20, 55-60.
68. Cutler, A. Beyond parsing and lexical look-up: an enriched description of auditory sentence comprehension. In R. J. Wales and E. Walker (Eds.), *New approaches to language mechanism*. Amsterdam: North-Holland, 1976.
69. Darwin, C. J. On the dynamic use of prosody in speech perception. In A. Cohen and S. G. Neeboom (Eds.), *Structure and process in speech perception*. Berlin, Heidelberg, New York: Springer Verlag, 1975.
70. Davies, D. and Isard, S. D. Utterances as programs. In D. Michie (Ed.), *Machine Intelligence 7*. Edinburgh: Edinburgh University Press, 1972.
71. Donders, F. C. *Over de snelheid van psychische processen*. Onderzoekingen gedaan in het Psychologisch Laboratorium der Utrechtse Hoogeschool, 1868-1869, Tweede reeks, II, 92-120.
72. Dooling, D. J. Rhythm and syntax in sentence perception. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 255-264.
73. Dooling, D. J. and Christiaansen, R. E. Context effects in sentence comprehension: A reply to Doll and Lapinski. *Bulletin of the Psychonomic Society*, 1975, 5, 261-262.
74. Fillenbaum, S. Inducements: on the phrasing and logic of conditional promises, threats, and warnings. *Psychological Research*, 1976, 38, 231-250.
75. Flores d'Arcais, G. B. Semantic and perceptual factors in the processing of comparative sentences. *Italian Journal of Psychology*, 1974, 1, 267-303.
76. Flores d'Arcais, G. B. Some perceptual determinants of sentence construction. In G. B. Flores d'Arcais, (Ed.), *Studies in perception*. Milano: Martello, 1975.
77. Flores d'Arcais, G. B. The perception of complex sentences. This volume.
78. Fodor, J. A. *The language of thought*. Hassocks: Harvester Press, 1976.
79. Fodor, J. A. and Bever, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 414-420.
80. Fodor, J. A., Bever, T. G., and Garrett, M. F. *The psychology of language. An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill, 1974.
81. Fodor, J. D., Fodor, J. A., and Garrett, M. F. The psychological unreality of semantic representations. *Linguistic Inquiry*, 1975, 6, 515-531.
82. Fodor, J. A. and Garrett, M. F. Some syntactic determinants of sentential complexity. *Perception and Psychophysics*, 1967, 2, 289-296.
83. Fodor, J. A., Garrett, M. F., and Bever, T. G. Some syntactic determinants of sentential complexity II: Verb structure. *Perception and Psychophysics*, 1968, 3, 453-461.
84. Forster, K. I. Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics*, 1970, 8, 215-221.
85. Forster, K. I. and Olbrei, I. Semantic heuristics and syntactic analysis. *Cognition*, 1973, 2, 319-347.
86. Forster, K. I. and Ryder, L. A. Perceiving the structure and meaning of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 285-296.
87. Foss, D. J. Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 457-462.
88. Foss, D. J. Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 699-706.

89. Foss, D. J., Bever, T. G., and Silver, M. The comprehension and verification of ambiguous sentences. *Perception and Psychophysics*, 1968, 4, 304-306.
90. Foss, D. J. and Cairns, H. S. Some effects of memory limitation upon sentence comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 541-547.
91. Foss, D. J. and Jenkins, C. M. Some effects of context on the comprehension of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 577-589.
92. Foss, D. J. and Lynch, R. H. Decision processes during sentence comprehension: effects of surface structure on decision times. *Perception and Psychophysics*, 1969, 5, 145-148.
93. Foss, D. J. and Swinney, D. A. On the psychological reality of the phoneme: Perception, identification and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 246-257.
94. Franks, J. J. Toward understanding understanding. In W. B. Weimer and D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
95. Frederikson, C. H. Effects of task-induced cognitive operations on comprehension and memory processes. In R. O. Freedle and J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. New York: Winston, 1972.
96. Frederikson, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, 7, 139-166.
97. Freedle, R. O. and Craun, M. Observations with self-embedded sentences using written aids. *Perception and Psychophysics*, 1970, 7, 247-249.
98. Friend, K. E. Perceptual encoding in comparative judgments of race. *Memory and Cognition*, 1973, 1, 80-84.
99. Gamlin, P. J. Level of listening comprehension as a function of two process variables: Syntax and meaningfulness. *Language and Speech*, 1972, 15, 232-261.
100. Garrett, M. F. Syntactic structures and judgments of auditory events. Unpublished Ph.D. Dissertation, University of Illinois, 1965.
101. Garrett, M. F. Does ambiguity complicate the perception of sentences? In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
102. Garrett, M. F. Experimental issues in sentence comprehension: Complexity and segmentation. In C. Cherry (Ed.), *Pragmatic aspects of human communication*. Dordrecht: Reidel, 1974.
103. Garrett, M. F. Syntactic processes in sentence production. In R. J. Wales and E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland, 1976.
104. Garrett, M. F., Bever, T. G., and Fodor, J. A. The active use of grammar in speech perception. *Perception and Psychophysics*, 1966, 1, 30-32.
105. Garrod, S. and Trabasso, T. A dual-memory information processing interpretation of sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 155-167.
106. Glass, A. L. and Holyoak, K. J. The effect of *some* and *all* on reaction time for semantic decisions. *Memory and Cognition*, 1974, 2, 436-440.
107. Glass, A. L., Holyoak, K. J., and O'Dell, C. Production frequency and the verification of quantified statements. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 237-254.
108. Glucksberg, S., Trabasso, T., and Wald, J. Linguistic structures and mental operations. *Cognitive Psychology*, 1973, 5, 338-370.

109. Gough, P. B. Experimental psycholinguistics. In W. O. Dingwall (Ed.), *A survey of linguistic science*. College Park, Md.: Linguistic Program University of Maryland, 1971.
110. Green, D. W. The effects of task on the representation of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 275-283.
111. Greene, J. M. Syntactic form and semantic function. *Quarterly Journal of Experimental Psychology*, 1970, 22, 14-27.
112. Greene, J. M. The semantic function of negatives and positives. *British Journal of Psychology*, 1970, 61, 17-22.
113. Greene, J. M. and Wason, P. C. Negation: A rejoinder to Wales and Grieve. *Perception and Psychophysics*, 1970, 8, 238-239.
114. Grice, H. P. Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics*. Vol. 3: *Speech acts*. New York: Seminar Press, 1975.
115. Grieve, R. and Wales, R.J. Passives and topicalization. *British Journal of Psychology*, 1973, 64, 173-182.
116. Haggard, M. Understanding speech understanding. In A. Cohen and S. G. Nooteboom (Eds.), *Structure and process in speech perception*. Berlin, Heidelberg, New York: Springer Verlag, 1975.
117. Hakes, D. T. Does verb structure affect sentence comprehension? *Perception and Psychophysics*, 1971, 10, 229-232.
118. Hakes, D. T. Effects of reducing complement constructions on sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 278-286.
119. Hakes, D. T. and Cairns, H. S. Sentence comprehension and relative pronouns. *Perception and Psychophysics*, 1970, 8, 5-8.
120. Hakes, D. T. and Foss, D. J. Decisions processes during sentence comprehension: Effects of surface structure reconsidered. *Perception and Psychophysics*, 1970, 8, 413-416.
121. Hakes, D. T., Evans, J. S., and Brannon, L. L. Understanding sentences with relative clauses. *Memory and Cognition*, 1976, 4, 283-290.
122. Halle, M. and Stevens, K. N. Speech recognition: A model and a program for research. In J. A. Fodor and J. J. Katz (Eds.), *Readings in the philosophy of language*. Englewood-Cliffs: Prentice Hall, 1964.
123. Hamilton, H. W. and Deese, J. Comprehensibility and subject-verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 163-170.
124. t'Hart, J. and Cohen, A. Intonation by rule: A perceptual quest. *Journal of Phonetics*, 1973, 1, 309-327.
125. t'Hart, J. and Collier, R. Integrating different levels of intonation analysis. *Journal of Phonetics*, 1975, 3, 235-255.
126. Haviland, S. E. and Clark, H. H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 512-521.
127. Healy, A. F. and Miller, G. A. The relative contribution of nouns and verbs to sentence acceptability and comprehensibility. *Psychonomic Science*, 1971, 4, 94-96.
128. Hogaboam, T. W. and Perfetti, C. A. Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 265-274.
129. Holmes, V. M. Order of main and subordinate clauses in sentence perception. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 285-293.
130. Holmes, V. M. and Forster, K. I. Detection of extraneous signals during sentence recognition. *Perception and Psychophysics*, 1970, 7, 297-301.

131. Holmes, V. M. and Forster, K. I. Click localization and syntactic structure. *Perception and Psychophysics*, 1972, 12, 9-15.
132. Holmes, V. M. and Forster, K. I. Perceptual complexity and underlying sentence structure. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 148-156.
133. Holyoak, K. J. The role of imagery in the evaluation of sentences: imagery or semantic factors? *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 163-166.
134. Holyoak, K. J. and Glass, A. L. The role of contradictions and counterexamples in the rejection of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 215-239.
135. Hörmann, H. The concept of sense constancy. Paper for the XXIth International Congress of Psychology, Paris, July 1976.
136. Hörmann, H. *Meinen und Verstehen. Grundzüge einer psychologischen Semantik*. Frankfurt: Suhrkamp, 1976.
137. Hornby, P. A. Surface structure and presupposition. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 530-538.
138. Hupet, M. and Le Bouedec, B. Definiteness and voice in the interpretation of active and passive sentences. *Quarterly Journal of Experimental Psychology*, 1975, 27, 323-330.
139. Huttenlocher, J., Higgins, E. T., Milligan, C. and Kauffman, B. The mystery of the 'negative equative' construction. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 334-341.
140. Jarvella, R. J. Effects of syntax on running memory span for connected discourse. *Psychonomic Science*, 1970, 19, 235-236.
141. Jarvella, R. J. Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 409-416.
142. Jarvella, R. J. and Herman, S. J. Clause structure of sentences and speech processing. *Perception and Psychophysics*, 1972, 11, 381-384.
143. Johnson-Laird, P. N. Experimental psycholinguistics. *Annual Review of Psychology*, 1974, 25, 135-160.
144. Johnson-Laird, P. N. The three term-series problem. *Cognition*, 1972, 1, 57-82.
145. Johnson-Laird, P. N. and Tridgell, J. N. When negation is easier than affirmation. *Quarterly Journal of Experimental Psychology*, 1972, 24, 87-91.
146. Jones, S. Visual and verbal processes in problem-solving. *Cognitive Psychology*, 1970, 1, 201-214.
147. Jorgensen, C. C. and Kintsch, W. The role of imagery in the evaluation of sentences. *Cognitive Psychology*, 1973, 4, 110-116.
148. Just, M. A. Comprehending quantified sentences: The relation between sentence-picture and semantic memory verification. *Cognitive Psychology*, 1974, 6, 216-236.
149. Just, M. A. and Carpenter, P. A. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 244-253.
150. Just, M. A. and Carpenter, P. A. The relation between comprehending and remembering some complex sentences. *Memory and Cognition*, 1976, 4, 318-322.
151. Just, M. A. and Clark, H. H. Drawing inferences from the presuppositions and implications of affirmative and negative sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 21-31.
152. Kaplan, R. M. Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence*, 1972, 3, 77-100.
153. Kaplan, R. M. On process models for sentence analysis. In D. A. Norman and D. E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: Freeman, 1975.
154. Katwijk, A. van. *Accentuation in Dutch*. Eindhoven: I.P.O., 1974.

155. Kempen, G. Syntactic constructions as retrieval plans. *British Journal of Psychology*, 1976, 67, 149-160.
156. Kimball, J. P. Seven principles of surface structure parsing in natural language. *Cognition*, 1973, 2, 15-47.
157. Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
158. Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., and Keenan, J. M. Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 196-214.
159. Klee, H. and Eysenck, M. W. Comprehension of abstract and concrete sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 522-529.
160. Klenbort, I. and Anisfeld, M. Markedness and perspective in the interpretation of the active and passive voice. *Quarterly Journal of Experimental Psychology*, 1974, 26, 189-195.
161. Komfeld, J. R. Clause boundary and dominance effects on sentence perception. *M.I.T. Research Laboratory of Electronics, Quarterly Progress Report*, 1973, **110**, 177-181.
162. Krueger, L. E. Sentence-picture comparison: A test of additivity and processing time for feature matching and negation coding. *Journal of Experimental Psychology*, 1972, 95, 275-284.
163. Krulee, G. K. and Schwartz, H. R. Scanning processes and sentence recognition. *Journal of Psycholinguistic Research*, 1975, 4, 141-158.
164. Lackner, J. R. and Garrett, M. F. Resolving ambiguity: Effects of biasing contexts in the unattended ear. *Cognition*, 1972, 1, 359-372.
165. Ladefoged, P. *Three areas of experimental phonetics*. London: Oxford University Press, 1967.
166. Ladefoged, P. and Broadbent, D. E. Perception of sequence in auditory events. *Quarterly Journal of Experimental Psychology*, 1960, 13, 162-170.
167. Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*, New York: Wiley, 1951.
168. Levelt, W. J. M. Generative grammatica en psycholinguïstiek. II Psycholinguïstisch onderzoek. *Nederlands Tijdschrift voor de Psychologie*, 1966, 21, 367-400.
169. Levelt, W. J. M. Hierarchical chunking in sentence processing. *Perception and Psychophysics*, 1970, 8, 99-103.
170. Levelt, W. J. M. *Formal grammars in linguistics and psycholinguistics*. Vol. I: *An introduction to the theory of formal languages and automata*. Vol. II: *Applications in linguistic theory*. Vol. III: *Psycholinguistic applications*. The Hague: Mouton, 1974.
171. Levelt, W. J. M. *What became of LAD?* Lisse: Peter de Ridder Press, 1975.
172. Levelt, W. J. M. and Kempen, G. Semantic and syntactic aspects of remembering sentences: A review of some recent continental research. In A. Kennedy and A. Wilkes (Eds.), *Studies in long term memory*. New York: Wiley, 1975.
173. Levelt, W. J. M. and Ouweneel, G. R. E. The perception of French sentences with a surface structure ambiguity. *Nederlands Tijdschrift voor de Psychologie*, 1969, 24, 245-248.
174. Levelt, W. J. M., Schreuder, R., and Hoenkamp, E. Structure and use of verbs of motion. In R. N. Campbell and P. T. Smith (Eds.), *Recent advances in the psychology of language: Formal and experimental approaches*. New York: Plenum Press 1978
175. Levelt, W. J. M., Zwanenburg, W., and Ouweneel, G. R. Ambiguous surface structure and phonetic form in French. *Foundations of Language*, 1970, 6, 260-273.
176. Lewis, J. L. Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology*, 1970, 85, 225-228.

177. Lippman, M. Z. The influence of grammatical transform in a syllogistic reasoning task. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 424-430.
178. Locatelli, F. Le rôle des oppositions temporelles dans la compréhension de la succession. *L'Année Psychologique*, 1973, 73, 493-506.
179. Loosen, F. Cognitive organisatie van zinnen in het geheugen. Unpublished Ph. Dissertation, Louvain, 1972.
180. MacKay, D. G. To end ambiguous sentences. *Perception and Psychophysics*, 1966, 3, 426-436.
181. MacKay, D. G. Mental diplopia: Towards a model of speech perception at the semantic level. In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
182. MacKay, D. G. Aspects of the theory of comprehension, memory and attention. *Quarterly Journal of Experimental Psychology*, 1973, 25, 22-40.
183. MacKay, D. G. and Bever, T. G. In search of ambiguity. *Perception and Psychophysics*, 1967, 2, 193-200.
184. Marslen-Wilson, W. D. Linguistic structure and speech shadowing at very short latencies. *Nature*, 1973, 244, 522-523.
185. Marslen-Wilson, W. D. Sentence perception as an interactive parallel process. *Science*, 1975, 189, 226-228.
186. Marslen-Wilson, W. D. Linguistic descriptions and psychological assumptions in the study of sentence perception. In R. J. Wales and E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland, 1976.
187. Marslen-Wilson, W. D. and Tyler, L. K. Processing structure of sentence perception. *Nature*, 1975, 257, 784-786.
188. Marslen-Wilson, W. D. and Tyler, L. K. Memory and levels of processing in a psycholinguistic context. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 112-119.
189. Martin, J. G. Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 1972, 79, 487-509.
190. Martin, J. G. Rhythmic expectancy in continuous speech perception. In A. Cohen and S. G. Nooteboom (Eds.), *Structure and process in speech perception*. Berlin, Heidelberg, New York: Springer Verlag, 1975.
191. McNeill, D. and Lindig, K. The perceptual reality of phonemes, syllables, words and sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 419-430.
192. Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, 1, 242-300.
193. Miller, G. A. Toward a third metaphor for psycholinguistics. In W. B. Weimer and D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
194. Miller, G. A. and Johnson-Laird, P. N. *Language and perception*. Cambridge, Mass.: Harvard University Press, 1976.
195. Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision*. New York: McGraw Hill, 1975.
196. Miron, M. S. and Brown, E. The comprehension of rate-incremented aural coding. *Journal of Psycholinguistic Research*, 1971, 1, 65-76.
197. Mistler-Lachman, J. L. Levels of comprehension in processing of normal and ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 614-623.
198. Mistler-Lachman, J. L. Depth of comprehension and sentence memory. *Journal of Verbal Learning and Verbal Behaviour*, 1974, 13, 98-106.
199. Mistler-Lachman, J. L. Queer sentences, ambiguity, and levels of processing. *Memory and Cognition*, 1975, 3, 395-400.

200. Morton, J. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
201. Morton, J. and Long, J. Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 43-51.
202. Noordman, L. G. M. Foreground and background information in reasoning. In R. N. Campbell and P. T. Smith (Eds.), *Recent advances in the psychology of language: Formal and experimental approaches*. New York: Plenum Press, 1978.
203. Norman, D. A. and Rumelhart, D. E. *Explorations in cognition*. San Francisco: Freeman, 1975.
204. Olson, D. R. and Filby, N. On the comprehension of active and passive sentences. *Cognitive Psychology*, 1972, 3, 361-381.
205. Olson, G. M. and Clark, H. H. Research methods in psycholinguistics. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception*. Vol. 7: *Language and speech*. New York: Academic Press, 1976.
206. Olson, J. N. and McKay, D. G. Completion and verification of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 457-470.
207. Paivio, A. and Begg, I. Imagery and comprehension latencies as a function of sentence concreteness and structure. *Perception and Psychophysics*, 1971, 10, 408-412.
208. Petrick, S. R. *A recognition procedure for transformational grammars*. Dept. of Modern Languages, M.I.T., 1965.
209. Reber, A. S. Locating clicks in sentences: Left, center and right. *Perception and Psychophysics*, 1973, 13, 133-138.
210. Reber, A. S. and Anderson, J. R. The perception of clicks in linguistic and non-linguistic messages. *Perception and Psychophysics*, 1970, 8, 81-89.
211. Riding, R. J. A method for investigating the perceptual segmentation of speech. *Language and Speech*, 1975, 18, 153-157.
212. Riesbeck, C. K. Conceptual analysis. In R. C. Schank (Ed.), *Conceptual information processing*. Amsterdam: North-Holland, 1975.
213. Rips, L. J. Qualification and semantic memory. *Cognitive Psychology*, 1975, 7, 307-340.
214. Rips, L. J., Shoben, E. J., and Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
215. Roncato, S. and Sonino, M. On some strategies of comparing sentences against pictures. *Italian Journal of Psychology*, 1976, 3, 165-186.
216. Rosenberg, S. and Jarvella, R. J. Semantic integration as a variable in sentence perception, memory and production. In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
217. Rosenberg, S. and Jarvella, R. J. Semantic integration and sentence perception. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 548-553.
218. Rubenstein, H., Garfield, L., and Millikan, J. A. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 487-494.
219. Rubenstein, H., Lewis, S. S., and Rubenstein, M. A. Homographic entries in the internal lexicon: Effects of systematicity and relative frequency of meanings. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 57-62.
220. Russell, B. *Human knowledge: Its scope and limits*. Allen and Unwin, 1948.
221. Sachs, J. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 1967, 2, 437-442.
222. Savin, H. B. and Bever, T. G. The non perceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 295-302.
223. Savin, H. B. and Perchonock, E. Grammatical structure and the immediate recall of

- English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 348-353.
224. Schank, R. C. Conceptual dependency". A theory of natural language understanding. *Cognitive Psychology*, 1972, 3, 552-631.
225. Schank, R. C. Identification of conceptualizations underlying natural language. In R. C. Schank and K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973.
226. Schank, R. C. *Conceptual information processing*. Amsterdam: North-Holland, 1975.
227. Schvaneveldt, R. W. and Meyer, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and Performance IV*. New York: Academic Press, 1973.
228. Schwartz, D., Sparkman, J. P., and Deese, J. The process of understanding and judgment of comprehensibility. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 87-93.
229. Seitz, M. R. and Weber, B. A. Effects of response requirements on the location of clicks superimposed on sentences. *Memory and Cognition*, 1974, 2, 43-46.
230. Shanon, B. The two meanings of a homophone. *Perception and Psychophysics*, 1974, 16, 571-574.
231. Sherman, M. A. Bound to be easier? The negative prefix and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 76-84.
232. Sherman, M. A. Adjectival negation and the comprehension of multiply negated sentences. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 143-157.
233. Shields, J. L., McHugh, A., and Martin, J. G. Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology*, 1974, 102, 250-255.
234. Smith, K. H. and McMahan, L. E. Understanding order information in sentences: Some recent work at Bell Laboratories. In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
235. Soli, S. D. and Balch, W. R. Performance biases and recognition memory for semantic and formal changes in connected discourse. *Memory and Cognition*, 1976, 4, 673-676.
236. Springston, F. J. Some cognitive aspects of presupposed coreferential anaphora. Unpublished Ph.D. Dissertation, Stanford University, 1975.
237. Springston, F. J. and Clark H. H. *And* and *or*, or the comprehension of pseudo-imperatives. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 258-272.
238. Sternberg, S. The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, 1969, 30, 276-316.
239. Stevens, A. L. and Rumelhart, D. E. Errors in reading: An analysis using an augmented transition network model of grammar. In D. A. Norman and D. E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: Freeman, 1975.
240. Stillings, N. A. Meaning rules and systems of inference for verbs of transfer and possession. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 453-470.
241. Streeter, L. A. and Bever, T. G. The effects on the detection of linguistic and non-linguistic stimuli are opposite at the beginning and end of a clause. Mimeographed, Columbia University, 1975.
242. Sturges, P. T. and Martin, J. G. Rhythmic structure in auditory pattern perception and immediate memory. *Journal of Experimental Psychology*, 1974, 102, 377-383.
243. Tanenhaus, M. K. and Carroll, J. M. The clausal processing hierarchy and nouniness. In R. Grossman, J. San, and T. Vance (Eds.), *Papers from the parasession on functionalism*. Chicago: Chicago Linguistic Society, 1975.
244. Tanenhaus, M. K., Carroll, J. M., and Bever, T. G. Sentence-picture verification

- models as theories of sentence comprehension: A critique of Carpenter and Just. *Psychological Review*, 1976, 83, 310-317.
245. Thorne, J. P., Bratley, P., and Dewar, H. The syntactic analysis of English by machine. In D. Michie (Ed.), *Machine Intelligences*. New York: American Elsevier 1968.
246. Toppino, T. C. The underlying structures of sentences are not the primary units of speech processing: A reinterpretation of Bever, Lackner, and Kirk's findings. *Perception and Psychophysics*, 1974, 15, 517-518.
247. Trabasso, T. Mental operations in language comprehension. In R. O. Freedle and J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. New York: Winston, 1972.
248. Trabasso, T., Rollins, H., and Shaughnessy, E. Storage and verification stages in processing concepts. *Cognitive Psychology*, 1971, 2, 239-289.
249. Treisman, A. M. and Squire, R. Listening to speech at two levels at once. *Quarterly Journal of Experimental Psychology*, 1974, 26, 82-97.
250. Treisman, A. M., Squire, R., and Green, J. Semantic processing in dichotic listening? A replication. *Memory and Cognition*, 1974, 2, 641-646.
251. Tversky, B. Encoding processes in recognition and recall. *Cognitive Psychology*, 1973, 5, 275-287.
252. Tversky, B. Pictorial encoding of sentences in sentence picture comparisons. *Quarterly Journal of Experimental Psychology*, 1975, 27, 405-410.
253. Wales, R. and Grieve, R. What is so difficult about negation? *Perception and Psychophysics*, 1969, 6, 327-332.
254. Wang, M. D. Influence of linguistic structure on comprehensibility and recognition. *Journal of Experimental Psychology*, 1970, 85, 83-89.
255. Wang, M. D. The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 398-404.
256. Wannemacher, J. T. Processing strategies in picture-sentence verification tasks. *Memory and Cognition*, 1974, 2, 554-560.
257. Wannemacher, J. T. Processing strategies in sentence comprehension. *Memory and Cognition*, 1976, 4, 48-52.
258. Wanner, E. Review of Fodor, J. A., Bever, T. G., and Garrett, M. F. *The psychology of language*. New York: McGraw-Hill, 1974. *Journal of Psycholinguistics Research*, 1977, 3, 261-270.
259. Wanner, E., Kaplan, R., and Shiner, S. Garden paths in relative clauses. Mimeographed. Harvard University, Cambridge, Mass., 1975.
260. Wanner, E. and Maratsos, H. An augmented transition network model of relative clause comprehension. Mimeographed. Harvard University, Cambridge, Mass.. 1975.
261. Wanner, E. and Shiner, S. Measuring transient memory load. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 159-167.
262. Wanner, E. and Shiner, S. Ambiguities in relative clauses. Mimeographed. Harvard University, Cambridge, Mass., 1975.
263. Wason, P. C. and Johnson-Laird, P. N. *Psychology of reasoning*. London: Batsford, 1972.
264. Wickelgren, W. A. The long and the short of memory. *Psychological Bulletin*, 1973, 80, 425-438.
264. Wilks, Y. An artificial intelligence approach to machine translation. In R. C. Schank and K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973.
266. Wingfield, A. The intonation-syntax interaction: Prosodic features in perceptual processing of sentences. In A. Cohen and S. G. Nooteboom (Eds.), *Structure and process in speech perception*. Berlin, Heidelberg, New York: Springer Verlag, 1975.

267. Wingfield, A. and Klein, J. F. Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics*, 1971, 9, 23-25.
268. Winograd, T. Understanding natural language. *Cognitive Psychology*, 1972, 3, 1-191.
269. Winograd, T. A procedural model of language understanding. In R. C. Schank and K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973.
270. Woods, W. A. Transition network grammars for natural language analysis. *Communications of the ACM*, 1970, 13, 591-606.
271. Wright, P. Some observations on how people answer questions about sentences. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 188-195.
272. Wright, P., Holloway, C. M., and Aldrich, A. R. Attending to visual or auditory verbal information while performing other concurrent tasks. *Quarterly Journal of Experimental Psychology*, 1974, 26, 454-463.
273. Wright, P. and Kahneman, D. Evidence for alternative strategies of sentence retention. *Quarterly Journal of Experimental Psychology*, 1971, 23, 197-213.
274. Zwanenburg, W., Ouweneel, G. R., and Levelt, W. J. M., La frontière du mot en français. *Studies in Language*, 1977, 1, 209-221.

CONTENTS

1. Aims	1
2. Tasks and dependent variables	4
2.1 <i>Simultaneous measurement</i>	5
2.2 <i>Successive measurement</i>	8
3. Linguistic and experimental variables	21
3.1 <i>Prosody and rhythm</i>	22
3.2 <i>Syntactic complexity</i>	23
3.3 <i>Ambiguity</i>	29
3.4 <i>Constituent and clause structure</i>	37
3.5 <i>Pictorial context</i>	38
3.6 <i>Non-pictorial context. Given and new information</i>	46
3.7 <i>Some conclusions</i>	47
4. Theoretical considerations	47
4.1 <i>Immediate understanding and depth of processing</i>	48
4.2 <i>Comprehension in context</i>	53
4.3 <i>Formal theories of sentence understanding</i>	54
4.4 <i>Conclusions</i>	60
5. References	62