Chapter 6

The Perception of Relations: The Interaction of Structural, Functional, and Contextual Factors in the Segmentation of Sentences

John M. Carroll, Michael K. Tanenhaus, and Thomas G. Bever

THE ORGANIZATION OF THE STIMULUS: FROM CHUNKING TO SENTENCE SEGMENTATION

Much theory and research in cognitive psychology is predicated on the assumption that human information-processing capacities are limited relative to the apparent complexity of human behaviour and experience. This is particularly clear if the limits of working memory capacity are contrasted with the length and complexity of the sequential stimuli that people are able to process without apparent difficulty. For example, recall memory for a sequence of random words is limited to about eight. However, this direct assessment of 'working memory' capacity does not jibe with actual human verbal performance. People readily produce and understand sentences containing many more than eight words. One fundamental task of cognitive theory is to describe a mechanism which organizes basic and limited processing capacities to create complex human behaviour and deal with ordinary experience.

Miller (1956) proposed such a mechanism in his 'chunking' analysis. He noticed that when items in a sequence are grouped, or chunked, together, subjects can successfully deal with more items. He inferred that these chunked items are recorded into higher order units, as schematized in (1).

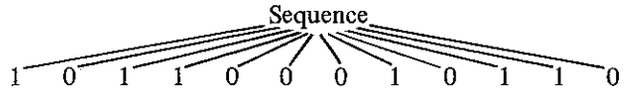
(1) Input Sequence → Recoding Scheme → Chunks

A large number of unstructured items in the input sequence is distilled, as it were, into a smaller number of structured items. But such chunks can be chunked themselves. The hierarchical and iterative imposition of structure on sequences effectively allows for *unlimited* processing capacity. Thus, Miller's

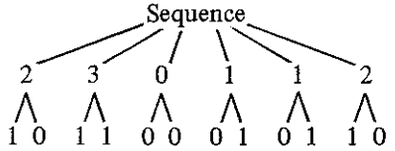
chunking proposal offers a potential framework in which to describe the organization of complex human information processing. (Of course, the application of chunking has to be restricted in some manner or else the chunking analysis will predict an infinite memory capacity.)

Miller reports an experiment by Sidney Smith that illustrates how chunking can increase processing capacity, in this case, short-term memory capacity. Smith presented subjects with sequences of binary digits, ones, and zeros, which they were to keep track of. Not surprisingly, he found that they were able to recall about eight digits. He then trained his subjects to recode pairs of binary digits into base four digits, thus: 00 → 0, 01 → 1, 10 → 2, 11 → 3. In this condition, recall performance goes up. When subjects were told to recode triples of binary digits into octal numbers (i.e. 000 → 0, 001 → 1, 010 → 2, 011 → 3, 100 → 4, 101 → 5, 110 → 6, 111 → 7), recall was even better. Smith's recording schemes are illustrated in (2).

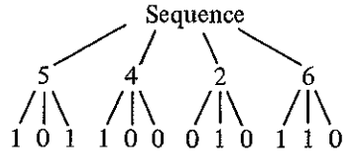
(2) a. No Instructions



b. Base four Recording



c. Octal Recording



Lists of digits have very minimal intrinsic structure. In Smith's experiment, the recoding schemes, and therefore the nature of the chunking units, are defined by the experimenter. This is just the sort of example Smith and Miller needed to bring into relief the potential advantages of chunking to human processing: sequences that would otherwise exceed working memory capacity can be processed when structured by recoding schemes.

Miller's chunking analysis can be easily extended to the processing of *complex* stimuli, i.e. stimuli with apparent intrinsic structure. As items enter working memory, the processor groups or chunks them into higher order units.

However, the recoding schemes which define these units, unlike the recoding schemes Smith defined, must enable the processor to recover the intrinsic structure of the sequence, in addition to the literal items constituting the sequence. The relationships, for example, between the words and phrases of a sentence or the contours and objects in a visual scene have such intrinsic structure. In order to comprehend a sentence or a visual scene the perceiver must recover not only the components, but also the relations between them. The task of perceptual theory is to describe how the relations between items in a structured array are recovered given the limitations of the perceptual system.

Returning to Miller's analysis, one might expect there to be a rather close correspondence between the major structural units of the stimulus and the chunking units constructed by perceptual recoding schemes. The structural units of the stimulus are defined by the relations which the perceiver must recover in order to comprehend the stimulus. The linguistic constituents of a sentence enter into various grammatical relations like subject-of, object-of, and modifier-of. The constituent contours and objects in a visual scene enter into various spatial relations. Of course, the exact nature of this correspondence between structural and perceptual units will be determined by the extent to which the structural units of the stimulus can be used as effective chunking units by the perceptual system (i.e. can be constructed by perceptual recoding schemes). In order to study the nature of this correspondence, we need to consider the perception of a complex stimulus whose intrinsic structure is fairly well understood.

From this perspective, the psychological study of language provides an excellent case study. Language is clearly an intrinsically structured and complex human capacity. The science of linguistics is fundamentally concerned with describing the intrinsic structure of language. Indeed, one of the central concerns of the last 15 years of psycholinguistic research has been to describe the role of linguistic structures in behavioural theories of language.

In particular, several theorists have applied Miller's chunking analysis to the description of sentence perception. Basically, the hypothesis explored is that some particular structural unit in the linguistic description of the sentence functions as a chunking unit during sentence perception. (In the area of sentence perception these chunking units are often called *segmentation units*.) Bever, Lackner, and Kirk (1969) and Fodor, Bever, and Garrett (1974) have developed the claim that the listener treats the linguistic clause as a segmentation unit. Thus, on their model, sentence perception consists in part of isolating and segmenting together word sequences which correspond to linguistic clauses. Fodor and Bever (1965) and Chapin, Smith, and Abrahamson (1972) have also defined the segmentation unit in terms of a level of linguistic structure. They propose that major surface-structure constituents are segmentation units in sentence perception (see Carroll and Bever, 1976, for review).

Both of these hypotheses about the nature of the segmentation units of sentence perception are couched in purely *structural* terms. They describe a

simple and straightforward relation between what we earlier referred to as intrinsic stimulus structure (in this case, syntactic structure) and the determination of recoding schemes for chunking (in this case, segmentation). But they purchase this simplicity at the price of overlooking potentially important *functional* segmentation variables. Recall that structural units should be expected to function as perceptual units just to the extent that they can be effectively organized by the perceptual system. Neither the clause nor the constituent satisfies this requirement.

First, consider the constituent theory. In its strongest form, this theory claims that major surface-structure constituents are the segmentation units of sentence perception. Thus, in the course of perceiving a sentence like (3), the listener is segmenting the initial eight words as a noun-phrase surface constituent and the final eight words as a verb phrase constituent.

(3) The big old worn and splintering carriage wheel rolled down the hill and across the road.

Thus, a sentence 16 words long is intermediately organized into two more manageable units of eight words each. Sentences like (4), however, present problems for this analysis.

(4) The big old worn and splintering heavy wooden carriage wheel with the rusted spokes is a genuine antique.

In sentence (4) the initial noun-phrase constituent is itself 14 words long. Indeed, noun-phrase constituents can be arbitrarily long and complex. Accordingly, there are cases, like (4), in which it is preposterous to claim that major constituents are segmentation units.

The clausal theory is no better off, although in many cases it too makes intuitively satisfying predictions. For example, in (5) a 16-word sentence is rendered as two clausal segmentation units of eight words each.

(5) After John hung up the clothes to dry he went inside to start cooking his supper.

However, surface-structure clauses, like noun-phrase surface constituents, can be arbitrarily long. The initial surface clause in (6) is 18 words long.

(6) After the long dry summer of our sorrow and utter discontent had finally come to a welcomed end we simply moved to Pittsburg.

Consequently, the limitations of processing capacity would be exceeded before segmentation could occur. If, as Bever *et al.* (1969) suggest, the pertinent notion

of clause is actually the deep-structure clause, other problems appear. On a deep-clause segmentation theory, the initial sequence of nine gerunds in (7) would correspond to nine separate segmentation units since each corresponds to a deep-structure clause.

(7) Walking, talking, eating, drinking, hurrying, studying, listening, theorizing and sleeping were John's daily activities.

Here the number of segmentation units which would have to be integrated into a perceived sentence will exceed processing capacity. Instead of predicting too many words per hypothesized segmentation unit, the deep-clausal theory will predict too many hypothesized segmentation units per sentence.

The objections we have just raised regarding the clausal and constituent theories address issues of processing capacity. A second line of argument centres on whether clausal or constituent segmentation units enable the listener to recover the defining relations of the sentence. In sentence (7) each of the initial nine gerunds is a surface noun-phrase constituent which corresponds to a deep-structure clause. Each has the noun 'John' as its subject relation. However, this relation is not specified until the final three words of the sentence. It is not clear how these initially subjectless gerunds could be perceptually organized as complete clausal segmentation units. The deep-structure clause position encounters additional difficulties with sentences like (8a).

(8) a. *The tall* man ordered a Scotch.
b. (the man(the man is tall) ordered a Scotch)
S_1 S_2 S_2 S_1

In sentence (8a), the adjective 'tall' corresponds to the deep-structure clause S_2 in (8b), which presents a deep-structure bracketing of (8a). Thus, the deep-structure clause theory makes the unlikely prediction that the sequence 'the tall' should be segmented and recoded as a perceptual unit.

It is important to note that we have attacked the *strongest* versions of the two structural-linguistic segmentation theories (see Carroll, 1976, for discussion). However, it seems reasonable to proceed in this fashion if we want to bring into relief the directions in which the theory of sentence perception must now be elaborated. The structural-linguistic segmentation theories provide solid ground from which to begin this elaboration: comprehending a sentence necessarily involves representing the information contained in the sentence's linguistic structure (i.e. one cannot be said to have understood a sentence unless one tacitly knows which noun was the object of a given preposition, the subject of the main verb, etc.; see Fodor, 1968). Thus, whatever the segmentation units of sentence perception are like, they are logically entrained by linguistic structures.

In the remainder of this chapter we will present an initial functionalist elaboration of the linguistic-structural theories of segmentation. On this view, the segmentation structure of a sentence is determined by a variety of factors including, but not limited to, the intrinsic structure of the sentence (see Carroll, 1976; Tanenhaus and Carroll, 1975). Indeed, we will show that the strong versions of the linguistic-structural segmentation theories can be experimentally rejected. We adopt a perceptual cue approach; we ask 'what are the cues utilized by the listener to identify complete propositions reflected in good potential segmentation units?' and 'how do these cues interact to predict the segmentation structure of perceived sentences?'

FUNCTIONAL AND CONTEXTUAL FACTORS IN SEGMENTATION

According to the chunking analysis, sequences of items are recoded into higher order units in a way that conserves processing capacities, like working memory. From this it follows that a sequence of words will comprise a better segmentation unit if it is potentially recodable into an independent memory structure. Current views of memory have identified such memory units with propositional structures (e.g. Kintsch, 1974; Rumelhart, Lindsay, and Norman, 1972). Thus, we make the assumption that linguistic sequences which can be directly mapped onto complete propositional structures are the ideal segmentation units in sentence perception. We refer to such sequences as *functionally complete*.

Functionally complete sequences contain a complete, explicit, and coherent set of grammatical relations. Functionally incomplete sequences fail to meet this requirement. The two initial clauses in example (9) are both examples of functionally complete sequences.

(9) a. After John spilled the beans, everyone ignored him.
b. The town's construction of a new school cost the taxpayers a mint.

In (9a), 'John' is the subject relation, 'spilled' is the verb and 'the beans' is the object. In (9b), 'the town' is the subject, 'construction' is the verb and 'a new school' is the object.

The initial three-word sequence in sentence (10a) is functionally incomplete in that it lacks an explicit verbal relation.

(10) a. The big dog bit me on the knee.
b. Fleeing was John's alternative to fighting.

It corresponds to a proposition which can be glossed as 'the dog is big', but as it appears in sentence (10a), it lacks the copulative verb 'is' which relates the noun 'dog' to its adjectival modifier. Similarly, the sentence-initial gerund 'fleeing' in

sentence (10b) provides the listener with a functionally incomplete sequence (recall sentence (7)). 'Fleeing' in sentence (10b) corresponds to the proposition 'John flee'. But the subject relation of this proposition is not present in the manifest initial sequence of the sentence. It is not specified until later in the sentence.

A possible elaboration of the functional completeness property concerns the *specificity* with which grammatical relations are represented in a sequence. Thus, the italicized noun phrases in (11a) through (11c) seem to be increasingly less specific (*qua* sentences out of context, see below). Indeed, unless the noun 'John' in (11d) is understood as referring to a particular person, it is quite abstract, and unspecific, for this listener.

(11) a. After *the little fellow with a moustache* left, we realized what a great guy he was.
b. After *the man* left, we realized what a great guy he was.
c. After *he* left, we realized what a great guy he was.
d. After *John* left, we realized what a great guy he was.

Ceterus paribus, sequences with more specific grammatical relations may be better potential segmentation units than sequences with less specific grammatical relations. Similarly, the difference in specificity between the italicized object noun phrases in (12) may render the clause in (12a) a better potential segmentation unit than that in (12b).

(12) a. After he bought *some gum*, he went home.
b. After he bought *some*, he went home.

(See Marslen-Wilson, Tyler, and Seidenberg, 1976; Tanenhaus and Seidenberg, 1978.) In this paper, we do not treat specificity, aside from specificity via context (see section on contextual cues).

In considering the role of the functional completeness property in structuring sentence segmentation, we will be concerned, first, with how functional completeness might be recognized in a word sequence during sentence perception, and, second, with how functional completeness might be compromised in the course of sentence segmentation. We will approach these questions by examining two sorts of perceptual segmentation cues: cues which describe the internal structure of sequences (specifically, whether or not given sequences are functionally complete) and cues which indicate how much processing resource a particular sequence requires.

There are quite clearly a variety of potential cues that can serve to inform the listener about the internal structure of a linguistic sequence. With particular regard to the property of functional completeness, the listener might be cued by a 'noun-verb-noun' configuration. Bever (1970) showed that listeners assume

that a sequence whose lexical configuration is noun-verb-noun corresponds to a canonical set of propositional relations: subject-predicate-object. Thus, listeners apparently interpret noun-verb-noun configurations as potential functionally complete segmentation units. They use the 'noun-verb-noun' configuration as a cue to functional completeness.

But there are other potential cues available to the listener. Consider, for example, the information available by way of context and inference. As we have noted earlier, the sentence initial noun-phrase sequence in (10a) is functionally incomplete because the copulative verb relation between the noun 'dog' and the modifier 'big' is not explicitly present in the sequence.

The deletion of the copulative verb, however, is a regular and general syntactic process and hence the omitted relation can be recovered and 'filled in' by inference. Such inferential cues, if they are used by listeners in perceptually organizing sentences, could effectively render certain functionally incomplete sequences functionally complete. By assumption, this would make sentence comprehension an easier task for the listener.

Certain other functionally incomplete sequences can be effectively rendered functionally complete by way of contextual cues. For example, the initial word of sentence (10b) is a functionally incomplete sequence. However, if the deleted subject relation 'John' is provided somehow, then the incomplete sequence 'Fleeing' can be organized into the complete propositional segmentation unit 'John flee'. Context could help do this in several ways, either by prior discourse (i.e. the sentence occurs in a talk about John), ostension (i.e. John is there) or even by means of the information provided in the following clause of (10b).

Another class of cues that can inform the listener of a sequence's internal structure can be referred to as *local sign cues*. Lexical items, like nouns and verbs, are themselves local sign cues which denote functionally complete sequences: a sequence cannot be functionally complete unless it contains *some* nouns and verbs. Thus, quite aside from what configurational cues are present (e.g. noun-verb-noun), the simple fact that nouns and verbs occur cues the property of functional completeness. Clearly, these lexical local signs are not perfectly valid cues of functional completeness. The initial noun phrase of (10a) contains the noun 'dog' but is functionally incomplete (in that it contains no explicit verbal relation). And the initial headless nominalization of (10b) contains the verbal element 'fleeing' but is also functionally incomplete (in that it contains no subject noun).

There may be further local sign cues below the level of the word. Consider the tense morpheme component of verbs. Tense morphemes actually have a greater cue validity than their stems in signalling a functionally complete sequence: if a verbal element bears a tense morpheme then its surface-subject relation is always explicitly present in the word sequence. Conversely, as sentences like (10b) demonstrate, the occurrence of a verbal stem does not guarantee functional completeness. However, the occurrence of a tensed verbal element does indeed guarantee this. (Note though, that certain functionally complete sequences do *not* contain tensed verbal items, as in '*The men's refusing the offer* cost the company a lot.')

The classes of cues we have considered thus far provide information that the listener can use to recognize functional completeness in a sequence: they are cues to the internal structure of sequences. However, these cues interact with cues that indicate how much processing capacity a particular sequence requires. We assume that the listener's processing attempts to optimize the amount of processing capacity expended on each segmentation unit. If too much capacity is required per segmentation unit, processing breaks down. If too little is required, capacities go unused, and processing is less efficient than it could be.

Consider working memory. If a sequence is very long, working memory can become overtaxed. In order to avert breakdown, the processing capacity cue of sequence length can induce segmentation, perhaps compromising functional completeness. Thus, the sentence-initial noun phrase in (3) may indeed serve as a segmentation unit even though it is functionally incomplete. The initial eight words of sentence (4) 'The big, old, worn and splintering, heavy, wooden' might even serve as a segmentation unit of necessity, though this sequence has virtually no 'intrinsic' structure at all. In contrast, if a sequence is very short, it may fail to stimulate segmentation processing even if it *is* functionally complete. Thus, the first five words of sentence (13) may be segmented as one unit even though these five words comprise two distinct functionally complete sequences.

(13) *John returned and Mary arrived* just in time for both of them to meet the mayor of Utica.

Like sequence length, sequence complexity may act as a processing capacity cue. If a sequence becomes too complex, segmentation units may not be isomorphic with functionally complete sequences. (See Carroll, 1976, and Carroll and Bever, 1976, for further speculation and discussion of processing capacity cues.)

Before proceeding, we must make two disclaimers. First, we have certainly not presented an exhaustive catalogue of segmentation cues. Much remains to be done just to understand the role in sentence perception of the cues we have identified, but even with this work complete we would not have a comprehensive theory of segmentation cues. Second, we are not committed to an all-or-none model of segmentation. One may ultimately need to talk about amount or degree of segmentation, rather than simply about whether or not segmentation obtains.

We now turn to some experimental work-in-progress designed to provide experimental evidence against purely linguistic definitions of the segmentation units of sentence perception and to validate initially the role of our functional cues in sentence perception.

SOME EXPERIMENTS

The configurational N-V-N cue

The first two experiments we will present were designed to test the claim that the N-V-N configurational cue perceptually denotes the functional completeness property. Thus, our main prediction is that noun-verb-noun functionally complete sequences will be better segmentation units than functionally incomplete sequences. We used the location technique since the linguistic models of segmentation are fundamentally based on the results of research using this technique (see Fodor and Bever, 1965; Bever *et al.*, 1969; Chapin *et al.*, 1972).

Fodor and Bever observed that when a click is superimposed on a recorded sentence, listeners systematically mislocate the click with regard to the sentence sequence. Fodor and Bever account for this by arguing that clicks are phenomenally displaced towards a segmentation boundary: the relatively high perceptual coherence of a segmentation unit 'resists' interruption by the click.

We presented subjects with sentences which in most cases contained a brief tone (see Carroll, 1976; and Carroll and Tanenhaus, 1978; for details of method and procedure). In the experimental sentences there are two tone locations: just before the final word of the sentence-initial sequence and just after the first word of the sentence-final sequence. The four types of initial sequences we studied are listed in (14), with tone locations indicated by slashes.

(14) a. Main Clauses (Functionally Complete N-V-N):
I felt very sorry for the old/bum so/I gave him a dime.
b. Subordinate Clauses (Functionally Complete N-V-N):
After the crook stole the woman's/bag he/ran for safety.
c. Headless Nominalization (Functionally Incomplete):
Meeting the pretty young/girl was/the highlight of Peter's trip.
d. Noun Phrase (Functionally Incomplete):
The old painted wooden/pipe was/on display at the local museum.

Our experimental sentences were divided evenly among the four sequence-type categories in (14). Thus, half of the sentences had functionally complete N-V-N initial sequences and half had functionally incomplete initial sequences. We also manipulated sequence length and varied it orthogonally with initial sequence-type. Half of our experimental sentences had 'long' initial sequences (eight words and ten syllables) and half had 'short' initial sequences (five words and seven syllables).

Following the Fodor and Bever procedure, subjects listened to the sentences through headphones and then wrote them out from memory in a booklet. If a tone had occurred in the sentence, the subject then marked the tone location with a slash in his written script. Tones were not placed in all sentences in order to avoid an increase in the probability of a tone with serial position in the sentence (Abrams, 1973).

We modified our location-scoring procedure from that of Fodor and Bever. Recent research indicates that the behavioural basis of the manifest location effect may be more complex than Fodor and Bever originally supposed. In particular, it may be that subjects tend to assume that interruptions of whose location they are actually uncertain occurred while they were busy processing. Abrams and Bever (1969), Bever and Hurtig (1975) and Seitz (1972) have shown that sensitivity to interruptions is lowest just at the end of a presumed segmentation boundary. Thus, part of the location effect may be due to a 'response bias' for locating interruptions at points of relatively high segmentation processing (see Bever, Hurtig, and Handel, in preparation).

To accommodate this possibility we count mislocations of tones into or towards the 'processing zone' implicated by the work of Abrams and Bever, Bever and Hurtig, and Seitz, namely the last word of the initial sequence and the sequence boundary itself, as confirming the hypothesis that the sequence boundary is a segmentation boundary. Any mislocation away from this processing zone is scored as disconfirming the segmentation hypothesis. Correct subjective locations and overshoots (mislocations which go towards the boundary and beyond it out of the processing zone) are not scored.

The results of Experiment I are summarized in Table 6.1 in percentages. A Friedman two-way analysis of variance by ranks for k correlated samples (see Ferguson, 1971, pages 331–335) reveals an overall difference in tone mislocation across the four sequence-types by-subject (i.e. taking subjects as the unit of analysis), $p < .05$. However, the Anova taking items, or sentences, as the unit of analysis (i.e. by-item) fails to reveal an overall difference (Kruskall-Wallis one-way Anova by ranks for k uncorrelated samples, see Ferguson, 1971, pages 331–335).

Since the functional completeness and sequence length contrasts were planned comparisons, further tests were performed. For functionally complete initial sequences 69% of the errors analysed were mislocated into or towards the processing zone. However, for functionally incomplete initial sequence-types this figure is only 56%. This difference is significant by subject, $p < .01$

Table 6.1 Experiment I. Percentages of tones mislocated into or towards the processing zone, for the different types of clauses

Functionally complete 69%		Functionally incomplete 56%	
Long 65%	Short 70%	Long 50%	Short 60%
Mains 74%	Subordinates 61%	Noun phrases 60%	Nominalizations 51%

(Wilcoxon signed ranks test, correlated samples) and by-item, $p < .005$ (Wilcoxon rank sums test, uncorrelated samples). The effect of sequence length is less encouraging. An overall Anova for the four functional completeness X sequence length possibilities (long functionally complete, short functionally complete, long functionally incomplete, short functionally incomplete) is non-significant. Internal paired comparisons and an overall long versus short contrast are also non-significant.

We draw several conclusions from this experiment. First, the results of Experiment I are fully consistent with our claim that functionally complete N-V-N sequences should be better segmentation units than functionally incomplete sequences. Thus, our approach is encouraged. However, on the basis of this experiment alone we cannot uniquely determine what sort of cue is responsible for this functional completeness effect. It appears that contextual/inferential cues are *not* responsible, since on the basis of context and inference the initial sequences of (14c) and (14d) would indeed be functionally complete.

The functionally complete experimental items in Experiment I clearly do differ from the functionally incomplete items in terms of the N-V-N configurational cue: only the former items provide the listener with an N-V-N cue. However, this differentiation confounds with other differences between the functionally complete and functionally incomplete items of Experiment I. Although both functionally complete and incomplete sequences contain lexical nouns (cf. 'bum', 'bag', 'girl', 'pipe' in 14), there are also local sign differences between the two classes. Main clauses and subordinate clauses always contain lexical verbs, and these verbs are always tensed verbs. In the functionally incomplete cases, only the nominalizations contain lexical verbs and these verbs bear a nominalizing morpheme (e.g. '-ing' in 14c) instead of tense. Also, functionally incomplete noun-phrase sequences, like (14d), tend to contain more lexical modifiers (e.g. 'old', 'painted', 'wooden'). Thus, while Experiment I establishes the pertinence of functional completeness to segmentation organization, it leaves open whether the efficient cue is a configuration (e.g. N-V-N) or a local sign (e.g. verb tense), or, perhaps more reasonably, some combination of the two. (It is not clear that any notion of configuration—such as a propositional segmentation structure—*can* be constructed out of purely local sign elements, Köhler, 1929.)

The apparent failure of Experiment I to reveal an effect of sequence length may simply indicate that the tone-location response measure is relatively insensitive to processing capacity cues. However, we must also consider the possibility that our hypothesized sequence-length cue has limited construct validity as a principle of sentence segmentation. In either case, of course, the only way to resolve this question is to study sequence length using other experimental response measures. Some of this research will be reported below.

A final conclusion we draw from Experiment I is that the two structural-linguistic segmentation theories are manifestly inadequate. First, consider the

clausal model. In most theories of grammar *all* of the four initial sequence-types in (14) correspond to deep-structure clauses. (Nominalizations are formed by a transformational rule of Nominalization. Noun phrases are formed by Relative Clause Reduction followed by Adjective Preposing. And sentence-initial subordinate clauses are formed by a rule of Adverb Preposing. See Stockwell, Schachter, and Partee, 1973.) Thus, adopting the deep-clausal view offers no way of distinguishing nominalizations and noun phrases from main and subordinate clauses. Therefore, it cannot account for the results of Experiment I.

A surface-clausal view could distinguish noun phrases, which are not surface clauses, from the other three sequence-types, which are. However, it cannot distinguish nominalizations from main and subordinate clauses. This contrast is significant by-subject, $p < .05$ (Wilcoxon signed ranks test, correlated samples), and nearly so by-item, $p < .1$ (Wilcoxon rank sums test, uncorrelated samples).

Finally, consider the constituent theory of Fodor and Bever and Chapin *et al.* Both of our functionally incomplete sequence-types comprise major surface constituents (noun phrases and nominalizations) but neither of our functionally complete sequence-types do. Thus, the constituent theory predicts the obverse of what was found.

Experiment II was designed to confirm and elaborate the functional completeness result of Experiment I. First, it replicates the functionally complete versus incomplete difference discovered in Experiment I. In doing so, it also tries to control more thoroughly structural-linguistic variables, and thus to establish the independent importance of functional completeness in sentence segmentation. Finally, it attempts to unconfound the configurational cue to functional completeness (N-V-N) and the local sign cue (verb tense), and therefore to demonstrate the independent contribution of different cues. The four types of initial sequences we examined are listed in (15), with tone locations indicated by slashes.

(15) a. Relative Clause (Functionally Complete, Tensed Verb):
The judge who resisted the l/arge bribe was q/uite well respected by his colleagues.
b. Noun Phrase (Functionally Incomplete):
The very pretty and talented b/londe girl was t/he runner-up in the Miss Wisconsin contest.
c. Headed Nominalization (Functionally Complete, Nominalized Verb):
Humphrey's refusal to take a f/irm stand caused h/im to lose the election.
d. Headless Nominalization (Functionally Incomplete):
The overall reaction to the n/ew play was m/ore positive than negative.

The experimental sentences for Experiment II were divided equally among the four categories in (15). (Stimulus preparation, procedure, and scoring are identical to that of Experiment I except that tones were placed in words rather than word boundaries. We found that tones could be placed more accurately and easily in words, since they are more integral acoustic events—for further discussion see Carroll, 1976, and Carroll and Tanenhaus, 1976, 1978.)

The results of Experiment II are summarized in Table 6.2 in terms of percentages. The Friedman two-way Anova by ranks for the four sequence categories is significant by-subject, $p < .05$. The Kruskall-Wallis one-way Anova by ranks is significant by-item, $p < .05$.

The main result for functional completeness is also summarized in Table 6.2. For functionally complete items (relative clauses and headed nominalizations), 78% of the errors analysed were mislocations into or towards the hypothesized processing zone. The corresponding figure for functionally incomplete sequences (noun phrases and headless nominalizations) is only 65%. This difference is significant by-subject, $p < .005$ (Wilcoxon Signed Ranks Test), and by-item, $p < .001$ (Wilcoxon rank sums test). Thus, Experiment II succeeds in replicating the functional completeness result of Experiment I.

A further issue, however, concerns just what sort of cue to functional completeness is the efficient experimental variable. Consider the differences between relative clauses and headed nominalizations. Both consist of N-V-N configurations (i.e. 'Judge resist bribe' and 'Humphrey refuse take—stand'), but they differ in terms of the local sign cues they provide. Relative clauses contain tensed verbs (e.g. 'resist*ed*') and relative pronouns (e.g. 'who'). Headed nominalizations contain the possessive morpheme (e.g. 'Humphrey'*s*') and nominalization morphemes (e.g. 'refus*al*'). We did not make predictions about all of these local sign differences, but we have noted earlier that tense morphemes, since they imply functional completeness, might serve as effective segmentation cues.

The results of Experiment II cannot distinguish between relative clause cases and headed nominalizations (see Table 6.2). Thus, we find no evidence for any of the above-mentioned local sign cues. On this basis, we conclude that the results of Experiment II implicate the configurational N-V-N cue. (We do not give up on local sign cues, however: see Experiment III below.)

Table 6.2 Experiment II. Percentages of tones mislocated into or towards the processing zone, for the different types of clauses

Functionally complete 78%		Functionally incomplete 65%	
Relative clauses 77%	Headed nominalizations 78%	Noun phrases 54%	Headless nominalizations 70%

As Experiment II provides further evidence for the functional completeness property and, in particular, for the N-V-N configurational cue, it also marshals further evidence against the linguistic-structural theories of the segmentation unit. Consider Relative Clauses and Noun Phrases. In standard accounts of English grammar (e.g. Chomsky, 1957) surface noun phrases with modifiers, like (15b), are *derived* from underlying relative clause structures. Thus, Relative Clauses and Noun Phrases, like (15a) and (15b), have a rather close linguistic relationship. They are both major surface-structure constituents (noun phrases) and both correspond to deep-structure clauses. However, for Relative Clauses 77% of the location errors in Experiment II were located into or towards the processing zone, while for noun phrases this figure is 54%. This difference is significant by-subject, $p < .005$ (Wilcoxon Signed Ranks Test), and by-item, $p < .001$ (Wilcoxon Rank Sums Test).

Relative Clauses and Noun Phrases, do, however, differ linguistically. Relative Clauses are surface-structure clauses where Noun Phrases are not. In this sense, the nominalizations of Experiment II provide a critical case: headed and headless nominalizations are syntactically identical with respect to their constituent analysis. They are deep clauses, surface clauses and, in our sentences, subject noun phrases. Neither structural-linguistic theory would predict a difference. Nevertheless, headed nominalizations, as predicted by functional completeness, seem to be better segmentation units. This difference is reliable by-subject, $p < .05$ (Wilcoxon Signed Ranks Test), but misses significance by-item, $p = .102$ (Wilcoxon Rank Sums Tests, one-tailed prediction).

To summarize, the results of Experiments I and II are inconsistent with the two structural-linguistic positions characterized in our first section. Since these positions predict no difference where differences in fact obtain, they are empirically rejected in just the sense that the statistical null hypothesis is rejected. In contrast, the functional completeness property is fully consistent with the data at hand and to that extent is therefore empirically confirmed. Finally, as argued under Experiment II, the evidence appears to show in particular that the configurational N-V-N cue is an efficient variable in organizing functionally complete segmentation units.

The local sign verb tense cue

The conclusions we can draw from Experiments I and II are limited in that only one experimental paradigm was used. Thus, our results confound the intended manipulation with possible idiosyncracies of the experimental procedure. Therefore, we decided to investigate functional completeness using other experimental tasks. Experiments III and IV make use of the Suci, Ammon, and Gamlin (1967) *next-word naming task.*

Suci *et al.* presented subjects with sentences like (16) followed by probe words from the sentence.

(16) The small black ant digs holes.

They measured the time required for subjects to utter the word that immediately followed a given probe word in the original sentence. Thus, if sentence (16) was followed by the word 'black' the subject would respond 'ant', etc. Suci *et al.* found that the largest next-word naming latencies obtain when the probe word is the final word in the noun phrase ('ant' in (16)) and the response word is the initial word in the verb phrase ('digs' in (16)).

Their interpretation of this, in the terms we have been using, is that major surface constituents are segmentation units and that a probe word is a poorer prompt for a response word which has been organized into a different segmentation unit than it is for one organized into the same segmentation unit. We can extend this interpretation to amount or degree of segmentation: the more segmentation that occurs at the juncture between the probe word and the response word in the sentence, the longer it should take to retrieve the latter given the former. (For further discussion of this logic, see Carroll, 1976.)

In Experiment III we studied sentence-paradigms like (17).

(17) a. Relative Clause (Functionally Complete, Tensed Verb):
Howard who revised/ /the game-rules aggravated the old pros.
b. Sentential Subject (Functionally Complete, Tensed Verb);
That Howard revised/ /the game-rules aggravated the old pros.
c. Headed Nominalization (Functionally Complete, Nominalized Verb):
Howard's revision of the game-rules aggravated the old pros.
d. Headless Nominalizations (Functionally Incomplete, Nominalized Verb)
The revision of/ /the game-rules aggravated the old pros.

The probe word in each exemplar is 'game-rules', the response word is 'aggravated'. Each such sentence-paradigm was constructed by cross-splicing an acoustically identical copy of the sequence from the double slash through the completion of the probe-word presentation, beginning 200 msec after the end of the sentence. (The probe was read in the same voice, but with list intonation.) Each subject responded to only one member of each sentence-paradigm, but across subjects, this procedure allows us to use matched-pairs statistics by-item. (See Carroll, 1978, for further details of stimulus preparation, procedure, and scoring.)

Next-word naming latencies measure the interval from the probe onset to the beginning of the response, and are reported in Table 6.3. Analysis of variance

Table 6.3 Experiment III. Geometric mean latencies by-subject (milliseconds) for next-word naming

Mode of verb			
Tensed		Nominalized	
Relative clause (Functionally complete)	1559	Headed nominalization (Functionally complete)	1394
Sententional subject (Functionally complete)	1438	Headless nominalization (Functionally incomplete)	1350
Means	1498		1371

reveals significant effects of initial sequence-type, by-subject, $p < .05$, and by-item, $p < .05$. These overall differences militate against the unelaborated structural-linguistic segmentation theories. *All* of the four sentence categories in (17) have similar linguistic descriptions. Their underlined initial sequences correspond to deep-structure clauses, surface-structure clauses and noun-phrase (major surface) constituents (see Chomsky, 1976; Stockwell *et al.* 1973). Thus, once again the structural-linguistic theories predict the null hypothesis, and are rejected by the data.

What is of greater interest to us, however, is how these overall differences bear on the validity of the segmentation cues we have proposed. In particular, we would like to test the validity of our hypothesized local sign segmentation cues. We were unable to find effects of local sign cues in the data of Experiments I and II. For the data of Experiment III we tested the prediction that tense on a verb stem acts as a perceptual cue identifying a potential functionally complete segmentation unit.

Two of the sequence-types studied in Experiment III contained tensed verbs (relative clauses and sentential subjects) and two contained nominalized verbs (the two nominalization types). The difference between tensed cases and nominalized cases is significant by-subject, $p < .005$, and by-item, $p < .05$. However, this overall contrast partially confounds the local sign tense cue with the configurational N-V-N cue: Both tensed sequence sequence-types comprise N-V-N configurations, however, only one of the nominalized sequence-types comprises N-V-N configurations (headed nominalizations are N-V-N, but headless nominalizations are not). Therefore, we contrasted the two tensed sequence-types with headed nominalization cases alone. This difference is significant by-subject, $p < .025$, and nearly significant by-item, $p = .077$.

We also examined the data of Experiment III for effects of the N-V-N cue. First, we contrasted the three sequence-types which comprise N-V-N configurations (relative clauses, sentential subjects, headed nominalizations) with the

fourth which does not (headless nominalization). This difference is reliable by-subject, $p < .025$, and by-item, $p < .025$. However, this overall comparison confounds N-V-N with various local sign differences (e.g. relative clauses have relative pronouns 'who' and tense 'revised', sentential subjects have the complementizer 'that' and tense). Contrasting headless nominalization cases only with headed nominalizations minimizes these spurious differences. When we assessed the effects of the N-V-N cue by contrasting the two nominalization types, we found only a non-significant trend, ($p = .19$, by-subject; $p = .13$, by-item).

Post hoc we also examined the data of Experiment III for effects of sequence length. We correlated the length in words of the initial sequence for each of our experimental sentences within a sequence-type with the geometric mean next-word naming latency across subjects for each sentence (the overall range of sequence length was 4 to 11 words, $\bar{X} = 6.8$). Longer initial sequences should encourage segmentation to a greater extent than shorter initial sequences, since the former consume more processing capacity. For Relative Clauses, the correlation between sequence length and response latency is $r = .51$, $p < .01$. For Sentential Subjects, this correlation is $r = .56$, $p < .005$; for Headed Nominalizations, $r = .35$, $p < .05$; and for Headless Nominalizations, $r = .26$, n.s. These data suggest that the processing capacity cue of sequence length *does* influence segmentation organization. We elaborate this result in Experiment IV.

Experiment III provides more evidence against the structural-linguistic theories. Further, it provides some initial confirmatory evidence for our functional completeness cues of verb tense and N-V-N, as well as for the processing capacity cue of sequence length. The evidence concerning verb tense and N-V-N is somewhat paradoxical: recall that in Experiment II we were unable to find an effect of verb tense, when N-V-N configurational properties were controlled. Conversely in Experiment III the significant effect of N-V-N we found overall, reduces to a non-significant trend when verb morpheme (and other local sign) properties are controlled.

The difficulty we have had in separating configurational and local sign cues is in itself not surprising. These cues are ecologically very highly correlated. The ideal factorialized manipulations one would like to make, do not exist in English. But what is surprising, or at least intriguing, is the seeming asymmetry between the local sign cue and the configurational cue vis-à-vis the two response measures we used in Experiments II and III. We will return to this in our conclusion.

Sequence length as a processing capacity cue

Experiment III provided highly significant but indirect evidence of sequence length as a segmentation cue. Experiment IV was designed to test directly for an

effect of the processing capacity cue of sequence length. We used sentence-pairs like the one in (18).

(18) a. Long:
The dark and dangerously empty//city park seemed to get more frightening by the minute.
b. Short:
The//city park seemed to get more frightening by the minute.

'Long' members of a pair had an initial noun-phrase sequence seven words long; 'Short' members had noun phrases three words long. The probe word was the final word of the noun phrase (e.g. 'park') and the response word was the verb (e.g. 'seemed'). Sentence-pairs were cross-spliced so that they were acoustically identical from before the original occurrence of the probe word in the sentence through its presentation beginning 200 msec after the sentence (in (18), from the double slash on).

We predicted that the 'long' cases would be better segmentation units than the 'shorts' and hence obtain longer naming latencies. However, this prediction is confounded with a non-segmentation 'sequence search' model: there are more words to search through in the long cases in order to find the probe word–response word sequence. We therefore included control sentences as in (19).

(19) a. *The very unusual but impractically shaped clay* dish sold for 75 dollars.
b. *The rowdy school* boys were punished for disrupting the assembly.

The italicized sequences are 'potential' noun phrases, but cannot be interpreted as such in the sentences of (19). We probed the same serial positions as in the experimental sentences, the seventh word (clay) or the third word (school).

In the control sentences, unlike the experimental sentences, a segmentation boundary does not intervene between probe word and response word. Thus, on the segmentation account of length differences, but not on the sequential search model, we predict, first, a long versus short difference for experimental but not control sentences and, second, an overall effect of longer latencies for experimental than for control due to segmentation effects in the former but not the latter. (For further details of design and procedure, see Carroll, 1976.)

Both of these predictions obtain; refer to Table 6.4 for summary statistics. Analysis of variance reveals an overall effect of condition (experimental versus control, $p < .05$. T-tests show a significant effect of sequence length by-subject, $p < .05$, and by-item, $p < .025$, for experimental sentences. However, no sequence-length effect obtains for control sentences. Thus, using the next-word technique we measure an effect of length but not 'mere length'. Length cannot

Table 6.4 Experiment IV. Geometric mean latencies by-subject (milliseconds) for next-word naming

		Sentences	
		Experimental	Control
Length of sequences	Long	1309	1074
	Short	1148	995
	Means	1225	1034

stimulate segmentation when that segmentation would constitute a misparsing (e.g. (the very unusual but impractically shaped clay) + (dish sold for 75 dollars) = (19a)).

Context as a segmentation cue

Processing capacity cues interact with cues to internal structure to define the potential segmentation units of the perceived sentence. 'Long' and 'short' experimental sentences both have similar internal structures (in the sense of functional completeness), but the 'long' cases consume more processing resource and therefore encourage segmentation more. On the other hand, while 'long' experimental and control exemplars are comparable in terms of the sequence length estimation of processing requirements, they differ in terms of internal structure and the experimental cases stimulate segmentation more than the control cases.

We now turn to the question of whether and, if so, how contextual cues help the listener to recognize and recode functionally complete sequences. In two experiments, we preceded sentences like (20a), which has functionally incomplete a first clause, with specifying contexts that provided the missing grammatical relations, (20b), or with neutral contexts that did not do so, (20c).

(20) a. After buying some he returned home.
b. The boy went to the store to buy some gum.
c. The store was around the corner.

If contextual cues can help the listener to recognize and recode good potential segmentation sequences, the functionally incomplete sequences preceded by specifying contexts should be better segmentation units than those preceded by neutral contexts. Only in the former case can the functionally incomplete sequence be 'completed'.

The first experiment used the next-word naming technique. The probe was the last word of the first clause and the correct response word was the first word

of the second clause. No differences in probe latencies were found as a function of the two types of contexts used. The second experiment used the location technique. Tones were placed either before the last word of the first clause or after the first word of the second clause (see Experiment I). Context did not affect the pattern of tone locations.

These non-results led us to speculate that the initial recoding schemes of sentence processing operate independently of context. However, a disturbing aspect of these experiments is that both the location and next-word naming tasks require that the listener maintains aspects of the literal form of the test sentence in memory. It is possible that this actually discourages subjects from integrating the test sentence with the preceding context sentence. *Prima facie*, this does not seem to be what happens in ordinary sentence comprehension. As a result, we have recently conducted several additional experiments to determine whether context can influence the early stages of sentence perception and comprehension.

In Experiment V, subjects were presented with two-sentence discourses, like those in (21).

(21) a. Why did Smith resign?
b. Smith who Ford asked to remain in Vermont, for political reasons, decided to resign his post.
c. Why did Ford want Smith to remain in Vermont?
d. Smith who Ford asked to remain in Vermont, for political reasons, decided to resign his post.

The context question, (21a) and (21c), was typed on an index card. The subject first read the question aloud. A tape-recorded version of the second sentence, (21b) and (21d), respectively, was then played to the subject. The second sentence was always ambiguous in that it contained an adverbial 'swing phrase' that could modify either the main clause or the embedded clause. In (21), the swing phrase is 'for political reasons'. The sentences were recorded using subdued intonation, with a brief pause, before and after the swing phrase. The same recording was used for both context question conditions.

After reading the context question and hearing the ambiguous sentence the subject first paraphrased the ambiguous sentence, then answered the context question. The question directed the subject's attention either to the main clause of the second sentence, as in (21a), or to the embedded clause of the second sentence, as in (21b). We used the subjects' paraphrase of the ambiguous second sentence to determine which organization was assigned to it. Subjects' paraphrases were overwhelmingly consistent with the assignment called for by the question-context. In addition, only one subject even noticed that some of the sentences were ambiguous. Thus, in a task for which discourse context must be held in memory (question-answering), apparent context effects on segmentation organization *do* obtain.

Experiment VI provides evidence indicating that contextual cues are actually used on-line during sentence processing. We have already discussed the N-V-N configurational cue, and the evidence for it in Experiments I, II, and III. Consider, from this perspective, sentences (22).

(22) a. *The girl who kissed the boy* blushed.
b. *The girl whom the boy kissed* blushed.

The initial italic sequence in (22a) is an N-V-N functionally complete unit, while the corresponding sequence in (22b) is functionally complete but is not an N-V-N configuration. On this basis, we would predict that sentence (22a) would be easier to organize perceptually than sentence (22b). This prediction accords with a variety of evidence showing that subject relative-clause sentences, like (22a), are less difficult to process than object relative-clause sentences, like (22b) (see Wanner, 1976).

In a pilot study, subjects were presented with a series of sentences including subject and object relative clauses, like (22a) and (22b), respectively. Their task was to press a telegraph key as soon as they understood each sentence. Reaction times to understand subject-relative sentences were faster than reaction times to understand object-relative sentences.

Two types of context were then constructed for each pair of sentences containing subject and object relative clauses. The *congruent* context condition presented the information contained in the relative clause of the following sentence. For example, the congruent contexts for (22a) and (22b), were (23a) and (23b), respectively.

(23) a. One of the girls kissed the boy.
b. A boy kissed one of the girls.
c. The kids were playing spin the bottle.

The *neutral* context (23c), which did not specify anything about the information in the relative clause, was the same for both subject- and object-relative sentences. The two types of context and the two types of relative clauses were then varied factorially, resulting in four presentation versions. The versions were then recorded with normal intonation. Each subject heard only one version. The subject's task was to press a key as soon as he understood the second sentence of each pair. (For details of method, procedure, and results, see Tanenhaus, 1978.)

If the initial stages of sentence processing are independent of context, then the complexity difference between the subject and object relative clauses should not be affected by specifying versus neutral contexts. The results, which are presented in Figure 6.1, clearly do not support this prediction. Anova reveals a

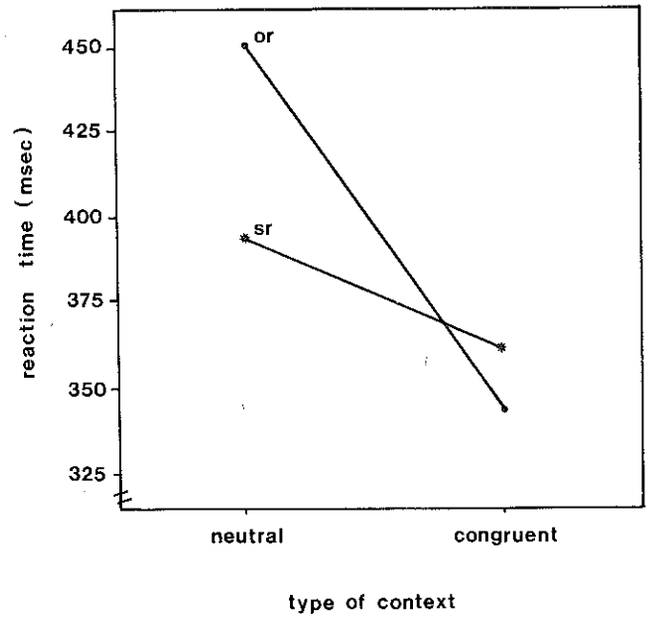


Figure 6.1. Reaction times to understand sentences containing subject (SR) and object (OR) relative clauses when preceded by neutral and congruent context sentences

main effect of context, $p < .001$ by-subject and $p < .025$ by-item, and an interaction between context and sentence type, $p < .001$ by-subject and $p < .025$ by-item. In the neutral context condition, reaction times to understand object-relative-clause sentences were significantly longer than reaction times to understand subject-relative clauses. This presumably reflects the difference between the two sentence types measured without any context manipulation. However, in the congruent context condition this object-relative versus subject-relative difference was completely eliminated. This pattern of results clearly suggests that intra-sentence syntactic-based processing strategies can be influenced by context.

Experiments V and VI suggest, contrary to our initial hypothesis, that listeners use context on-line as they are processing and organizing sentences. These results suggest that functionally incomplete sequences should become better processing units when they are preceded by contexts that clearly provide the information missing in the incomplete sequence. We are currently investigating this hypothesis, using paradigms that do not require the subject to hold the test sentences in memory.

CONCLUSIONS: THE PERCEPTION OF RELATIONS

Our conclusions divide into three sections. First, we summarize the specific points which have been discussed in the present chapter. Next, we consider some broader implications of this work for the study of sentence perception. We speculate on the relation of our cue constructs to levels of processing and representation in sentence perception. Finally, we attempt to place this research in the general context of the study of perception and comprehension.

Summary and specific conclusions

Early in this discussion, we raised an apparent paradox: human behaviour and experience are 'complex', while human processing resources are 'limited'. We reviewed the classic 'chunking' proposal of Miller (1956) concerning the deployment of limited resources. Miller and Smith's work shows that a saving in processing resource can be realized if stimuli are hierarchically and iteratively grouped, or 'chunked'. We then posed a further question: What are the principles that underlie chunking? In Smith's digit-recall experiments, recoding schemes were given by fiat of experimenter. Subjects were instructed as to how they should organize sequences. This work points out the potential advantages of chunking as a construct accounting for the manifest complexity of behaviour and experience. However, it suffers two limitations: first, digit lists are not intrinsically structured, in contrast to more natural stimuli; second, in ordinary behavioural situations people are usually not explicitly given recoding schemes to organize their processing capacity. General principles of chunking based exclusively on this work might overlook the implicit contribution of intrinsic stimulus structure to the organization of behaviour and experience.

Our own work has focused on a sequential stimulus which has a well-studied intrinsic structure. We have investigated how, if at all, the intrinsic structure of the stimulus determines the nature of the recoding schemes utilized in organizing the stimulus (e.g. perceptually). In particular, we have been concerned with the relation of linguistic structure and the recoding schemes that organize sentence perception. We reviewed two previous proposals to the effect that these recoding schemes pick out particular linguistic structures (clauses or major surface constituents). These proposals describe a simple and straightforward relation between stimulus structure and the recoding schemes that organize sentence perception: the latter must directly reconstruct the former (for further discussion, see Carroll, 1976). However, we were compelled to abandon this approach on both intuitive and experimental grounds. In the second section of this paper, we developed the claim that the recoding schemes of sentence perception are determined by an interaction of factors including, but not limited to, structural details of the stimulus.

We adopted a 'functionalist' perspective, and reapproached sentence perception from the viewpoint of the perceiving organism. We made the assumption that the organism attempts to organize perceptual input into optimally propositional units. However, we allowed, on the basis of the examples discussed in the first section, that these units might not always correspond to the propositions of the sentence's logical form. (Although, when comprehension succeeds, the 'logical' propositional structure of a sentence will always be recoverable from its 'perceived' propositional structure.) In the second section, we considered intuitively how listeners might recognize and assemble propositional (or grammatical) relations. In the third section, we experimentally tested some of these initial hypotheses.

The central property we have considered is functional completeness: the property of having a complete, coherent, and explicit set of grammatical relations. Functional completeness guarantees the listener a good propositional unit. But how can functional completeness be recognized? Can the listener always organize a sentence into functionally complete units? And, if not, how does the listener's processing compromise functional completeness when it must?

In approaching these questions we introduce a taxonomy of perceptual cues. We opposed cues to internal structure and processing-capacity cues. The former assist the listener in identifying good perceptual units (by hypothesis, functionally complete sequences). The latter restrict the processing of the former, and thereby can cause the listener to compromise functional completeness. Cues to internal structure were further divided into local sign, configurational and contextual/inferential cues.

In the third section, we presented six initial experimental studies designed to show that the sorts of cues we hypothesized could indeed influence the perceptual organization of sentences. In particular, we wanted to collect evidence indicating that local sign, configurational, and contextual/inferential cues influence segmentation and recoding to the extent that they inform the listener about the functional completeness of potential segmentation units. Additionally, we wanted to show that processing capacity cues influence perceptual organization in interaction with cues to internal structure.

Experiments I and II used the tone-location response measure. They showed that functionally complete sequences are better perceptual units than functionally incomplete sequences. Experiment I also manipulated sequence length, a presumed processing-capacity cue. However, no effect of sequence length obtained. Experiment II manipulated particular cues to internal structure in order to determine more specifically what accounted for the overall functional-completeness result. Contextual/inferential cues cannot provide the basis for the effect, since if contextual/inferential information were structuring the perceptual organization of the sequences we studied in Experiments I and II, *all* of them would effectively be functionally complete.

On our taxonomy, that leaves local sign and configurational cues. In Experiment II, we tried to separate these two cue classes and determined that

the N-V-N configurational cue could account for our data. The additional assumption of local sign cues was not supported by the results of Experiment II.

Experiments III and IV used the next-word naming response measure. Experiment III replicated the main functional completeness result of Experiments I and II. Again, we tried to determine which functional completeness cue or cues accounted for the obtained effect. Contextual/inferential cues were not manipulated, but we were able to contrast the N-V-N configurational cue with the local sign cue of verb tense. The data suggest that the configurational cue has a somewhat marginal influence, and that the local sign cue provides a much better account of obtained differences.

Experiment III also revealed significant correlational effects of the processing-capacity cue of sequence length. Experiment IV directly manipulated sequence length, and replicated this finding. Additionally, Experiment IV showed that 'mere' length does not influence the suitability of sequences as recoding units, but that length in conjunction with internal structural properties does.

Experiments V and VI addressed the issue of contextual/inferential cues. Experiment V demonstrated that contextual information can, at least in part, determine which of two possible organizations a listener imposes on an ambiguous sentence. Experiment VI shows that context can render a sequence which is configurationally a relatively poor segmentation sequence (i.e. not N-V-N), as good a segmentation sequence as one which is configurationally N-V-N.

As should be clear, we find the previous structural-linguistic theories of sentence segmentation inadequate. Neither of these theories distinguishes functionally complete sequences from functionally incomplete sequences. Neither of these theories distinguishes sequences of differing lengths. Finally, neither of these theories characterized discourse context as a segmentation variable. As we argued in the first section, the logical basis of these theories is not sound. But more concretely, the data from the experiments presented in the third section reject these theories as null hypotheses. It seems beyond question that a new approach to sentence segmentation is required.

Speculation on segmentation cues and processing capacity

While the preceding statement of results and conclusions is as far as, and perhaps farther than, we can go at present, some of the apparent 'weaknesses' in the data we have presented suggest a possible further elaboration of our cue theory of sentence perception, a 'strengthening' of our claims. First, consider the intuitive trade-off of cue 'cost' and cue 'validity' among the three classes of cues to internal structure that we presented.

Contextual-inferential cues seem to be highly valid indicators of functional completeness. They pick out any sequence that *can* be represented by a well-formed proposition (i.e. subject-predicate-object). However, the cost of contextual-inferential cues seems to be quite high too. In order to recognize a headless nominalization as a potential functionally complete unit, for example, it is necessary to recall previous discourse context, or to inspect the following clause for a logically antecedent subject, etc.

In contrast, consider our class of configurational cues. At this level of analysis, there is a sharp distinction between sequences like headless nominalizations which are themselves functionally incomplete (even though context may complete them) and sequences which are themselves functionally complete, like headed nominalizations. Only the latter comprise a complete, coherent, and explicit *configuration* of grammatical relations.

In the sense that configurational cues pick out a smaller subset of potential functionally complete sequences, they are less valid cues of functional completeness than contextual/inferential cues, which pick out a larger set. On the other hand, configurational cues seem to have less cost, as well. In order to compute, recognize, or apprehend a configurational cue, the listener need not concern himself with preceding information, subsequent information, or concurrent extra-linguistic information.

In this regard, consider local sign cues. In this case, the listener need not even consider the sequence, in its entirety, in order to obtain some information about the suitability of the sequence as a recoding unit. The listener can limit the scope of processing to single lexical items, or even single morphemes. Thus, local sign cues seem to have even less cost than configuration cues. Of course, they also have less cue validity. Verb tense, for example, is a local sign cue of functional completeness, but fails to pick out headed nominalizations. Local sign cues pick out a smaller subset of potential functionally complete sequences than configurational cues. (See Brunswik, 1956, and Carroll, 1976, for further discussion of cue cost and cue validity trade-offs.)

The various experimental paradigms that we used can also be intuitively ordered in terms of their 'cost' in processing capacity to the listener. We used four tasks: tone location, next-word naming, paraphrasing, and comprehension time. Of these, the next-word naming task of Experiments III and IV seems to place the greatest load on the listener's processing capacities. The task requires, first, that the subject must maintain a relatively literal representation of the sentence, and second, that the subject must rapidly search this representation for two particular adjacent words.

The location task of Experiments I and II also requires that the subject must maintain the literal word sequence of the sentence. It further requires that the subject must remember the position of the tone in the sentence sequence. However, the subject is not under any time pressure to generate his recall and tone-position responses.

Neither the paraphrase task of Experiment V nor the comprehension task of Experiment VI requires the subject to retain any verbatim representation of the

sentence. (However, in the latter task but not in the former, the subject's response is timed.)

Organizing our cue taxonomy and our experimental paradigms into 'cost' hierarchies suggests an account of the asymmetries in the data we have presented. Cue cost and task demands vis-à-vis processing capacity may trade off: higher cost cues can be used in lower cost experimental task environments. Conversely, lower cost cues will be relied on to a greater extent in higher cost task environments. Higher cost cues are preferable because they are more highly valid indicators of functional completeness. However, in relatively high-cost task environments, their cost in processing capacity may be prohibitive.

With this hypothesis in mind, reconsider the evidence from Experiments I to VI. Experiments V and VI made use of tasks relatively low in cost vis-à-vis processing capacity. These experiments provide evidence of contextual cues—relatively high in cost vis-à-vis processing capacity. Experiment VI, in particular, indicates that in the comprehension time task, contextual cues are more prepotent than configurational cues.

We have analysed the location task as being higher in cost than the paraphrase and comprehension tasks. In this task environment, we measured effects of the N-V-N configurational cue, but failed to find effects of the local sign verb tense cue or of context. This is consistent with the view that the listener does not have the processing capacity to make use of context. Being able, then, to use configurational and local sign information, he relies on the more highly valid configurational indicators of functional completeness.

Finally, consider the next-word naming task. In this relatively high-cost task environment, the listener must rely on the lowest-cost local sign type of cue. A trace effect of the N-V-N configurational cue is still found, but effects of context are not.

Somewhat more direct support for this analysis comes from our study of the processing-capacity cue of sequence length in Experiments I, III, and IV. In a relatively lower cost task environment (location), we find no effect of sequence length. However, in a higher cost task environment (next-word naming), we do find length effects.

Our 'cost' hierarchy hypothesis requires some conceptual elaboration. First, the notion of 'cost', while it has very much intuitive appeal, must be spelled out in far more detail. This necessitates an analysis of the micro-structure of sentence segmentation, and of the psycholinguistic research paradigms used to study segmentation, that has really only begun. Why and how, precisely, are configurational cues higher in cost than local sign cues? Why and how, precisely, is one task higher in cost than another? Would, for example, tone location, with some time pressure introduced, be higher in cost than next-word naming, with accuracy weighted far greater than speed? Without such an elaboration the cost-hierarchy hypothesis rests on very thin ice.

Second, we need to specify just what the cue hierarchy commits us to. We feel that while the effectiveness of local sign cues in segmentation processing may typically *precede* that of configurational cues, the former may not necessarily construct the input to the latter. Another way of putting this is to say that we doubt that the hierarchy of cues reflects a serially structured segmentation process. Again, we can only speculate on this matter in the absence of an analysis of the micro-processes of segmentation and recoding.

Finally, we need to investigate what relation exists between our task hierarchy and ordinary levels of processing in sentence perception. For instance, it is tempting to suggest that higher cost task environments (which tend to require the subject to maintain a nearly verbatim representation and to generate a response under time pressure), are actually tapping representations and processes of relatively early levels of sentence processing. Lower cost tasks might actually be more sensitive to later levels of perceptual processing. On the other hand, tasks may artifactually deform 'ordinary' perceptual representations to the extent that there would be no interesting correspondence with our notion of task cost and isolable levels of processing. But again, these are matters for further research.

Recoding schemes and relations: General remarks

We have several times recalled Miller's chunking proposal and raised the question of whether there are any general properties of recoding schemes, and if so, of course, what these properties are. On the basis of our investigations thus far, we are encouraged to believe that there are general properties. Specifically, we propose that all recoding schemes have the common property that they construct coherent relational mental structures. In the context of the study of sentence segmentation, our claim is that recoding schemes construct linguistic propositions consisting of coherent sets of grammatical relations.

More follows from this, however. We have already noted that the propositions constructed by perceptual recoding schemes may not correspond isomorphically to the underlying syntactic propositions of the logical form of the sentence. Thus, we must distinguish between the structural relations intrinsic to the sentence and the functional relations which are jointly determined by the intrinsic structure of the sentence and the exigencies, heuristics, limitations, and contexts of perception. At any given stage or level of processing, in any given task-environment and in any discourse or extra-linguistic context we can only measure the consequences of functional relations and the propositional units they define. This does not alter the fact that functional propositions are entrained by logical structures, that functional relations are entrained by structural relations. From the former we must be able to recover the latter. If a perception/comprehension device cannot satisfy this requirement it is not general.

But recoding schemes must do even more than construct functional relations and map them into structural relations. Once a functionally complete unit has been isolated and recoded, it must be integrated with other similar units to recover the 'perceived sentence'. Similarly, the perceived structure of a sentence must be integrated with other similar units to recover the 'perceived discourse'.

Recoding schemes map from stimulus sequences onto relational structures, and from relational structures onto other relational structures. At relatively 'immediate' stages of perceptual organization these ideals or goals may be most severely compromised by limitations of processing capacity. But even at relatively later stages, functional limitations of the task-environment, motivations, other context, etc., compromise these mappings. In every case, we would claim, the output structures are 'better' mental units: more complete relational structures, more complete amalgamations of relational structures and relational structures more consistent with the intrinsic structure of the stimulus.

The fundamental question in the description of comprehension processes is how the defining relations of the stimulus are internally organized. The propositions of a sentence, the relations between them and the grammatical relations internal to them, *inhere* in the sentence (or, perhaps more precisely, in our knowledge of the sentence). A theory of sentence comprehension must describe a process that can recover and properly organize these relations.

This is, perhaps, not a unique property of the study of sentence comprehension. The objects of a visual display, the spatial relations between them (above, behind, within) and the relations of the contours, areas, and volumes (etc.) internal to them, *inhere* in the display. Comprehending a visual display entails recovering and properly organizing these defining relations. In both cases, however, various limitations of human processing capacity complicate comprehension. For example, an obvious limitation in visual information processing is the relatively small spatial extent of what can be foveally fixated. Resultingly, the comprehension of visual displays is intermediately structured by a sequence of foveal glances from which the relations underlying the entire display must be recovered (Hochberg, 1968).

From this perspective, the units or 'objects' of perception and comprehension are coherent relational structures. In order to study perception and comprehension, we must identify the relations intrinsic to the structure of the stimulus and the mechanisms and heuristics employed in recovering and organizing these defining relations. In the case of sentence perception, the science of linguistics provides a theory of the fundamental relations which define the stimulus. Similarly, the intrinsic structure of visual objects and events has been described (Carroll, 1977; Clowes, 1971; Winston, 1970). In both cases, the most formidable question is *how* these relational structures are recognized and integrated into percept and concept.

ACKNOWLEDGMENTS

This research was supported in part by NSF-BNS 76-04334. An earlier version was presented at the 1976 Annual Meetings of the Eastern Psychological Association (Carroll and Tanenhaus, 1976). We thank Ellen Perecman and David Tucker who helped with stimulus preparation and in administering the experiments.

REFERENCES

Abrams, K. Real-time sentence processing. Unpublished Ph.D Dissertation. University of Michigan, 1973.

Abrams, K. and Bever, T. G. Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology*, 1969, **21**, 280–290.

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

Bever, T. G. and Hurtig, R. Detectability of a non-linguistic stimulus is poorest at the end of a clause. *Journal of Psycholinguistic Research*, 1975, **4**, 1–7.

Bever, T. G., Hurtig, R. R., and Handel, A. B. Response bias does not account for the effect of clause structure on the perception of non-linguistic stimuli. In preparation (available as Research Bulletin, Educational Testing Service, RB-75-30, 1975).

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.

Brunswick, E. *Perception and the representative design of psychological experiments*. Berkeley: University of California Press, 1956.

Carroll, J. M. The interaction of structural and functional variables in sentence perception: Some preliminary studies. Unpublished Ph.D. Dissertation, Columbia University, 1976.

Carroll, J. M. A program for cinema theory. *The Journal of Aesthetics and Art Criticism*, 1977, **35**, 337–351.

Carroll, J. Sentence perception units and levels of syntactic structure. *Perception and Psychophysics*, 1978, **23**, 506–514.

Carroll, J. M. and Bever, T. G. Sentence comprehension: A study in the relation of knowledge to perception. In E. C. Carterette and M. P. Friedman (Eds.), *The handbook of perception*. Vol. 5, *Language and speech*. New York: Academic Press, 1976.

Carroll, J. M. and Tanenhaus, M. K. Function clauses and sentence segmentation. *Journal of Speech and Hearing Research*, 1978, in press.

Carroll, J. M. and Tanenhaus, M. K. Functional completeness and sentence segmentation. Annual Meetings, Eastern Psychological Association, 1976.

Chapin, P., Smith, T., and Abrahamson, A. Two factors in perceptual segmentation of speech. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 164–173.

Chomsky, N. *Syntactic structures*. The Hague: Mouton, 1957.

Chomsky, N. *Reflections on language*. New York: Pantheon, 1976.

Clowes, M. On seeing things. *Artificial Intelligence*, 1971, **2**, 79–98.

Ferguson, G. *Statistical analysis in psychology and education*. New York: McGraw-Hill, 1971.

Fodor, J. A. *Psychological explanation*. New York: Random House, 1968.
Fodor, J. A. and Bever, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 414–420.
Fodor, J. A., Bever, T. G., and Garrett, M. *The psychology of language. An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill, 1974.
Hochberg, J. In the mind's eye. In R. Haber (Ed.), *Contemporary theory and research in visual perception*. New York: Holt, Rinehart and Winston, 1968.
Kintsch, W. *The representation of meaning of memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
Köhler, W. *Gestalt Psychology*. New York: Liveright, 1929.
Marslen-Wilson, W. D., Tyler, L. K., and Seidenberg, M. The semantic control of sentence segmentation. Paper presented at the Annual Meetings, Midwestern Psychological Association, 1976.
Miller, G. A. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81–96.
Rumelhart, D. E., Lindsay, P. H., and Norman, D. A. A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
Seitz, M. AER and the perception of speech. Unpublished Ph.D. Dissertation, University of Washington, 1972.
Stockwell, R., Schachter, P., and Partee, B. *The major syntactic structures of English*. New York: Holt, Rinehart and Winston, 1973.
Suci, G., Ammon, P., and Gamlin, P. The validity of the probe-latency technique for assessing structure in language. *Language and Speech*, 1967, **10**, 69–80.
Tanenhaus, M. K. Linguistic context and sentence perception. Ph.D. Dissertation, Columbia University, 1978.
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.
Tanenhaus, M. K. and Seidenberg, M. Informational determinants of the processing units in sentence perception. Annual Meetings, Midwestern Psychological Association, 1978.
Wanner, E. Paper presented at the Symposium on the Perception of Language. XXI International Congress of Psychology, Paris, July 1976.
Winston, P. Learning structural descriptions from examples. Unpublished Ph.D. Dissertation, M.I.T. Cambridge, Mass., 1970.

Chapter 7

Sentence Processing and the Clause Boundary

William Marslen-Wilson, Lorraine K. Tyler, and Mark Seidenberg

The identifying characteristic of most recent theorizing about sentence perception has been its emphasis on the role of the syntactic clause. This approach, known in general as the 'clausal hypothesis' (Carroll and Bever, 1976; Fodor, Bever, and Garrett, 1974), can in principle be separated into two main components. The first of these is the hypothesis that the speech input is segmented into relatively large processing units, which correspond to the syntactic clausal structure of the material (cf. Fodor and Bever, 1965; Garrett, Bever, and Fodor, 1966). We will refer to this as the clausal *structuring* hypothesis. The second main component is a theory about the order of processing events underlying the clausal organization of the input. This is an essentially serial theory, which argues that speech is not understood as it is heard, but rather that the major interpretative efforts of the processing system are focused around the syntactic clause boundary. We will label this the clausal-*processing* hypothesis.

The purpose of the arguments and the experiments to be presented in this chapter is to investigate the consequences for both these clausal hypotheses of a rather different approach to sentence perception, which we will refer to as the *on-line interactive* approach. This approach claims, in essence, that the listener attempts to interpret fully each utterance word-by-word as he hears it, from the first word of the input onwards (cf. Marslen-Wilson, 1976; Marslen-Wilson and Tyler, 1975). This claim, and the evidence supporting it, stand in obvious contrast to the clausal processing hypothesis. One of the experiments to be described here will include an additional test of the two approaches, focusing on the point in the processing of a sentence at which their predictions most clearly diverge: namely, as to whether or not there is an asymmetry in the availability of a semantic interpretation of the incoming material at the end of a clause as opposed to the beginning of a clause.

Secondly, we will examine the relationship between the on-line interactive