LANGUAGE AND COGNITIVE PROCESSES, 1991, *6* (1) 49–77

The Use of Higher-level Constraints in Monitoring for a Change in Speaker Demonstrates Functionally Distinct Levels of Representation in Discourse Comprehension

David J. Townsend
Department of Psychology, Montclair State College, Upper Montclair, New Jersey, U.S.A.

Thomas G. Bever
Department of Psychology, University of Rochester, Rochester, New York, U.S.A.

We test and disprove the common assumption that pragmatic probability facilitates the processing of lower linguistic levels. Two experiments show that detection of the acoustic properties that distinguish two speakers is harder in sentences that are pragmatically more probable. At the same time, detection of the same acoustic properties is easier at later points in a clause than at earlier points. The pragmatic inhibition in detecting acoustic properties suggests that different levels of linguistic processing compete for processing resources, while the within-sentence facilitation suggests limited interactions between processes at adjacent levels. The two results together demonstrate that discourse-level and sentence-level representations are

Requests for reprints should be addressed to David J. Townsend, Department of Psychology, Montclair State College, Upper Montclair, New Jersey 07043, U.S.A.

This research was supported by Grant BNS-8120463 from the National Science Foundation and by Separately Budgeted Research grants from Montclair State College to D.J.T. The preparation of the manuscript was made possible by fellowships from the Princeton Faculty Fellowship Program and the Institute for Critical Thinking at Montclair State College.

The authors are grateful to Caroline Carrithers for recording the stimuli, Rick Tiesen for preparing tapes, Brian McElree, Steve Dopkins, and Beth DeForest for running experiments, Suzanne Kannick and John Pesaniello for assisting in data analysis, George Rotter for statistical discussions, Carl Bredlau and the members of the Cognitive Science Seminar at Montclair State College for valuable comments, and Wayne Cowart and two anonymous reviewers for their detailed analysis of earlier drafts of the manuscript. Reports of this research were presented at the annual meetings of the Psychonomics Society and the City University of New York Conference on Human Sentence Processing.

functionally distinct during comprehension, and they force a modification of the view that linguistic processes at different levels of representation occur in architecturally segregated "modules" that do not share resources. We argue that two levels of representation can interact during comprehension when the representations that are available at the two levels are compatible. Thus, the occasional informational independence of processes occurs for functional reasons, not because they are carried out by architecturally segregated processes.

INTRODUCTION

How do listeners co-ordinate the processing of acoustic, lexical, syntactic, semantic, and pragmatic information during speech understanding? Recent discussions have focused on how and when more abstract kinds of information – such as the general topic of discourse, the speaker's intentions, the contextually appropriate meaning of a sentence, and the literal meaning of a sentence – are used in processing for more superficial information, such as the acoustic and lexical form of speech. An intuitively clear approach proposes that such abstract information never informs concrete levels of processing. According to this view, speech processing proceeds roughly from acoustic to lexical, to syntactic, to semantic, and then to pragmatic levels of representation, as in:

1. acoustic → lexical → syntactic → semantic → pragmatic

where a unidirectional arrow indicates that processing at a level of representation to the right of the arrow receives information only from processes to the immediate left of the arrow. Yet, basic facts about speech perception have long demonstrated that a simple "inward" progression of processing is not a correct account of speech perception. For example, Miller, Heise, and Lichten (1951; Miller & Isard, 1963) have shown that the perceptibility of words in noise is greatly enhanced when the words appear in a well-formed sentence. More recently, Marslen-Wilson and Tyler (1980) have shown that listeners' reaction times for detecting a specified word in speech decrease as the coherence of preceding context increases. These studies conclusively proved that some aspects of speech processing involve the use of information from more abstract levels for the organisation of more concrete levels.

While the general conclusions from studies like Miller et al. (1951) and Marslen-Wilson and Tyler (1980) are clear, there are several ways to interpret these results for a detailed model of speech perception. One interpretation is that all levels of representation interact at every point during perception. For example, information at a semantic level could constrain ongoing processing at the acoustic and lexical levels. This view-point makes explicit the intuition that when listeners know the general context, perception is facilitated. We shall call this intuition the *predictability assumption*: If information at a more abstract level is specified, it can reduce the number of choices that are available for consideration at a less abstract level, thereby facilitating decisions at the less abstract level. For example, semantic information can eliminate choices at the acoustic level, making it easier to perceive acoustic features. The predictability assumption is implicit in several models of linguistic processing that propose that higher levels of information can activate lower-level features. For example, the activation of a word candidate can activate the hypothesis that each of its component phonemes is present, and it can de-activate hypotheses about the presence of phonemes that are not a part of the word (e.g. Elman & McClelland, 1984). Although it is not essential that connectionist models of speech and language allow higher levels to influence all possible lower levels, they frequently do allow activation to "cascade" downwards, so that, for example, word hypotheses activate component phoneme hypotheses, which in turn activate hypotheses about acoustic features (Elman & McClelland, 1984; see also McClelland & Rumelhart, 1981, pp. 378–379). Extending this line of reasoning, speech comprehension may involve information flow as follows:

2. acoustic ↔ lexical ↔ syntactic ↔ semantic ↔ pragmatic

where a bidirectional arrow indicates that hypotheses at a particular level may be activated from both lower and higher levels of information. The predictability assumption also appears in models that maintain that semantic and pragmatic information can in effect prime syntactic options (Tyler & Marslen-Wilson, 1977; see also Schank & Birnbaum, 1984) because "there is no level of symbolic representation mediating between lexical representations and mental models" (Marslen-Wilson & Tyler, 1987, p. 59). We can summarise this model of information flow as:

3. acoustic → lexical ↔ pragmatic

The predictability assumption appeared in an early version of the cohort model of word recognition, which maintained that semantic and pragmatic information can activate word candidates (Marslen-Wilson & Welsh, 1978). The more recent version of the cohort model prohibits contextually based activation of word candidates, thereby relinquishing the predictability assumption, but it still maintains that a selected word candidate is recognised when it is integrated into a "higher-level representation of the current utterance and discourse" (Marslen-Wilson, 1987, p. 98). Besides its appearance in models of linguistic processing, the predictability assumption is also widely accepted as a methodological tool for investigating processing complexity, as we shall discuss in the next section. But first we summarise two other approaches to the organisation of processes during language comprehension.

It is possible to interpret results like those of Marslen-Wilson and Tyler (1980) as showing that semantic and pragmatic information have no effect at all on the processing of linguistic information (Fodor, 1983; Forster, 1979). On that view, linguistic processes only have access to linguistic information and linguistic processing resources. Fodor argues that demonstrations such as that by Miller et al. (1951) do not show that semantic constraints influence acoustic processing but, instead, they are artefacts that occur for a variety of reasons. For example, semantic constraints may increase the listener's ability to guess the identity of obscured words after perceptual mechanisms have operated on stimulus information. In addition, some results that allegedly demonstrate the influence of background information on perceptual processing may occur because of knowledge that is internal to the linguistic processing system (Fodor, 1984). For example:

> . . . phoneme restoration illustrates top-down information flow in speech perception. It does *not*, however, illustrate the cognitive penetrability of the language input system . . . . If . . . the "background information" deployed in phoneme restoration is simply the hearer's knowledge of the words in his language, then that counts as top-down flow within the language module; on any remotely plausible account, the knowledge of a language includes knowledge of its lexicon. (Fodor, 1983, p. 77)

Thus, another possible model of information flow is:

4. acoustic ↔ lexical ↔ syntactic ↔ semantic → pragmatic

where the pragmatic level does not influence processing at any linguistic level, though higher *linguistic* levels may influence lower levels. Fodor argues that the "encapsulation" of linguistic processes with regard to pragmatic information occurs because of the "fixed neural architecture" that is characteristic of language: ". . . hardwired connections indicate privileged paths of informational access; the effect of hardwiring is thus to facilitate the flow of information from one neural structure to another" (Fodor, 1983, p. 98). We will use the term *architectural modularity* to refer to the view that linguistic processes and "central" processes occur with distinct physiological bases. Such a physical segregation of linguistic processes from central processes would prevent the two from sharing information and processing resources.

Evidence from studies of the on-line processing of sentences with ambiguities favours some form of informational independence. These studies show that both meanings of an ambiguous word are accessed even when the preceding words in the sentence disambiguate it (e.g. Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Shapiro, Zurif, & Grimshaw, 1989; Swinney, 1979; Tanenhaus & Donnenwerth-Nolan, 1984). Similarly, both meanings of an ambiguous clause are accessed even when the discourse context disambiguates it (Holmes, 1984; Hurtig, 1978; Kurtzman, 1985). In each of these cases, higher-level pragmatic information neither suppressed the contextually irrelevant meaning, nor primed the contextually appropriate meaning. Higher-level context apparently does not limit the choices at lower levels of structure, just as predicted by the view that the architectural segregation of lower-level processes prevents them from accessing more abstract information.

A third model of the organisation of linguistic processes, which we call the *multiple representation hypothesis*, emphasises the functional differences between processes at different levels of structure. According to this model, processing proceeds simultaneously at several levels of structure. This model is distinct from architectural modularity, as Fodor (1983, p. 76) notes when he says "The claim that input systems are informationally encapsulated must be very carefully distinguished from the claim that there is top-down information flow *within* these systems." According to the multiple representation hypothesis, processes at each level apply distinct procedures to representations of specific types in order to produce different types of representations. For example, lexical processes apply principles of phonological and lexical organisation to sound energy to produce representations of possible words; syntactic processes apply principles of syntactic organisation to words to produce representations of the potential grammatical phrases of a sentence; semantic processes apply principles of the organisation of semantic roles to determine the propositional content of a sentence; and pragmatic processes apply knowledge of the organisation of discourse, human motivations, and social conventions to the meanings of sentences to produce a conceptual representation of discourse. The distinct nature of the representations at various levels entails that processes at different levels can interact only when they produce representations of the same type. For example, semantic information can influence syntactic processes only when the latter has produced a semantic representation (Townsend & Bever, 1982). The constraint that interactions between processes require similar representations prevents interactions between processes that are not adjacent, such as the semantic and the acoustic. The multiple representation hypothesis therefore accepts a restricted version of (2) in which information at a higher level of representation can influence the integration of information at the immediately lower level of representation. But the multiple representation hypothesis also contrasts with architectural modularity in claiming that processes at different levels of structure may share processing resources. This means that processing at the pragmatic level could actually impair responding to information at the acoustic level, as, for example, when knowing that a speaker intended an utterance as a request involves the use of resources that otherwise could be allocated to the detection of acoustic features. Experimental studies do suggest that several factors influence the assignment of resources to different levels of representation. For example, logical connectives (Townsend & Bever, 1978), syntactic complexity (Bever & Townsend, 1979), and the format in which printed text is presented (Townsend & Bever, 1989) can all influence accessibility to representations of semantic and syntactic structure. We will show below that this competition between on-going processes at different levels produces results that appear to support the predictability assumption, and that it presents a critical problem for architectural modularity.

The view that different processes compete for limited resources is not new. It is consistent with the common intuition that it is hard to proof-read for meaning and form simultaneously (see, for example, Goldman & Healy, 1985; Healy & Drewnowski, 1983; Levy & Begin, 1984). It is also consistent with the well-known Stroop phenomenon, in which the meaning of a word affects the time that is required to respond to aspects of its form (e.g. Cowan & Barron, 1987). Consider the finding that it is harder to name the colour of printed letters as blue when they spell CHARTREUSE than when they spell BLUE. Such a result generally is interpreted to show that some aspects of lexical access and ink-naming proceed in parallel, but independently. Lexical access involves recognising an entry in the mental lexicon, whereas ink-naming involves retrieving and producing a stored label for the perceptual event that corresponds to a particular wavelength of light energy. Though these processes are independent at some points, they ultimately do compete for limited processing resources, perhaps when both have yielded a representation of similar type. However, making lexical access easier by presenting a more common word, such as RED, presumably would increase even more the time that is needed to name the colour of the ink as blue. This suggests that when the meaning of a word is more readily available, it increases the competition between the representations of hypothetical word candidates and the label for ink colour. The multiple representation hypothesis is a linguistic version of this general model of the Stroop phenomenon: In both cases, processes at different levels of structure draw on representations of different types, they proceed in parallel, and they share some common processing resources.

The preceding discussion highlights the very different claims that the predictability assumption, architectural modularity, and the multiple representation hypothesis make about the organisation of linguistic processes. The predictability assumption proposes that representations at more abstract levels of structure restrict access to hypotheses at more concrete levels of structure. Architectural modularity proposes that both informational and processing resources for different levels of structure are physically segregated. The multiple representation hypothesis proposes that processes at different levels can interact when they share similar representations, but that they can compete for limited processing resources.

This report presents an experimental test of these models. We first discuss the conditions that are needed to test the predictability assumption against the multiple representation hypothesis. Then we briefly document some studies that superficially appear to support the predictability assumption, and we show how it has been used as a tool for investigating the organisation of linguistic processes. Next, we present two experiments that demonstrate critical problems for both the predictability assumption and architectural modularity. Then we consider the implications of the results for possible models of the organisation of linguistic processes.

THE PREDICTABILITY ASSUMPTION IN PSYCHOLINGUISTIC RESEARCH

An experimental test that distinguishes the predictability assumption from the two versions of informational independence – architectural modularity and the multiple representation hypothesis – requires three conditions:

(A) *There must be variation in the constraints at a relatively abstract level of representation*. This condition is necessary to fulfil the essential claim of the predictability assumption – that increased probability of an element at an abstract level can limit the choices among possible elements that are initially accessed at a more concrete level.

(B) *There must be a task that requires referring to each of two different levels of representation*. This requirement is necessary to determine whether processes at two levels compete for processing resources. According to the multiple representation hypothesis, increased constraints at an abstract level of representation facilitate processing *at that level* during normal comprehension, thereby freeing resources for processing at the concrete level. Requiring tasks at two levels prohibits such shifts in processing resources.

(C) *The two levels must each represent a natural linguistic form*. All three models make claims about the organisation of various levels of linguistic processing. No claim is made about the relation between linguistic and non-linguistic processing.

Several previous studies superficially appear to support the predictability assumption, but none has fulfilled all three of the above conditions. It will be useful to examine briefly why the predictability assumption has not been demonstrated before we consider examples of how it has been used in research methodology.

A large set of studies does not meet condition B. As noted earlier, one set of experiments varied the probability that a particular element occurred at one level, and observed facilitation in task performance at a lower level (e.g. Miller et al., 1951; Miller & Isard, 1963; Pollack & Pickett, 1963). For example, well-formed sentential contexts, like "Accidents kill motorists on the highways" facilitated the identification of words, compared to ill-formed contexts like "Around accidents country honey the shoot" (Miller & Isard, 1963). Because the listeners in these studies were not required to report on any aspect of the syntactic structure or meaning, there was no semantic or syntactic task that could compete with lower-level processes, such as word recognition. Hence, the increased semantic and syntactic constraints in "Accidents kill motorists on the highways" may have facilitated semantic and syntactic processing, thereby allowing more resources to be devoted to word recognition. Another example of a failure to meet condition B comes from sentence reading time studies, which generally show that pragmatically probable sentences are read faster than less probable sentences (e.g. Bower, Black, & Turner, 1979; Townsend, Carrithers, & Bever, 1987). Such results, however, do not necessarily show that the way in which a sentence fits into the larger discourse context influences lexical or syntactic processing. Two other interpretations are possible: Facilitating pragmatic processing may either render lexical and syntactic processing unnecessary, or it may release resources for lexical and syntactic processing. A similar interpretation applies to studies that demonstrate that lexical associations improve performance on lexically dependent processes such as phoneme detection (e.g. Cairns, Cowart, & Jablon, 1981; Cole & Jakimik, 1978; Cole & Perfetti, 1980; Foss & Blank, 1980; Morton & Long, 1976; Samuel, 1981). For example, Morton and Long (1976) found that subjects detected a target phoneme like /b/ more quickly in "A sparrow sat on the branch" than in "A sparrow sat on the bed". Because the subjects had to perform at only the phonemic level, it is impossible to know whether lexical associations reduced the number of candidates to be considered in recognising the initial phoneme of the final word, or whether they facilitated word recognition, thereby freeing resources for phonemic processing.

Several studies seem to refute the predictability assumption, but actually do not because they have not fulfilled condition C. This set of studies shows that increasing the comprehensibility of a passage impairs the detectability of non-speech sounds (e.g. Britton & Price, 1981; Britton, Westbrook, & Holdredge, 1978; Britton, Holdredge, Curry, & Westbrook, 1979; Britton et al., 1983; Green, 1977). For example, Green (1977) found that increasing syntactic constraints decreased the detectability of clicks that were superimposed on the speech signal. As facilitating syntactic processing disrupted the relatively superficial process of click detection, it seems to refute the predictability assumption. However, one might argue that processing the acoustic information that is needed to detect clicks may not be an intrinsic part of speech comprehension (e.g. Liberman & Mattingly, 1989), and so the predictability assumption does not apply.

An intriguing set of investigations seems to demonstrate the dependence of lower levels of linguistic processing on higher-level linguistic information, but actually supports their independence (e.g. Cole, 1973; Garnes & Bond, 1977; Goldman & Healy, 1985; Healy & Drewnowski, 1983; Marslen-Wilson & Welsh, 1978; Warren, 1970). During comprehension, people perceptually restore missing, degraded, mispronovnced, or misspelled signals, just as the reader may have restored "v" to "u" in "mispronovnced" earlier in this sentence. Similarly, preferences for noun *vs* verb categorisation for words like "joke" and "drive" can influence the interpretation of a preceding acoustic segment as "to" or "the" (Isenberg, Walker, Ryder, & Schweickert, 1980, as reported in Elman & McClelland, 1984). Such demonstrations do not necessarily show that comprehension at a higher level limits choices at a lower level. Just as before, these results are open to two other interpretations: They may show that increasing the availability of a higher-level representation makes accurate processing at the visual or acoustic level less necessary, or they may show that it frees resources for processing at the lower level.

Despite the fact that the predictability assumption has never been proven, researchers frequently have used it as the explanatory rationale for investigating the effects of syntactic, semantic, and pragmatic probability on lexical and phonetic decisions. For example, Marslen-Wilson and Tyler (1975; 1980; Marslen-Wilson, Tyler, & Seidenberg, 1978) used the time to detect a cue word or a rhyming word as a measure that should be sensitive to variations in discourse constraints. Marslen-Wilson and Tyler (1980) defined sentential constraints by the location of the target word as early or late in a sentence and by how well-formed the sentence was; they defined discourse constraints by the presence or absence of a well-formed context sentence. Marslen-Wilson and Tyler (1980) argued that because increased sentence-level and discourse-level constraints reduced word monitoring times, both types of constraints must facilitate word recognition. Similarly, Tyler and Marslen-Wilson (1977; Marslen-Wilson & Tyler, 1987) used the time to name a visually presented word as a measure of the effect of preceding semantic information on the perception of syntactic structure. In Tyler and Marslen-Wilson's (1977) study, it was assumed that the time to name IS or ARE following phrases such as "growing flowers" demonstrated the effect of semantic probability on the activation of certain hypotheses about syntactic structure. The authors' interpretation in each of these studies depends on the validity of the predictability assumption – that a reduction in the time that is needed for detecting a word reveals the effect of semantic probability on the number of options that are considered at lower levels of processing.

The results of these studies can also be explained by the multiple representation hypothesis. In each case, the higher-level constraints may have facilitated higher-level processing and freed resources for the lower-level task. Townsend and Bever (1982) indirectly refuted Tyler and Marslen-Wilson's (1977) interpretation by showing that (a) adequate control of the materials greatly reduces the observed effect of semantic probability on word naming time, and (b) the effects of preceding context are confined to the interpretation that corresponds to a natural unit of semantic representation. Given an ambiguous phrase such as "growing flowers", a preceding context that was biased toward the gerund (V-Obj) interpretation reduced naming times for IS, but a context that was biased toward the adjectival interpretation (Adj-N) did not reduce naming times for ARE. The multiple representation hypothesis explains this difference on the grounds that a V-Obj assignment represents a more complete semantic unit than does the Adj-N assignment and, consequently, it can interact more readily with semantic context. The results therefore suggested that the effect of semantic context on word naming time occurs because increased semantic probability facilitates integration of the percept with context, not because it facilitates perception *per se*.

A TEST OF THE PREDICTABILITY ASSUMPTION

The widespread use of the predictability assumption as a methodological tool warranted a direct test of it, as outlined in conditions A–C. We examined the effect of higher-level constraints on judgements about an acoustically based property of speech – whether or not there was a change in speaker during connected discourse. The task of monitoring for a change of speaker was selected because it provides an unbiased measure of the effect of higher-level constraints, unlike tasks such as phoneme-, word-, rhyme-, and synonym-monitoring, which may provide subjects with clues to the location of the target (see Townsend, Hoover, & Bever, in prep.). Although there are several principles of discourse structure that could be used as a basis for varying the constraints on a sentence in discourse, we chose an operational definition: One group of subjects constructed stories around a set of critical sentences. Other subjects then rated the predictability of the critical sentences in these stories. The experimental materials consisted of the story contexts that produced ratings of high *vs* low probability for the critical sentence. Some examples of the obtained high- and low-probability texts are (5) and (6), respectively:

5. Mary went to South Side High School. She was voted the most popular girl in the class. Because she is a sweet cute girl the boys asked her out often . . .
6. My friend has not lived at home for years. Through high school she constantly battled with her parents. Because she is a sweet cute girl the boys asked her out often . . .

A male speaker recorded the texts. In some recordings, a single word, such as GIRL or BOYS in the last sentence of (5) and (6), was replaced by the same word that was taken from a female speaker's recording of the same text. As subjects listened to recordings of these modified texts, they determined whether or not there was a word spoken by a female speaker and, if so, which one it was. Their primary task, however, was to judge the plausibility and cohesiveness of the text and aspects of the author's personality.

This procedure fulfils the conditions for testing the predictability assumption:

Condition A: There is variation in constraints at a relatively abstract level. The processing complexity of critical sentences varies at the discourse level because of the procedure for generating and selecting texts. The processing complexity also systematically varies within sentences: there are more grammatical and selectional constraints that occur on a word when it appears later in a clause. For example, there are fewer lexical options for completing the fragment "She is a sweet cute . . .", compared to the fragment "Because she is a sweet cute girl, the . . .". Note, however, that the processing complexity at the sentence level is *not* relevant for testing the predictability assumption because of condition B.

Condition B: There are tasks that require processing at both an abstract level and at a more concrete level of structure. There is a pragmatic task of reporting on personality characteristics of the author and the cohesiveness of the text, and an acoustic task of judging a change of speaker. The requirement to perform both tasks forces subjects to choose between focusing processing resources on the acoustic task or on the pragmatic task. As noted above, our procedure does not test the validity of the predictability assumption with regard to within-sentence constraints, as condition B would require a judgement that refers to the appropriate information at the level of the sentence as well as the acoustic/lexical judgement. Our procedure utilised judgements only at the pragmatic and acoustic levels.

Condition C: Both levels of structure normally are computed during comprehension. During speech comprehension, listeners compute the kinds of information that are needed to make the acoustic and the semantic judgements. Consider the relevance of the task at the acoustic level.

Speakers differ in fundamental frequency, speech rate, and the timing of phonemically relevant events such as voice onset time (see Fant, 1973), and listeners use these properties as a baseline for interpreting subsequent acoustic information (Diehl, Souther, & Convis, 1980; Elman, Diehl, & Buchwald, 1977; Ladefoged & Broadbent, 1957; Miller, 1987; Suomi, 1984; Tartter, 1984). For example, when a single recording of a target syllable follows a recording of the context utterance "Please say what this word is", perception of the following target syllable as /bit/, /bat/, /bet/, or /but/ depends on the formant structure of the context sentence (Ladefoged & Broadbent, 1957). Notice that it is not necessary that normal comprehension involves explicit judgements of speaker identity – it is only required that normal comprehension involves computation of the kinds of information on which such judgements are made. Similarly, explicit judgements of how cohesive a text is or how happy the author was may not be made during normal comprehension, but such judgements surely require reference to information that is formed during comprehension (see Fodor, 1983, p. 59 for a discussion).

The change of speaker monitoring task allows us to compare the predictability assumption against architectural and multiple representation versions of informational independence. Each of these three models maintains that increasing constraints within a sentence can facilitate acoustic recognition, but the three models make these claims for different reasons. The predictability assumption claims that within-sentence constraints reduce the number of choices at the acoustic level. According to architectural modularity, facilitation from within-sentence constraints occurs because acoustic, lexical, and syntactic processes are indigenous to the language module, and lower-level processes can utilise higher-level constraints as long as the higher-level constraints are contained within the language module. The multiple representation hypothesis makes the same prediction as architectural modularity about within-sentence constraints on the grounds that the interface between syntactic and lexical/acoustic processing uses a single representational format. However, the three models differ in their predictions about the effects of discourse-level probability on acoustic judgements. If increased probability of a sentence in discourse reduces the number of choices that the listener must consider at the acoustic level, there should be greater accuracy in detecting a change in speaker in more probable sentences. This prediction follows from the view that pragmatic information can activate a particular proposition, which in turn can activate a particular word with particular properties of formant structure, timing, and so on. On the other hand, if information about the probability of a sentence in discourse is exogenous to an independently functioning linguistic processing module, increasing the probability of a sentence in discourse should have no effect at all on acoustic judgements. This follows from the claim of architectural modularity that acoustic/lexical processing shares neither information nor processing resources with the processes that determine the pragmatic properties of sentences in discourse. In contrast, the multiple representation hypothesis proposes the following: If facilitating an explicit pragmatic judgement attracts resources that otherwise would be deployed to more concrete levels, there will be poorer accuracy in detecting acoustic properties in more probable sentences.

EXPERIMENT 1

Method

Subjects. The subjects were 216 native English-speaking Montclair State College undergraduates who participated as part of a course requirement.

Materials. We obtained a pool of texts by asking 21 undergraduates at Columbia University to create a story around each of nine sentences that ranged in length from 10 to 13 words. The students constructed their stories so that at least two sentences preceded the critical sentence, and at least one sentence followed it. Nine graduate students in psycholinguistics rated on a scale from 1 to 5 how predictable each of the critical sentences was in the 21 story contexts, how acceptable it was, and how realistic the story as a whole was. From this pool of 21 texts for each sentence, we selected two pairs of texts for each of eight critical sentences so that the pairs differed in rated probability; for each critical sentence, one pair of texts that differed in rated predictability was rated low in acceptability, and one was rated high in acceptability. Thus, within a pair of texts that was matched for acceptability, one text had a relatively higher average rated predictability than the other. The outcome of this selection procedure was 32 texts whose characteristics are summarised in Table 1. An independent group of 16 subjects rated the 16 high-probability texts as more predictable than the low-probability texts with which they were matched [$F(1,32) = 95.6$, $P < 0.001$]. With another group of subjects, a forced-choice rank-ordering task using the four contexts that preceded each critical sentence also showed a significant effect of probability [$t(53) = 6.23$, $P < 0.01$]. The high- and low-probability texts did not differ overall in rated acceptability [$F(1,32) < 1$], rated realisticness [$F(1,32) < 1$], or in the number of sentences that preceded the critical sentence [$F(1,32) = 1.8$, $P > 0.10$].

To verify that the texts produce the typically observed effect of probability on behaviour, we obtained sentence reading times for the critical

TABLE 1
Characteristics of Texts

	Discourse Constraints	
	Low	*High*
Rated predictability	2.70	3.94
Rated acceptability	3.57	3.61
Rated realisticness	3.26	3.39
Number of preceding sentences	2.33	2.67

sentences in the various contexts. Across four lists, each sentence appeared in four different contexts that varied in rated probability and acceptability. Each list contained two practice texts in addition to the eight critical texts. Sixteen undergraduates at Columbia University read the texts in a subject-paced, sentence-by-sentence reading task. After reading an entire text, the students wrote a paraphrase of it. Sentence reading times were 16.3 msec per word faster when the sentence was highly probable [$F1\ (1,15) = 9.28$, $P < 0.01$; $F2\ (1,15) = 3.94$, $P < 0.10$].

The texts were recorded by a male speaker and by a female speaker who spoke similar midwestern U.S. dialects. Both speakers attempted to impose normal intonation patterns on their readings. The male recording was used for presenting texts to the subjects, and the female recording was used only for selecting target words. Each text contained a critical sentence that was made more or less probable by the preceding context, as determined by predictability ratings. Four versions of each text were defined by (a) which of two content words in the male recording of the critical sentence was replaced by another recording of the same word, and (b) whether the inserted word was taken from the female speaker's recording of the sentence, or from a copy of the same male speaker's recording of the sentence. The inserted word was always a noun or an adjective that occurred near the centre of the critical sentence, either as the first word of the second clause [e.g. BOYS in (5) and (6)] or the last word of the first clause [e.g. GIRL in (5) and (6)]. Word-by-word predictability ratings from an independent group of 16 subjects showed that the target words that occurred early in a clause (BOYS) were assigned lower ratings of predictability than those that occurred late in a clause (GIRL) [$F(1,7) = 17.7$, $P < 0.01$]. Tape editing was carried out using digitised computer files. The result was 16 lists that each contained the same set of eight sentences embedded in story contexts. Across lists, each critical sentence appeared in four different contexts (high probability–high acceptability, high probability–low acceptability, low probability–high acceptability, and low probability–low acceptability). In each of these contexts, the recording of one of two words in the critical sentence – low within-sentence constraint (early in a clause) *vs* high within-sentence constraint (late in a clause) – was changed in one of two ways: no change of speaker (male replacement) *vs* change of speaker (female replacement). Within each list, there was at least one change of speaker trial for each combination of discourse probability and within-sentence constraint.

Procedure. The subjects were tested in 18 groups of 7–21 in classrooms on the Montclair State College campus. Each subject received a booklet with the following instructions:

> We are investigating the relation between personality characteristics and language. We believe that people with different personalities construct different kinds of stories and that they tell stories differently. This study is designed to gather evidence for our hypothesis.
>
> You will be hearing 10 very short stories composed by 10 different college students, but recorded by a single male reader. Some of these stories seem to be entirely natural and comprehensible, while others do not. We want you to make various judgements about the naturalness of the stories and about emotional characteristics of the authors of the stories. As you hear each story, try to decide the following: how well-written is the story, how likely is it that the sequence of events in the story could actually have occurred, how emotionally stable is the author, and how happy was the author at the time he/she wrote the story. When the story is finished, we want you to indicate your judgements on rating scales which range from 1 to 5:
>
> THE STORY WAS . . . WRITTEN.
> (1 = very poorly, 2 = rather poorly, 3 = neither well nor poorly, 4 = rather well, 5 = very well)
>
> THE SEQUENCE OF EVENTS IN THE STORY IS . . . TO HAVE OCCURRED IN REAL LIFE.
> (1 = very unlikely, 2 = rather unlikely, 3 = neither likely nor unlikely, 4 = rather likely, 5 = very likely)
>
> THE AUTHOR OF THE STORY IS . . . EMOTIONALLY.
> (1 = very unstable, 2 = rather unstable, 3 = of moderate stability, 4 = rather stable, 5 = very stable)
>
> AT THE TIME OF WRITING THE STORY, THE AUTHOR WAS . . .
> (1 = very unhappy, 2 = rather unhappy, 3 = neither happy nor unhappy, 4 = rather happy, 5 = very happy)
>
> The tape recordings of these stories may sound rather odd due to the fact that they were prepared by computer. In particular, the computer inadvertently made some of the words sound like they were spoken by a woman, rather than the man who recorded the stories. Since we would like to use these recordings for further experiments, we are interested in how serious

> this problem is. Hence, we would like you to also indicate whether the story contained a word that sounded like it was spoken by a woman, what word it was, and how confident you are that the word in question sounded like it was spoken by a woman.

The subjects then heard each of the 10 texts played over a Wollensack tape-recorder. After hearing each text, the subjects recorded their judgements about discourse properties and change of speaker in their booklets. After the subjects completed the experiment, its purpose was explained to them.

Results and Discussion

A "hit" occurred when the subject correctly stated that there was a change of speaker and identified which word was changed. The subjects were moderately accurate in detecting a change of speaker: overall, the hit rate was 63% and the false alarm rate was 4%. The percentage of misses appears in Fig. 1. The miss rate was 32% when discourse constraints were low, but 41% when they were high [$F1$ (1,215) = 6.27, $P < 0.05$; $F2$ (1,7) = 6.19, $P < 0.05$]. A more lenient scoring procedure counted as "hits" only those trials on which the subject correctly stated that there was a change of speaker, regardless of whether the subject correctly reported the changed word. By this lenient scoring procedure, misses occurred less frequently for texts with low discourse constraints [$F1$ (1,215) = 6.25, $P < 0.01$; $F2$ (1,7) = 2.16, $P < 0.18$]. Including only those trials in which the subjects were most confident of their judgement still produced a significant 9% advantage in accuracy for low-probability contexts.

Within-sentence constraints had the opposite effect. Using the stringent scoring procedure, the subjects missed 55% of the target words that occurred early in a clause, but only 17% of those that occurred late in a clause [$F1$ (1,215) = 310.3, $P < 0.001$; $F2$ (1,7) = 12.6, $P < 0.001$]. Two additional analyses showed that neither type of constraint effect was an artefact of the design. First, the distribution of hits and misses differed across levels of discourse constraint [$\chi^2(1) = 8.61$, $P < 0.01$] and within-sentence constraint [$\chi^2(1) = 283.9$, $P < 0.001$]. Secondly, analysis of variance that treated each trial as an independent observation using the stringent scoring procedure produced significant effects of discourse constraints [$F(1,1046) = 5.64$, $P < 0.001$] and within-sentence constraints [$F(1,1046) = 374.6$, $P < 0.001$].

These results show that the ability to report on the identity of a speaker is actually poorer in pragmatically probable sentences. This result conflicts with the general view that increasing higher-level constraints facilitates the perception of lower levels of structure. On the other hand, within-sentence

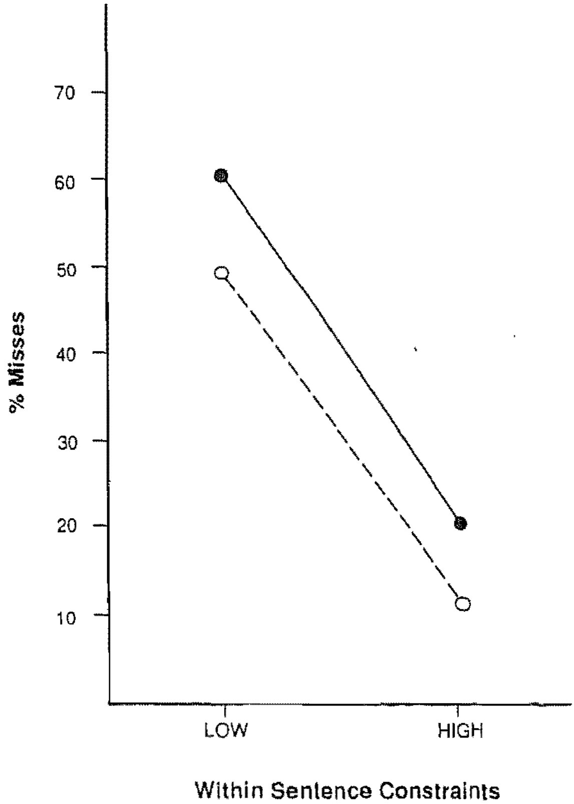

FIG. 1. Percentage of misses in detecting a change of speaker: ●—●, High discourse constraints; ○--○, low discourse constraints.

constraints did improve the subjects' ability to report on a change of speaker. This pattern of results suggests that information may flow between *adjacent* levels during on-going comprehension, and that processes at different levels of structure share resources. Our experimental procedure, however, does allow for alternative explanations. First, it is possible that the observed probability effects occur after perception. For example, increased discourse constraints may actually have facilitated acoustic processing by reducing the number of choices at the acoustic level, but the results of this processing may also have been forgotten more easily for more probable sequences. Secondly, the fact that the subjects were not required to make acoustic judgements on-line leaves open the possibility that they were not continually choosing between focusing on the acoustic *vs* the semantic level.

EXPERIMENT 2

In order to determine whether the observed probability effects are due to the organisation of on-line processes, a second experiment measured how quickly listeners could detect a change of speaker. According to the predictability assumption, architectural modularity, and the multiple representation hypothesis, response times to detect target words will be faster when within-sentence constraints are greater. However, the predictability assumption maintains that increasing discourse-level constraints will reduce response times. Architectural modularity maintains that increasing discourse-level constraints will have no effect on response times. The multiple representation hypothesis maintains that increasing discourse-level constraints will increase response times.

Method

Materials. One set of tapes contained distractor trials in which there was no change of speaker, and one set contained no distractor trials. The tapes for the distractor set consisted of the 16 lists that had been used in Experiment 1, except that a 50-msec, 500-Hz timing tone was added on a second channel at the onset of each target word. The set of tapes with no distractors consisted of eight lists in which all of the trials were recorded from the target-present trials of the distractor set. In each no-distractor list, there was one trial for each combination of rated discourse probability, rated acceptability, and target position within the critical sentence. Each distractor list contained at least one target-present trial for each combination of discourse probability and within-sentence constraints. There were two practice trials and eight test trials in each list.

Procedure. The subjects served in either the distractor or the no-distractor group. As in the first experiment, the subjects listened to stories for the purpose of rating discourse-level characteristics of each story on one of four dimensions. The subjects were instructed to press a response key as soon as they heard a change from a male speaker to a female speaker. The timing tone triggered a Hunter timer, which stopped when the subject pressed a response key.

Subjects. A total of 32 undergraduates at Columbia University were paid for their participation, 16 of whom served in the distractor condition and 16 of whom served in the no-distractor condition. The subjects were tested individually.

Results and Discussion

The miss rate was 4.3% for the distractor group and 2.3% for the no-distractor group. In 4.5% of the trials, response times were greater than two standard deviations over the cell mean; these were replaced by the corresponding value of two standard deviations over the cell mean. Statistical analysis was based on each subject's mean response times for correct responses, with missing data replaced by the cell mean. The variables that were included in statistical analyses were discourse-level constraints, sentence-level constraints, acceptability, and condition. The mean response times for hits appear in Fig. 2. Increased discourse-level constraints increased the time to detect a change in speaker. The mean response times were 29 msec faster in less probable sequences [$F1$ $(1,30) = 4.31$, $P < 0.05$;

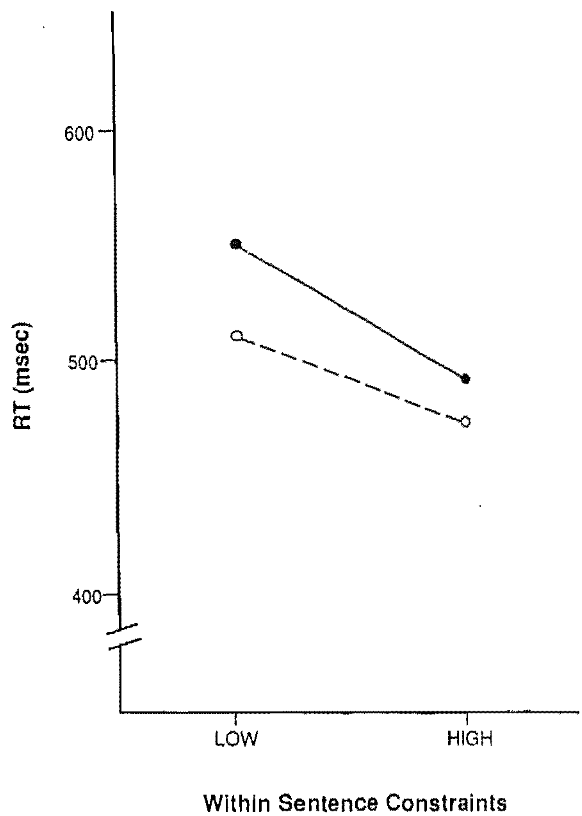

FIG. 2. Mean reponse times for detecting a change of speaker: ●—●, High discourse constraints; ○--○, low discourse constraints.

*F*2 (1,7) = 3.71, $P < 0.10$]. In contrast, response times were slower when sentence-level constraints were low [*F*1 (1,30) = 7.98, $P < 0.01$; *F*2 (1,7) = 3.71, $P < 0.10$]. Neither discourse-level nor sentence-level probability interacted with experimental condition (all *F*s < 1), though response times were faster in the no-distractor condition [*F*1 (1,30) = 9.65, $P < 0.01$; *F*2 (1,7) = 33.3, $P < 0.001$]. The advantage for low-probability discourses occurred across conditions, except for targets in contexts that had high sentential constraints but low ratings of discourse acceptability [*F*1 (1,30) = 1.77, $P > 0.10$; *F*2 (1,7) = 12.7, $P < 0.01$].

GENERAL DISCUSSION

The argument against the predictability assumption can be summarised as follows: The critical sentences were judged to be more probable in some discourse contexts than in others. The subjects read the critical sentences more quickly when they appeared in the more probable contexts. If we invoke the predictability assumption, we conclude that increasing the probability of a sentence in discourse reduces the number of options that are considered at the syntactic, lexical, or sublexical levels. Hence, decisions about the acoustic properties of a segment of speech should be easier when pragmatic information supports the sentence that contains the target. The speaker detection results showed just the opposite: Increasing the probability of a sentence in discourse actually impaired listeners' ability to report a change in speaker. Therefore, both interpretations of the identity of the speaker of the target word must have been considered despite higher-level constraints, directly contradicting the predictability assumption. Because activations of features at higher levels do not "cascade" freely down to features at lower levels, we can reject

2. acoustic ↔ lexical ↔ syntactic ↔ semantic ↔ pragmatic

In contrast to the effects of discourse-level constraints, increased constraints within a sentence facilitated detection of a change of speaker. These results are similar to those of word monitoring studies in which target words were easier to detect when they appeared later in a sentence (Marslen-Wilson & Tyler, 1975; 1980). The similarity in within-sentence effects in Marslen-Wilson's studies and the present one rules out one artefactual explanation of the observed speaker monitoring results: It eliminates the hypothesis that the speaker monitoring effects depend on an interaction between the fundamental frequencies of different speakers and the location of the target within a clause. Instead, the within-sentence effect must reflect the organisation of lexical, syntactic, and semantic processes at the level of the sentence. Notice again, however, that the facilitation from increased within-sentence constraints cannot be interpreted as support for the predictability assumption. Because the subjects were not required to make explicit judgements that refer to within-sentence constraints independently of the acoustic task, as required by condition B, it is possible that within-sentence constraints simply facilitated processing at the sentence level, freeing resources for processing at the acoustic level. Thus, two interpretations of the overall results are possible. The first interpretation is that higher-level information may facilitate processing at a lower level only when the two correspond to adjacent levels of linguistic representation. The second interpretation is that facilitation from higher-level information occurs only when the two levels are not both used for some explicit task.

The fact that discourse-level and sentence-level constraints have opposite effects on the detection of speaker-relevant acoustic properties presents difficulties for one potential explanation of the observed discourse-level inhibition. Consider the view that discourse-level inhibition occurred because speaker identification is not a "typical" aspect of comprehension. According to this view, typical processes such as word recognition are subject to contextual facilitation, but non-typical levels such as identifying the speaker of a word are not. We argued against this view earlier on the grounds that computation of speaker-relevant information must be "typical" because the acoustic properties of different speakers' speech influence the recognition of phonemes and words, even when listeners are not required to identify speakers. The fact that increased constraints within sentences improve the ability to detect a change in speaker at the same time that increased discourse-level constraints impair it renders such arguments unnecessary.

Consider the implications of the opposite effects of discourse-level and sentence-level constraints for the cohort model of word recognition (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987). In the cohort model, the acoustic information that corresponds to a word-initial phoneme such as /g/ activates a set of candidates for the word, such as GIRL, GUM, GORILLA, GRAB, GULP, GOOD, and so on. According to the earlier model (Marslen-Wilson & Welsh, 1978), contextual information and additional acoustic information eventually reduce this cohort to a single candidate. For example, the acoustic information corresponding to /gu/ would eliminate GRAB and GOOD from the word-initial cohort, that corresponding to /gur/ would eliminate GUM and GULP, and that for /gurl/ GORILLA. However, in this cohort model, the sentence and discourse context could eliminate candidates as well. For example, a sentence-level context like "Because she is a sweet cute . . ." would eliminate all verb candidates from the initial cohort, as the sentence context demands a noun or another adjective. The present results suggest that, while sentence-level context might eliminate candidates from the cohort, discourse-level context does not. Thus, the hypothesis that there is *unconstrained access* to various information sources is not correct.

Consider the implications of the opposite effects of constraints at different levels for the more recent version of the cohort model (Marslen-Wilson, 1987; see also Norris, 1986). In this model, a single candidate is selected from the candidates in the cohort on the basis of additional acoustic information. Contextual information operates only in the process of "mapping word-senses into higher-level representations" (Marslen-Wilson, 1987, p. 99). This integration of word candidates can occur "even though the sensory input may not have fully differentiated the word-form associated with this interpretation" (Marslen-Wilson, 1987, p. 98). Variations in the amount of acoustic processing that occurs before integration provide a possible explanation for discourse-level inhibition: The fact that integration is easier in more probable discourse contexts means that recognition occurs with less acoustic processing. Thus, there is less opportunity for making the acoustically based judgement in semantically probable discourses, and the speed and accuracy of detecting speaker changes are reduced. According to this hypothesis, the amount of *sub-integrative* processing accounts for discourse-level inhibition. Such a mechanism, however, now has difficulty explaining the opposite effects of sentence-level probability. Presumably, integration is also easier in more probable sentence contexts, so these contexts should lead to a decrement in performance on the acoustic task as well. The fact that sentence- and discourse-level constraints have opposite effects forces a mechanism that distinguishes the two information sources and uses them differently; that is, we must acknowledge a distinct level of representation that mediates the lexical representation and the discourse model (cf. Marslen-Wilson & Tyler, 1987).

Still a third possible explanation of the discourse-level inhibition appeals to *shifts in processing resources*. An analogy that illustrates this hypothesis is as follows: You are struggling to unlock a door with a key, but cannot make the key fit into the lock. In this case, you have the choice of continuing to struggle, or to examine carefully the key to determine if it is the right one. Similarly, when it is hard to form a discourse model (fitting the key into the lock), processing resources are shifted to the input data (the key). This analogy seems plausible, but it still leaves the puzzle of why sentence-level constraints do not have the same effect. If it is relatively hard to form a representation of sentence meaning, as when there are low within-sentence constraints, why doesn't this elicit an orientation to the input data, and facilitate target recognition?

The basic phenomenon to be explained is that both low discourse-level constraints and high sentence-level constraints facilitate detection of acoustic features. We have considered several single-mechanism explanations –

the unconstrained access hypothesis, the sub-integrative processing hypothesis, and the resource-shift hypothesis. In each case, we have encountered difficulty with the fact that discourse-level and sentence-level constraints have opposite effects on the detection of acoustic features. To account for why the same variable has opposite effects under different conditions, we are forced to propose different underlying mechanisms for the two conditions. One solution is to distinguish between integrating different information sources and processing resources: the effects of sentence-level constraints are due to integrating acoustic/lexical processing with grammatical and semantic constraints within sentences, and the effects of discourse-level constraints are due to shifts in processing resources. To elaborate on this model:

1. Independent processes at the acoustic, lexical, syntactic, and pragmatic levels simultaneously produce candidate representations at the appropriate levels. Each of these processes operates on specific kinds of input representations to produce specific kinds of output. As discussed earlier, lexical processes operate on information about acoustic structure to yield candidate lexical items, syntactic processes operate on lexical representations to organise words into phrases, pragmatic processes operate on propositions to yield candidate representations of discourse structure, the speaker's intentions, and so on.

2. Interactions between processes occur *when one process presents information in a form that is relevant for another.* Overlapping representations may occur when a lower level of processing establishes candidate units on which a more abstract process operates, for example, when acoustic/lexical processing produces candidate lexical representations on which syntactic processes operate. Increased syntactic constraints can then influence the selection of one out of several candidate words and the corresponding speaker identity. To illustrate how such interactions might occur, suppose that acoustic/lexical processing makes available two candidate words for a target acoustic segment. Suppose that acoustic processes have determined that this acoustic segment has certain timing properties, such that the actual word is GIRL if the segment had been uttered by one speaker, and it is CURL if it had been uttered by the other speaker. If the preceding sentential context is "She is a sweet cute . . .", the emerging syntactic representation allows selection of the GIRL candidate, and simultaneously establishes the identity of the speaker. If the preceding sentential context provides fewer constraints, as in "The . . .", there is little reason to select one word candidate over the other, and little basis for deciding whether the timing properties signal a change in speaker. Hence, increased within-sentence constraints may facilitate selection among candidate pairings of

acoustic and lexical representations. The principle of overlapping representations prohibits such interactions between discourse-level constraints and acoustic/lexical processing, as the former governs the organisation of propositions and the latter yields lexical representations.

3. Fixing a representation at any level momentarily draws limited and shared processing resources to the fixed unit. Because the targets that occurred in highly probable within-sentence contexts corresponded to a complete organisation of lexical representations into a syntactic unit, there is facilitation for lexical representations that fit the syntactic representation. Highly probable discourse contexts, however, facilitate the organisation of propositions into a discourse representation; they draw processing resources towards the pairing of semantic and pragmatic representations and away from lexical representations and their interface with syntactic representations. As a result, increased discourse-level probability inhibits detection of acoustic features.

The model that is depicted by (1)–(3) leads to a reinterpretation of previous results. Take, for example, the word monitoring results of Marslen-Wilson and Tyler (1975; 1980) that were mentioned earlier. They found that well-formed contexts at the sentence level and at the discourse level had similar effects on word monitoring times, and suggested that there is a direct mapping between words and the discourse representation, without any intermediate syntactic or semantic representations. According to the present evidence and the model outlined above, Marslen-Wilson and Tyler's word monitoring results are due to the allocation of resources that depend on discourse-level constraints. In particular, the facilitation in word monitoring times that Marslen-Wilson and Tyler found for the earliest words in a sentence is due to the fact that well-formed contexts enable the listener to form a coherent discourse representation, which, in the absence of any explicit discourse-level task, frees resources for word monitoring. But well-formed discourse contexts do not limit the number of choices that listeners consider during word recognition.

In the proposed model, there is limited interaction between adjacent levels of representation (such as the lexical and the syntactic) but no interaction between non-adjacent levels (such as the lexical and the pragmatic). This occasional informational independence follows from the fact that computations at different levels are performed on different representations. These computations are "encapsulated" entirely as a result of their functions: Computations at distinct levels of representation will not interact in so far as they are couched in distinct representational languages. As noted earlier, this functional interpretation of the encapsulation of processes contrasts with architectural modularity, in which the internal operations of the language module are physically barred from pragmatic processes, so that they cannot share processing resources. However, the discourse-level inhibition in detecting acoustic properties demonstrates that pragmatic processes do share resources with acoustic/lexical processes. This falsifies the prediction of architectural modularity that pragmatic processes have no effect of any kind on linguistic processes. Notice that architectural modularity cannot account for the discourse-level inhibition with the principle that information from lower levels is simply more difficult to access (see Fodor, 1983, pp. 55–60): Poorer access to a lower level of representation only predicts that detecting speaker changes is always harder than detecting discourse-level properties.

The present results have distinguished two independent factors that operate within a processor: the information that facilitates the choice of a representation, and the computational resources that are used to construct it. Facilitating the choice of representation should not make building that representation less demanding of resources. Consider the analogy of an automatic construction machine that assembles piles of nails, wood, glass, and metal into houses. The choice of which house this machine constructs (e.g. a Victorian *vs* an A-frame) can be facilitated by presenting it with materials that fit together in only one way, e.g. windows with a trapezoidal shape might demand an A-frame. In this case, the machine sets about using all of its energy to construct the only house afforded by the materials. Presumably, such a house will be constructed faster than one for which the machine is unguided. But if the machine is also called on to serve some other function, such as counting the number of nails, a surprising fact emerges: there will be less energy available to check the nails when the self-guiding house is under construction, than when the machine has not yet started. This follows from the fact that, once the machine has begun constructing the house, it is fully occupied with it. That is, the self-guiding house construction immediately absorbs the full resources of the machine. Although constructing the house hit-or-miss may take longer, it does not occupy continuously all of the machine's resources. The model that underlies this analogy is a theory of what happens during the construction of a discourse-level representation. The variations in the discourse-level constraints in our materials influence how easy it is to determine which discourse model to construct, but they do not influence how easy it is to construct that model. In fact, the data suggest that constraining the discourse model temporarily draws resources away from lower levels of representation just because it initiates immediately the assembly of the discourse model.

The present discussion has emphasised the conservative view that there is competition for limited processing resources that can be devoted to the products of processing at different levels, rather than a sharing of resources for carrying out computations at different levels. We have adopted this conservative view because it is premature to say that processes at the discourse and lexical levels share resources for *any* of their computations. Because we selected texts by having subjects rate the probability of a particular sentence in the context of a preceding narrative, which itself was written by other subjects, the most we can say about the facilitating contexts is that they have the status of word associations, i.e. they are operationally defined in the same way as norms for word associations like SALT–PEPPER. Just as the associative priming that is based on word association norms may occur without computing intermediate representations, the discourse-level priming of our sentences may have drawn on no computational resources. On this interpretation, forming these discourse-level associations does not draw on shared processing resources, but carrying out computational operations and responding to discourse-level associations does. If there were discourse-level computations in these texts – comparable to the computation of an intermediate representation of being at the seaside, where one frequently buys ice cream, when one gives ICE CREAM as an associate to SKY – it would have been appropriate to say that variations in discourse-level constraints influenced the resources that are available for carrying out these computations. Because we have no account of these intermediate discourse-level computations, further research will be required to determine if any processes at the discourse and lexical levels share resources.

Manuscript received April 1990
Revised manuscript received August 1990

REFERENCES

Bever, T. G. & Townsend, D. J. (1979). Perceptual mechanisms and formal properties of main and subordinate clauses. In E. Walker & W. Cooper (Eds), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, pp. 159–226. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11*, 177–220.

Britton, B. K. & Price, K. (1981). Use of cognitive capacity in reading: A performance operating characteristic. *Perceptual and Motor Skills, 52*, 291–298.

Britton, B. K., Westbrook, R. D., & Holdredge, T. S. (1978). Reading and cognitive capacity usage: Effects of text difficulty. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 582–591.

Britton, B. K., Holdredge, T. S., Curry, C., & Westbrook, R. D. (1979). Use of cognitive capacity in reading identical texts with different amounts of discourse level meaning. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 262–270.

Britton, B. K., Graesser, A. C., Glynn, S. M., Hamilton, T., & Penlaud, M. (1983). Use of cognitive capacity in reading: Effects of some content features of text. *Discourse Processes*, *6*, 39–57.

Cairns, M. S., Cowart, R. D., & Jablon, A. D. (1981). Effects of prior context upon the integration of lexical information during sentence processing. *Journal of Verbal Learning and Verbal Behavior*, *20*, 445–453.

Cole, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception and Psychophysics*, *1*, 153–156.

Cole, R. A. & Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing*, pp. 67–116. London: Academic Press.

Cole, R. A. & Perfetti, C. A. (1980). Listening for mispronunciations in a children's story: The use of context by children and adults. *Journal of Verbal Learning and Verbal Behavior*, *19*, 297–315.

Cowan, N. & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception and Psychophysics*, *41*, 393–401.

Diehl, R., Souther, A., & Convis, C. (1980). Conditions on rate normalization in speech perception. *Perception and Psychophysics*, *27*, 435–443.

Elman, J. L. & McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. Lass (Ed.), *Speech and language*, Vol. 10, pp. 337–374. London: Academic Press.

Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustic Society of America*, *62*, 971–974.

Fant, C. G. M. (1973). A note on vocal tract size factors and nonuniform *F*-pattern scalings. In C. G. M. Fant (Ed.), *Speech sounds and features*, pp. 84–93. Cambridge, Mass.: MIT Press.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, Mass.: MIT Press.

Fodor, J. A. (1984). Precis of the Modularity of Mind. *Behavioral and Brain Sciences*, *8*, 1–5.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In E. Walker & W. Cooper (Eds), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, pp. 27–85. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Foss, D. J. & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology*, *12*, 1–31.

Garnes, S. & Bond, Z. S. (1977). The influence of semantics on speech perception. *Journal of the Acoustical Society of America*, *61*, 565(A).

Goldman, H. B. & Healy, A. F. (1985). Detection errors in a task with articulatory suppression: Phonological recoding and reading. *Memory and Cognition*, *13*, 463–468.

Green, D. W. (1977). The immediate processing of sentences. *Quarterly Journal of Experimental Psychology*, *29*, 135–146.

Healy, A. F. & Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 413–426.

Holmes, V. (1984). Parsing strategies and discourse context. *Journal of Psycholinguistic Research*, *13*, 237–257.

Hurtig, R. (1978). The validity of clausal processing strategies at the discourse level. *Discourse Processes*, *1*, 195–202.

Kurtzman, H. S. (1985). *Studies in syntactic ambiguity resolution*. Unpublished doctoral dissertation, Massachusetts Institute of Technology (distributed by the Indiana University Linguistics Club).

Ladefoged, P. & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*, 98–104.

Levy, B. & Begin, J. (1984). Proofreading familiar text: Allocating responses to perceptual and conceptual processes. *Memory and Cognition*, *12*, 621–632.

Liberman, A. M. & Mattingly, I. G. (1989). A specialization for speech perception. *Science*, *243*, 489–494.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*, 71–102.

Marslen-Wilson, W. D. & Tyler, L. K. (1975). Processing structure of sentence perception. *Nature*, *257*, 784–786.

Marslen-Wilson, W. D. & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*, 1–71.

Marslen-Wilson, W. D. & Tyler, L. K. (1987). Against modularity. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding*, pp. 37–62. Cambridge, Mass.: MIT Press.

Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

Marslen-Wilson, W. D., Tyler, L., & Seidenberg, M. (1978). Sentence processing and the clause boundary. In W. Levelt & G. Flores d'Arcais (Eds), *Studies in the perception of language*, pp. 219–246. New York: John Wiley.

McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part I: An account of basic findings. *Psychological Review*, *88*, 375–407.

Miller, G. A. & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, *2*, 217–228.

Miller, G. A., Heise, G., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, *41*, 329–335.

Miller, J. (1987). Mandatory processing in speech perception: A case study. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding*, pp. 309–322. Cambridge, Mass.: MIT Press.

Morton, J. & Long, J. (1976). Effect of word transitional probability of phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, *15*, 43–51.

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition*, *22*, 93–136.

Pollack, I. & Pickett, J. M. (1963). The intelligibility of excerpts from conversation. *Language and Speech*, *6*, 165–171.

Samuel, A. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, *110*, 474–494.

Schank, R. & Birnbaum, L. (1984). Memory, meaning, and syntax. In T. Bever, J. Carroll, & L. Miller (Eds), *Cognitive, philosophical, and computational foundations of language*, pp. 209–252. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Seidenberg, M., Tanenhaus, M., Leiman, J., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14*, 489–537.

Shapiro, L. P., Zurif, E. B., & Grimshaw, J. (1989). Verb processing during sentence comprehension. *Journal of Psycholinguistic Research*, *18*, 223–243.

Suomi, K. (1984). On talker and phoneme information conveyed by vowels: A whole spectrum approach to the normalization problem. *Speech Communication*, *3*, 199–209.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 523–534.

Tanenhaus, M. K. & Donnenwerth-Nolan, S. (1984), Syntactic context and lexical access. *Quarterly Journal of Experimental Psychology*, *36A*, 649–661.

Tartter, V. C. (1984). Laterality differences in speaker and consonant identification in dichotic listening. *Brain and Language*, *23*, 74–85.

Townsend, D. J. & Bever, T. G. (1978). Inter-clause relations and clausal processing. *Journal of Verbal Learning and Verbal Behavior*, *17*, 509–521.

Townsend, D. J. & Bever, T. G. (1982). Natural units of representation interact during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *21*, 688–703.

Townsend, D. J. & Bever, T. G. (1989). Expertise and constraints in interactive sentence processing. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, pp. 482–489. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Townsend, D. J., Carrithers, C., & Bever, T. G. (1987). Listening and reading processes in college- and middle school-age readers. In R. Horowitz & J. L. Samuels (Eds), *Comprehending oral and written language*, pp. 217–242. London: Academic Press.

Townsend, D. J., Hoover, M., & Bever, T. G. (in prep.). The use of monitoring tasks for investigating the organization of speech comprehension processes.

Tyler, L. K. & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 683–692.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 393–395.