Real Language Users

Ngoni Chipere

Research Centre for English and Applied Linguistics, University of Cambridge

The idea of a perfectly competent but resource limited language user is the basis of many models of sentence comprehension. It is widely assumed that linguistic competence is a) uniform; b) generative; c) autonomous; d) automatic and e) constant. It is also believed that the free expression of these properties is frustrated by limits in the availability of computational resources. However, no firm experimental evidence for the classical language user appears to exist. Negative evidence for each assumption is reviewed here and the notion of resource limitations is shown to be suspect. An experiment is reported which tested each of the five assumptions underlying the conventional notion of linguistic competence. It was found that native speakers of English a) differed in grammatical skill; b) often failed to display productivity; c) violated syntax in favour of plausibility; d) expended conscious effort to comprehend some sentences and e) appeared to adapt to novel structures as the experiment progressed. In line with previous studies, a relationship was found between comprehension skill and formal education. A new finding is that highly educated non-native speakers of English can outperform less educated native speakers of English in comprehending grammatically challenging English sentences. The results indicate that the classical language user is an inaccurate model of real language users, who appear to differ considerably in linguistic skill. A number of specific questions for further research are raised.[†]

INTRODUCTION

Research in sentence processing is often carried out on the assumption that language users know their native language perfectly. This idea appears to originate from Chomsky's (1965) statement that:

> *Linguistic theory is concerned with an ideal speaker-listener, in a completely homogenous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.* (1965:3).

It is not made clear here how real language users might differ from ideal language users. One possibility is that real language users know their language perfectly but suffer from memory limitations. Another possibility is that real language users may not know their language perfectly regardless of memory capacity. On at least one occasion, Chomsky clearly suggests the second possibility:

> *... I would be inclined to think, even without any investigation, that there would be a correlation between linguistic performance and intelligence; people who are intelligent use language much better than other people most of the time. They may even know more about language; thus when we speak about a fixed steady state, which is of course idealized, it may well be (and there is in fact some evidence) that the steady state attained is rather different among people of different educational level [...] it is entirely*

[†] I would like to thank John Williams, James Blevins and Anne Gilman for commenting on the draft.

conceivable that some complex structures just aren't developed by a large number of people, perhaps because the degree of stimulation in their external environment isn't sufficient for them to develop. (1980:175-6).

While Chomsky, at least in this instance, acknowledges the possibility of individual differences in linguistic competence between native users of the same language, many of his followers have taken ideal and real language users to be the same (see Ney, 1993) with the only difference being that real language users suffer from performance limitations which constrain their expression of grammatical knowledge. The perfectly competent but resource limited language user (henceforth the classical language user) is the template for many models of sentence comprehension (see, among many others, Cowper, 1976; Frazier and Fodor, 1978; Pulman, 1986; Blank, 1989; Gibson, 1991; Just & Carpenter, 1992; Stabler, 1994; Lewis, 1996; Fodor and Pylyshyn, 1988 and Caplan & Waters, in press). It is assumed in such models that linguistic competence is a) uniform; b) generative; c) autonomous; d) automatic and e) constant and that the expression of these properties is constrained by limits in computational resources. Despite its popularity among researchers, however, it is difficult to find experimental evidence for the five assumptions of the classical model. In fact, there is considerable experimental evidence against each of them. The idea that computational resources have fixed limits is also suspect. What appears to have happened instead is that the assumptions of the classical model have been allowed to penetrate investigative practices in ways which shield it from direct scrutiny. In addition, the highly specialised nature of sentence processing research has prevented neighbouring fields of research from exposing the weaknesses of the model. The following literature review will show that investigative practices are often influenced by assumptions which appear questionable in the light of findings established in other fields of psychology.

Uniform Competence. It is generally assumed that all normal adult users of a language have the same underlying linguistic competence. This assumption has been institutionalised in sampling and analytical procedures. Firstly, experimental subjects are usually selected from the university student population. Secondly, individual differences in experimental performance are often consigned to the error term during analysis of variance (see Eysenck & Keane, 1995:466; Osterhout, 1997 and Sasaki, 1997). In other words, experiments are designed and their results are analysed on the prior assumption that competence is uniform. However, experimental evidence for this assumption is hard to find. On the other hand, empirical evidence for individual differences in linguistic ability has been found along at least eleven dimensions: 1) in the making of grammaticality judgements (Hill, 1961 and Spencer, 1972); 2) in the making of acceptability judgements (Greenbaum & Quirk, 1970 and Mills & Hemsley, 1974); 3) in understanding novel noun compounds (Geer, Gleitman & Gleitman, 1971); 4) in knowledge of lexical category membership (Cupples & Holmes, 1992); 5) in knowledge of lexical probabilistic constraints (Pearlmutter & MacDonald, 1995); 6) in the ability to cope with syntactic ambiguity (Lefever and Ehri, 1974; Cupples & Holmes, 1992 and Pearlmutter & McDonald, 1995); 7) in the ability to cope with decreases in syntactic predictability (Graesser et al, (1980); 8) in brain wave responses to syntactic anomaly (Osterhout, 1997); 9) in parsing exceptional constructions (Chomsky, 1969; Kramer, Koff & Luria, 1970 and Sanders, 1974); 10) in assigning constituent structure (Huey, 1908; Dearborn & Anderson, 1937; Cromer, 1970; Levin & Caplan, 1970; Muncer & Bever, 1984; Cupples & Holmes, 1987 and Dabrowska, 1997) and 11) in assigning thematic roles (Bates et al, 1982; Wulfeck et al, 1986; Kilborn & Cooreman, 1987; Harrington, 1987; Kail, 1989; McDonald, 1989; Kilborn & Ito, 1989 and Sasaki, 1997). In the light of these studies, the almost exclusive use of data from university student subjects to make generalisations about all normal adult native users of a language is clearly harzadous, particularly in view of the relationship found between level of education and grammatical skill (Geer, Gleitman & Gleitman, 1971; Cox, 1976; Baruzzi, 1982; Karanth, Kudva & Arpana, 1995 and Dabrowska, 1997). Apart from the reported effects of education on comprehension skill, there is also evidence that individual differences in cognitive style are reflected in patterns of language comprehension. Lefever & Ehri, (1976) found a correlation between the ability to shift mental set and sentence disambiguation ability. This finding is related to more recent reports of correlations between reading ability and sentence disambiguation ability (Cupples & Holmes, 1987 and Pearlmutter & MacDonald, 1995). These findings are important because sentence disambiguation has been assumed to be part of linguistic competence and, according to Mitchell (1994:376), “much of the empirical work on parsing hinges on the way people handle structurally ambiguous sentences”. Some models of sentence processing depend heavily on ambiguity resolution phenomena (eg. Frazier and Clifton, 1996). The discovery that these phenomena are subject to individual variation raises problems for such models.

Generativity. It has been said that "Obviously, every speaker of a language has mastered and internalised a generative grammar that expresses his knowledge of the language." (Chomsky, 1965:8; but see the earlier quote). Accordingly, “most theories of sentence processing incorporate the claim that parsing is both fast and grammatically controlled” (Frazier, 1998:126). Ultimately, the assumption that parsing is based on grammar appears to rest on the argument that language users can understand novel sentences. The argument is not compelling, however, since the ability to understand novel sentences does not entail the ability to understand *all* grammatical sentences, any more than the ability to solve novel mathematical problems entails the ability to solve all mathematical problems. The argument needs to be substantiated by experimental tests which demonstrate that the ability to understand complex sentences increases indefinitely as processing constraints are reduced. There does not appear to ever have been any concerted effort to test the notion of generativity in this way. After forty years of sentence processing research, there is still no clear evidence for the psychological reality of grammatical rules (see reviews in Levelt, 1974 & 1978; Bock & Loebell, 1990 and Tanenhaus & Trueswell, 1995). In fact, there is evidence that some native speakers of English "do not spontaneously assign a phrasal organisation to text as they read" (Cupples and Holmes, 1987:180; see also Huey, 1908; Dearborn & Anderson, 1937; Cromer, 1970; Levin & Caplan, 1970; Muncer & Bever, 1984). If it is assumed that all language users suffer from the same limitations in memory capacity (see models reviewed in Lewis, 1996), then it is difficult to accommodate individual differences in the ability to structure linguistic input within the classical model. An obvious way around the problem is to attribute individual differences in performance to individual differences in working memory capacity. This is the approach of Just and Carpenter (1992). However, there is a serious problem with capacity limited accounts of processing difficulty, regardless of whether capacity constraints are thought to be uniform or subject to individual variation. The problem is that computational resources are often confounded with knowledge. It has been established that increases in levels of skill are accompanied by increases in domain specific working memory capacity (see Ericsson & Kintsch, 1995). It has also been shown that tasks which initially appear to exceed computational resources consume less effort and time with practice (see Spelke, Hirst & Neisser, 1976 and Logan 1988). The positive effect of practice on performance is so large that Lachman, Lachman and Butterfield (1979:206) note that "it is beginning to look as if the notion of limited resources is invalid." Unless due precautions are taken, there is a danger of labelling a lack of skill as a lack of computational resources. It appears that one celebrated comprehension phenomenon widely believed to reflect limitations in computational resources might actually arise from a simple lack of linguistic knowledge. The difficulty of comprehending centre embedded sentences has long been the paradigm illustration of the resource limited classical language user (see Miller and Isard, 1964 and a brief but recent discussion in Caplan and Waters, in press). However, Blumenthal (1966) found that this difficulty was attributable to a lack of grammatical knowledge. Stolz (1967) replicated and extended Blumenthal's findings and concluded that, "... S[ubject]s who did not display productivity (i.e. generativity) were hampered not so much by processing difficulties so much as they just could not relate the sentences to their knowledge of English." Significantly, Stolz found that his subjects' performance improved with training. This finding has been replicated by Powers & Peters (1973); Blaubergs & Braine (1974) and Roth (1982). There are, in fact, a very large number of studies which show positive effects of training on native speakers' syntactic skills (cited below under *Constancy*). Resource limitations cannot therefore be invoked to account for processing difficulties before it has been shown that such difficulties are not due to a lack of grammatical skill.

Autonomy. It is often assumed that linguistic competence is independent of semantic factors. Thus native users of English are supposed to be able to parse 'Colourless green ideas sleep furiously" purely on the basis of syntax (Chomsky, 1957:15 but see Hill, 1961). However, Moore and Carling (1982) changed the choice of words but retained the syntactic structure of Chomsky's example to produce "Antepenultimate idiosyncratic elocution paragraphs bright" which sounds highly ungrammatical and may even be unparsable to some native speakers. According to Fodor and Pylyshyn (1988:24), "subject, object and all the rest are Classically defined with respect to the geometry of constituent structure trees." However, consider that in "The boy reads well", 'the boy' is assigned the subject role, yet in "The book reads well", 'the book' is assigned object role. In the second sentence, grammatical roles are assigned with respect to lexical choice rather than constituent structure. In needs to be emphasised that, despite the widespread use of formal grammars in sentence processing research, no complete syntactically autonomous grammar of any natural language has ever been compiled (see Gross, 1979). Instead, there has been a progressive incorporation of semantic factors into grammar over the years in the form of elaborated lexical entries (see Trask, 1993:159). Positive evidence for the important role of lexical information in sentence processing is reviewed in Tanenhaus & Trueswell (1995); see also Bever, Sanz and Townsend (1998). Additionally, Bates and Goodman (1997) offer considerable evidence in favour of the argument that grammar and the lexicon are inseparable in language acquisition, aphasia and language comprehension. It might be noted, in passing, that since individual language users undoubtedly differ in lexical knowledge, lexically based linguistic and psycholinguistic theories open the logical possibility of individual differences in linguistic competence. Experimental evidence for this possibility is offered by Shapiro, Nagel & Levine (1993) and Pearlmutter & MacDonald (1995); a theory of lexically based individual differences in linguistic knowledge is developed in MacDonald & Christiansen (in press). Despite this evidence for a close relationship between sense and structure, however, syntactic autonomy is often treated as a methodological assumption. Performance on sentences of a given structural description is often generalised to performance on all sentences with the same structural description, the assumption being that parsing is sensitive primarily to constituent structure. To give a key example, the mechanics of the garden path effect are often described with respect to the constituent structure of reduced relative clause sentences (eg. Frazier and Clifton, 1996). However, while "The horse raced past the barn fell" might produce a reliable garden path effect in naive subjects, "The landmine buried in the sand exploded" does not (see Tanenhaus & Trueswell, 1995 and Bever, Sanz & Townsend, 1998). Lexical choice seems to matter in the comprehension of centre embedded sentences as well. Contrast "The man whom the farmer whom the girl saw sued died" with "The fact that the man who Andrew looked up to was a criminal bothered Sarah." (see Baird and Koslick, 1970; Hakes et al, 1976; Stabler, 1994; Lewis, 1996 and Hudson, 1997 for a discussion of factors other than centre embedding *per se* which make the first example difficult to understand).

Automaticity. It has also been assumed that syntactic processing is a purely automatic process. According to the 'immediate linguistic awareness hypothesis': "The initial part of any sentence comprehension consists of deriving a complete underlying representation of the sentence" (Levelt, 1976:49). This idea is captured in a statement by Merril Garret that, "What you have to remember about parsing is that basically it's a reflex". Fodor (1983) says that this statement is "the deepest remark that I have yet heard about the psychological mechanisms that mediate the perception of speech" and considers his modularity hypothesis to be "a sustained meditation on Merril's insight ." According to Frazier (1998:134): "Fodor's (1983) important modularity hypothesis has shaped the way many psycholinguists think about the human language capacity". However, while the notion of automaticity plays a key role in psycholinguistics, reference is seldom made to work actually done in automatic processing research. For instance, Fodor (1983) associates automaticity with, among other properties, innateness, domain specificity, autonomy, speed, mandatoriness and unavailability to conscious awareness. However, automaticity is generally regarded in automatic processing research as an effect of practice, which is "acquired only in consistent task environments, as when stimuli are mapped consistently onto the same responses throughout practice" (Logan, 1988:492). On this view, automaticity is not simply domain specific but task specific. It is predicted on this view that skilled individuals will be more likely to display automaticity than less skilled individuals but only with respect to specific tasks within the domain of skill. For instance, while chess experts display superior memory span than chess novices for familiar chess positions, experts and novices perform equally when chess positions are randomised (Chase and Simon, 1973). This pattern of between and within subject differences has been reported in a wide variety of domains, such as bridge, go, music, electronics, computer programming, dance, basketball, field hockey and figure skating (see review in Ericsson and Kintch, 1995). An easy way to test the modularity hypothesis would therefore be to look for between and within subject differences by broadening the range of subjects and materials. The use in parsing research of a single population (university students) and a restricted range of materials (largely attachment ambiguities according to Mitchell, 1994) has discouraged such differences from revealing themselves. In the few cases where individual differences and materials have been controlled for, between and within subject differences have been reported (Cupples & Holmes, 1987; Shapiro, Nagel & Levine, 1993 and Pearlmutter & MacDonald, 1995). Secondly, the properties of speed and autonomy appear to be features of expert rather than novice performance. Chi, Feltovich & Glaser (1981) for instance, found that physics experts solved physics problems four times as fast as physics novices. The effect of practice on processing speed happens to be one of the most stable findings in cognitive psychology. It has been captured mathematically in the Power Law of Practice, whereby the time taken to perform a task decreases in relation to the number of times the task is performed according to the power function (see Eysenck and Keane, 1995:385). With regards to autonomy, Chi, Feltovich & Glaser (1981) found that novices analysed physics problems in terms of spurious surface features, whereas experts analysed the same problems in terms of deeper underlying principles. A similar finding has been reported in language comprehension, whereby relatively unskilled readers are affected by task-irrelevant information which skilled readers are quick to discard (Gernsbacher, 1990; Meiran, 1996). However, much of the data in parsing research is obtained from highly skilled language users and individual differences in processing speed and their implications for the modularity hypothesis have not been explored. The notions of mandoriness and unavailability to conscious awareness also need to be evaluated in the light of actual findings in automatic processing research. According to Neuman (1984:270): automatic processing is not an "invariant consequence of stimulation, independent of a subject's intentions". For example, it has been found that the Stroop effect - a paradigm example of mandatory processing - is subject to task demands (Logan, 1980 and Henik, Friedrich & Kelogg, 1983 and Bauer & Besner, 1997). Even such low level perceptual processes as phonetic labelling are reportedly subject to attentional influence (Gordon, Eberhardt & Rueckel, 1993). Baars (1988) even cites studies which report that biofeedback training can result in temporary control of autonomic functions such as heart rate, skin conductivity and peristalsis and even the behaviour of single neurons. Biofeedback training is reportedly "so ubiquitous a phenomenon that there seems to be no form of CNS activity (single-unit, evoked-potential, or EEG) or part of the brain that is immune to it." (Buchwald, 1974 cited in Baars, 1988). Coming back to language processing, Remez (1994) reports that task demands can determine whether artificially generated speech is perceived as random noise or as speech. There is also evidence that task demands can determine whether or not syntactic analysis is carried out (Mistler-Lachman, 1972; Forster & Olbrei, 1973; Holmes, 1979; Cupples & Holmes, 1987; Foertch & Gernsbacher, 1994; see also Reder & Cleeremans, 1990; Sanford & Garrod, 1994 and Bever, Sanz & Townsend, 1998). Thus another way of testing the modularity hypothesis is to see what effect different task demands have on syntactic processing. If parsing is truly a reflex, then it should always occur regardless of the nature of the experimental task.

Constancy. Finally, it is often assumed that parsing does not adapt to experience. According to Fodor (1998:286):

> *Rightly or wrongly, modern psycholinguistics has been largely guided by the working assumption, from Chomsky (1965), that nothing is learned unless there is reason to believe that it is [...]. To say that perceivers have to learn to parse sentences is thus a concession to be made only under duress, when the empirical evidence insists.*

In line with this assumption, many models of parsing stipulate inflexible parsing operations based on grammars developed in diachronically oriented linguistics. Such models imply that inputs are analysed in exactly the same way on each occasion in much the same way that the same data is analysed in the same way on each occasion by a computer program. There is an obvious problem with such inflexibility. Languages change. Bybee and Thompson (1997) argue that frequency is involved in shaping linguistic units and Boyland (1996) presents evidence that syntactic change can be speeded up experimentally. However, the facts of synchronic linguistics do not seem to have any bearing on mainstream parsing research. The fact that languages do change, however, implies a degree of plasticity in the linguistic representational system. Strong evidence for this plasticity is provided by studies which report that training can enhance the syntactic performance of normal adult native speakers (Stolz, 1967; Powers & Peters, 1973; Blaubergs & Braine, 1974; Stewart, 1978*a*; Stewart, 1978*b* and Swan, 1979); normal child native speakers (Elardo, 1972; Weaver & Ruder, 1980; Roth, 1982; Van Wijk & Kempen, 1983; Neville & Searles, 1985; Richgels, 1987; Kren et al, 1989 and Moerk, 1991); native speaker children with specific language impairment (reviewed in Leonard, 1981; see also Weller, 1979; Connell, 1987; and Gail & Dodd, 1995) and agrammatic adult native speakers (see the special issue of *Brain and Language* edited by Martin, 1996). The relationship between education and grammatical skill is another form of evidence for environmentally induced changes in the linguistic representations in normal adult native speakers (Geer, Gleitman & Gleitman, 1971; Cox, 1976; Baruzzi, 1982 and Karanth, Kudva & Arpana, 1996; Dabrowska, 1997). Despite such findings, the idea that parsing is constant from one parse to the next has, until recently, been taken as a methodological assumption. Results are often averaged over items. While this averaging is necessary for other reasons, it means that one is unlikely to observe any learning which might take place during an experiment unless one is specifically looking for it. It is now widely accepted that sentence comprehension is sensitive to frequency in the form of priming (Bock & Loebell, 1990 and Branigan, Pickering & Stewart, submitted) and knowledge of lexical probabilities (Pearlmutter & MacDonald, 1995). If grammatical representations are sensitive to patterns of usage, then they may be expected to differ *between* individuals exposed to different patterns of usage and also *within* individuals at different points in time. Experimental practice so far has been such as to discourage the discovery of such differences.

The foregoing paragraphs have attempted to show that the classical language user relies more on its theoretical appeal than on any empirical support. The model's empirical weaknesses have been hidden from view by investigative practices which shield it from direct scrutiny. The strong theoretical appeal of the model seems to be related to the information processing tradition in cognitive psychology. Working within this tradition, Miller, Galanter & Pribram (1960) and Fodor & Pylyshyn (1988) have advocated a computer metaphor of mind (see also Lachman, Lachman & Butterfield, 1979). It appears that each of the five properties of the classical language user is related to a key property of the digital computer. The distinction between competence and performance corresponds to the distinction between software and hardware. Computer languages are typically generative and context-free. Finally, the fixed nature of algorithms is closely related to the assumptions of automaticity and constancy. The classical language user is essentially a digital computer. According to Fodor & Pylyshyn (1988:62): "... the claim that the mind has the architecture of a Classical computer is not a metaphor but a literal empirical hypothesis". The foregoing review has shown that researchers in sentence processing have often treated the classical language user not as a hypothesis but as an established set of facts. Negative evidence against each assumption of the classical model has been presented. The assumption that computational resources are severely limited has been also shown to be problematic. One reason why classical language user has managed to withstand so much negative evidence for so long may be that this evidence often relates only to one assumption of the model or another but not to all assumptions together. It is not as difficult, however, to design an experiment which addresses all these assumptions as it might first seem. Competence and generativity can be tested but seeing if between and within subject differences remain when performance constraints are neutralised. Autonomy can be tested by seeing if sentences with the same structure but different lexical items are parsed equally well. Automaticity can be tested by seeing if subjects always carry out complete syntactic analyses of test sentences. Finally, constancy can be tested by seeing if there are any order effects during an experiment. An experiment is described below which was designed to address these issues by asking if highly educated native and non-native speakers of English can outperform less educated native speakers of speakers of English in a test of grammatical skill.

EXPERIMENT

Hypothesis. The experiment was designed to test the hypothesis that graduate native and non-native speakers of English comprehend grammatically challenging English sentences more accurately than non-graduate native speakers of English. Comprehension skill was defined as the ability to assign thematic roles to sentence constituents. Comprehension accuracy was measured by the number of correct responses to questions of the type *who* did *what* to *whom*? It was predicted that non-graduate natives would obtain the least scores.

Subjects. Three groups of subjects were tested. Group 1 consisted of 12 graduate native speakers of English. 11 members of this group were postgraduate students in Applied Linguistics and 1 member was a lecturer in Linguistics. Group 2 consisted of 12 postgraduate non-native speakers of English. 11 members of this group were also postgraduate students in Applied Linguistics and 1 member was a lecturer in Linguistics. Members of Group 2 had learnt English as a second or foreign language at school and some of them had taught English as a second or foreign language professionally. Group 3 consisted of 12 native speakers of English whose formal education did not extend beyond high school. At the time of the experiment, all members of Group 3 were employed as cleaners, gardeners or porters. Groups 1 and 2 volunteered to participate without pay and 10 members of Group 3 accepted £5.00 each for their participation.

Materials. Grammatically novel and complex structures were used. Test sentences were Complex Noun Phrase (CNP), Tough Movement (TM) and two types of Parasitic Gap constructions (PG1 and PG2). An earlier study by Dabrowska (1997) had shown education-related differences between native speakers in the ability to parse these sentence types. Subjects had to answer 5 questions on each test sentence. With minor exceptions, questions for all construction types were ranked in difficulty as follows. The key question was the most difficult question to answer and was considered to be diagnostic of correct parsing. The backup question was less difficult than the key question and was used to compensate for the possibility of guessing on the key question. The third type of question had two possible answers and was used to test subjects' awareness of structural ambiguities. The fourth type of question was a give-away question and the fifth type of question was a grammaticality judgement. Data from grammaticality judgements was not used to make inferences about parsing ability. Test sentences, questions and answers are listed in the Appendix.

Design. Three sentences from each structural type were created (=12 base sentences). Semantically plausible (SP); semantically neutral (SN) and semantically implausible (SI) versions of each base sentence were then created (=36 test sentences). Examples of each structural type in all three Plausibility conditions are shown in table 1.

table 1

ST	PL	Example Sentence
TM	SN	Tim will be hard to get Mary to give a loan to.
	SI	The bank manager will be hard to get the convict to give a loan to.
	SP	The convict will be hard to get the bank manager to give a loan to.
CNP	SN	Tom knows that the fact that taking good care of himself is essential surprises Peter.
	SI	Tom knows that the fact that taking good care of himself is essential surprises the doctor.
	SP	The doctor knows that the fact that taking good care of himself is essential surprises Tom.
PG1	SN	The girl who Paul saw after discovering Alex proposing to dismiss had lunch in a cafe.
	SI	The lady who Paul saw after discovering the servant proposing to dismiss had lunch in a cafe.
	SP	The servant who Paul saw after discovering the lady proposing to dismiss had lunch in a cafe.
PG2	SN	The man who Peter saw before overhearing his girlfriend proposing to jilt walked away.
	SI	The counsellor who Peter saw before overhearing his girlfriend proposing to jilt walked away.
	SP	The man who the counsellor saw before overhearing his girlfriend proposing to jilt walked away.

Each group of subjects was divided into 3 lots (= 4 subjects each) and the experiment was also divided into 3 blocks (= 4 sentences each). Within each block, all 3 lots of subjects were exposed to all 4 structural types and all 3 Plausibility conditions. Over the course of the experiment, each lot of subjects was exposed to each of the 4 structures in each of the 3 Plausibility conditions (see table 2). The sequence of conditions for each lot was rotated over blocks to prevent any possible comprehension strategies based on awareness of conditions. The sequence of questions for each sentence was also varied over blocks to prevent any possible comprehension strategies involving questions. Because only one instance of each construction type was represented in each block, the effects of individual sentences were confounded with any possible effects of sentence order (item/order confound) making it impossible to test the constancy hypothesis directly. There were three filler sentences, one presented the start of the experiment, another in the middle of block 2 and another at the end of the experiment.

table 2

Block	*Sentence*	*Structure*	*Neutral*	*Implausible*	*Plausible*
1	1	TM	Lot 1	Lot 2	Lot 3
		PG1	Lot 2	Lot 3	Lot 1
		PG2	Lot 3	Lot 1	Lot 2
		CNP	Lot 1	Lot 2	Lot 3
2	2	TM	Lot 2	Lot 3	Lot 1
		PG1	Lot 3	Lot 1	Lot 2
		PG2	Lot 1	Lot 2	Lot 3
		CNP	Lot 2	Lot 3	Lot 1
3	3	TM	Lot 3	Lot 1	Lot 2
		PG1	Lot 1	Lot 2	Lot 3
		PG2	Lot 2	Lot 3	Lot 1
		CNP	Lot 3	Lot 1	Lot 2

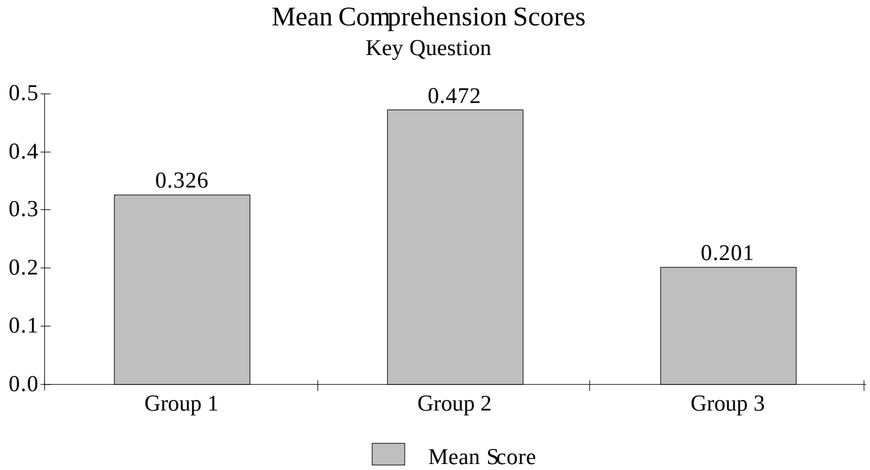
Procedure. Test sentences were presented on a computer screen using self paced presentation. An introductory screen explained the following procedure. Subjects were to read the introductory message and to ask the experimenter to explain anything they did not understand. Having read the instructions, subjects were to click on a mouse button in order to display each test sentence on the screen one word at a time. Each word would remain visible on the screen as the sentence unfolded. After studying the sentence to their satisfaction, subjects were to click the mouse button again and a question would appear below the sentence, which remained visible. The questions had to be answered verbally since responses were recorded on tape. Clicking the mouse button again would cause the first question to disappear and be replaced by another one, and so on, until all five questions had been answered. Clicking the mouse button after the fifth question had appeared allowed subjects to move on the next sentence. This procedure was designed with the following considerations in mind. The self-pace mode of presentation was needed to obtain reading time data. Words were allowed to remain visible on screen as the sentence unfolded in order to eliminate short term memory as a performance factor. This mode of presentation may have influenced the reading time patterns. However, the elimination of memory load had a higher priority than the reading time data, which was not intended for use in testing hypotheses. Finally, questions were presented one at a time to prevent participants from developing comprehension strategies based on answers to previous questions.

RESULTS

The comprehension questions yielded 60 responses from each subject (5 questions per sentence x 12 sentences). Comprehension scores were subjected to a 5 way ANOVA (3 Groups of 12 subjects each x 3 Lots of 4 subjects each x 4 Structures x 3 Plausibility Conditions (alternatively, 3 Sentence Conditions) x 5 Questions.) Three kinds of response time data were obtained: 1) word reading time (WRT); 2) sentence response time, i.e. time spent studying sentence after it was fully displayed on the screen and before answering any questions (SRT) and 3) question response time (QRT). Outliers were removed by replacing values lying below or above one and a half times the interquartile range with the value of the first and third quartiles respectively. However, a positive skew remained. In the absence of any other remedy, it was decided to go ahead with a 5 way ANOVA (3 Groups of 12 subjects each x 3 Lots of 4 subjects each x 4 Structures x 3 Plausibility Conditions x 7 Response Times (= 1 mean word reading time + 1 sentence response time + 5 question response times)). Reading time profiles were analysed in terms of the effects word position (Positions). Results from the analysis are presented in five sections corresponding to the five assumptions of the classical model. With some qualifications, a main effect of a given variable on comprehension scores constitutes negative evidence for the corresponding assumption. Uniform Competence is negated by an effect of Groups; Generativity by an effect of Structures; Autonomy by an effect of Plausibility; Automaticity by an effect of Questions and Constancy by an effect of Sentences. Presentation of comprehension data is restricted to data from the key questions unless otherwise stated.

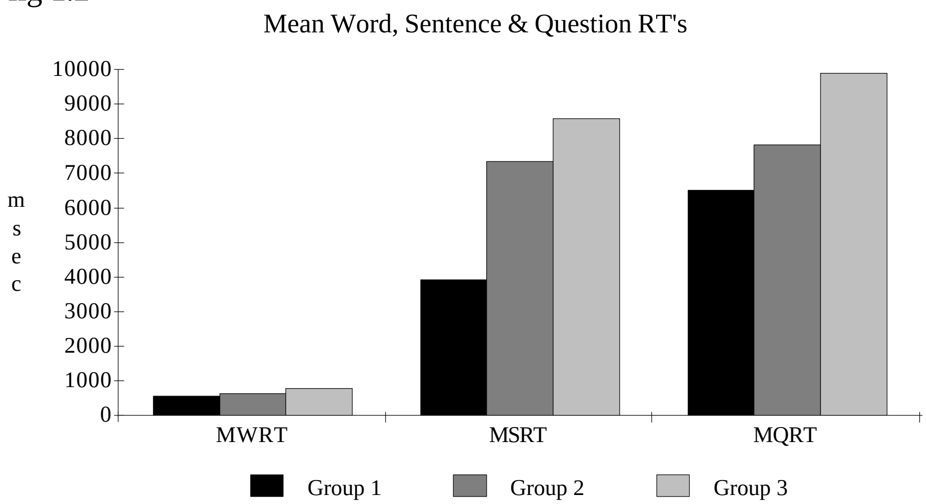
Competence. A main effect of Groups was found for the key comprehension questions, $F_{2,27} = 5.8515$, $p<0.0077$ (fig. 1.1). A pairwise test reveals that, for the key questions, only the difference between Group 2 (non-native graduates) and Group 3 (non-graduate natives) is significant, $F_{1,18} = 16.2384$, $p<0.0008$ There is a significant difference between Group 1 and Group 3 only when data from all four comprehension questions is analysed, $F_{1,18} = 10.2979$, $p<0.0049$. There is also a weak interaction between Groups 1 and 2, Structures, Verbs and Questions, wherein Group 1 performs less consistently than Group 2, $F_{18,324} = 1.6545$, $p<0.0461$. All groups performed equally well on the filler items, obtaining maximum or near maximum scores.

fig. 1.1



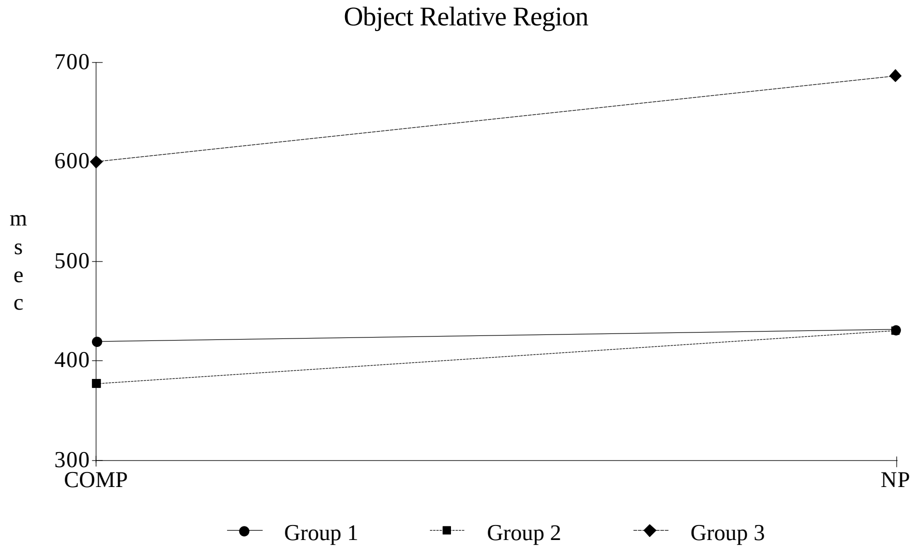
Groups had no significant effect on reading time per word, $F_{2,27}$ = 2.3232, p<0.1172, but it had significant effects on sentence response time $F_{2,27}$ = 5.6775, p<0.0087, and question response time, $F_{2,27}$ = 6.9336, p<0.0037 (fig 1.2).

fig 1.2



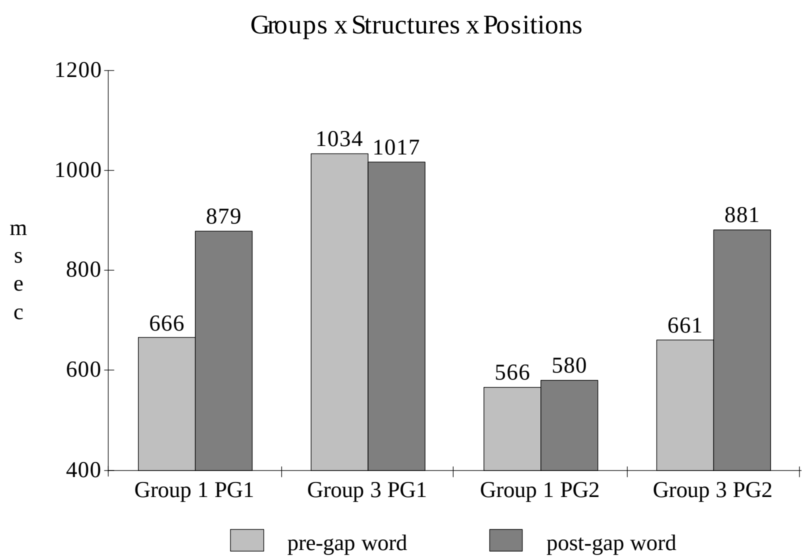
There was an interaction between Group 1 and Group 3 on reading times for a specific region of the parasitic gap sentences. These sentences were all object relative structures, which began as follows: "The man who Peter met ...". Reading times for the complementiser (COMP = who) and the noun phrase (NP = Peter), interact with Groups 1 and 2 significantly, $F_{1,18}$ = 8.3204, p<0.0099 (fig. 1.3). The effect of Position is not significant for Group 1, $F_{1,9}$ = 1.3663, p<0.2725 or for Group 2, $F_{1,9}$ = 3.7081, p<0.0863 but is significant for Group 3, $F_{1,9}$ = 13.131, p<0.0055.

fig 1.3



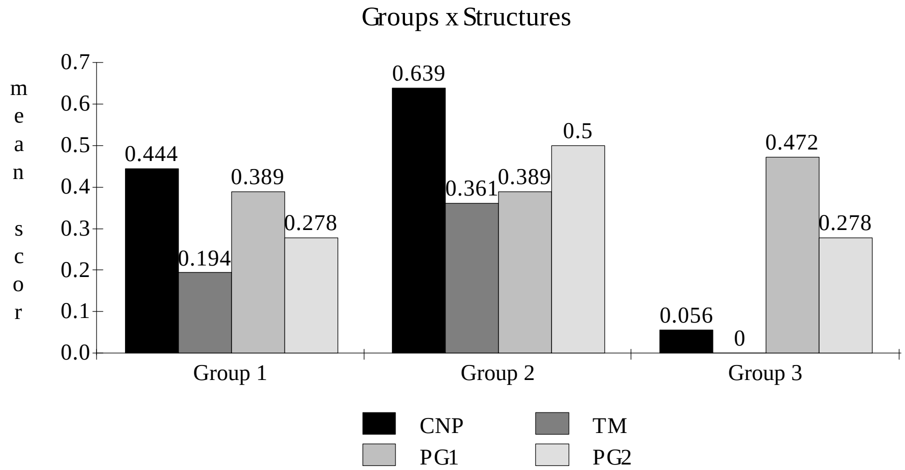
There was also a group difference with regard to the gap effect (i.e the second or 'parasitic gap'). When the difference in reading times between pre- and post-gap word positions was analysed, a weak interaction was found between Groups, Structures and Positions, $F_{1,18}$ = 4.5649, p<0.0466. For Group 1, the gap effect is confined to PG1 but for Group 3 it is confined to the PG2 structure (fig. 1.4.)

fig. 1.4



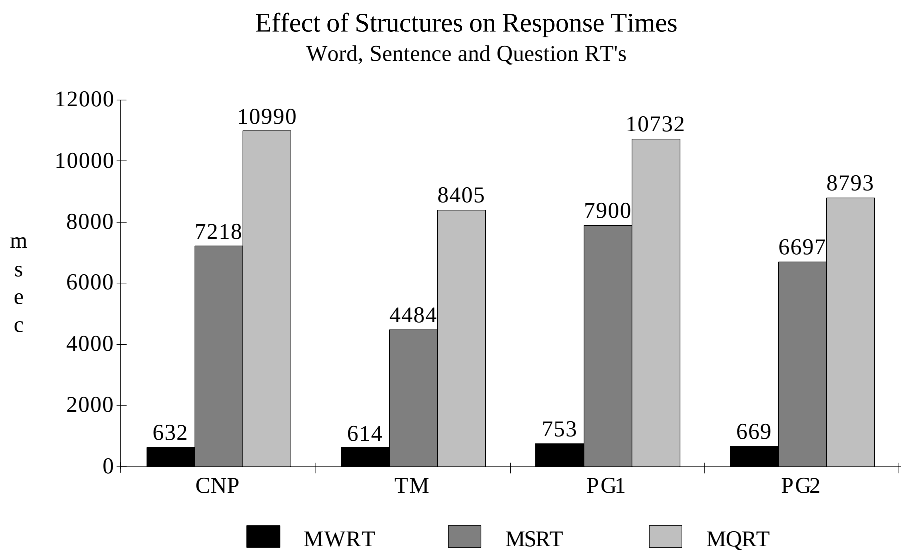
Generativity. There was a main effect of Structures on scores, $F_{3,81}$ = 6.0496, p<0.0009 and a significant interaction between Groups and Structures, $F_{6,81}$ = 4,2521, p<0.0009. A *post hoc* test reveals no significant effect of structures for Group 2, $F_{3,27}$ = 1.7885, p< 0.1731; a marginally significant effect for group 1, $F_{3,27}$ = 3, p<0.0480 and a highly significant effect for Group 3, $F_{3,27}$ = 19.0696, p<0.0000 (fig 2.1).

fig. 2.1



There was a main effect of Structures on Response Times, $F_{3,81} = 13.4192$, $p<0.0000$.

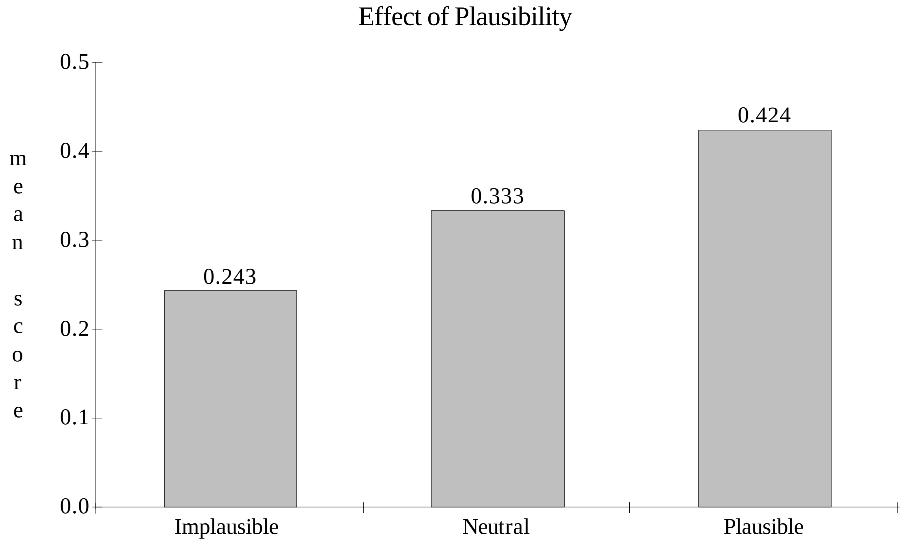
fig. 2.2



(NB. although PG1 and PG2 are structurally identical, PG2 elicited significantly faster response times, $F_{1,27} = 12.7305$, $p<0.0014$.)

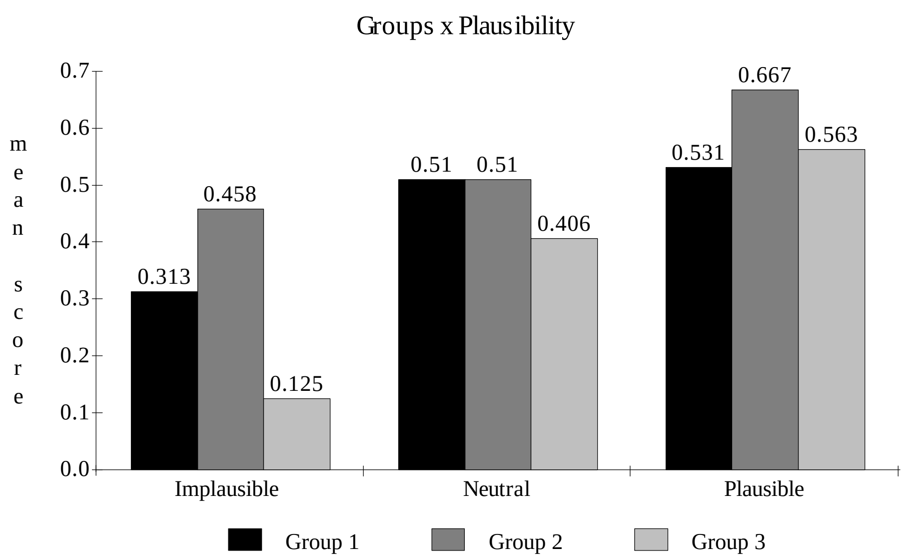
Autonomy. A main effect of Plausibility on scores was found, $F_{2,54} = 9.5063$, $p<0.0003$. fig. 3.1 shows mean scores for the three Plausibility conditions collapsed over Structures and Groups.

fig. 3.1



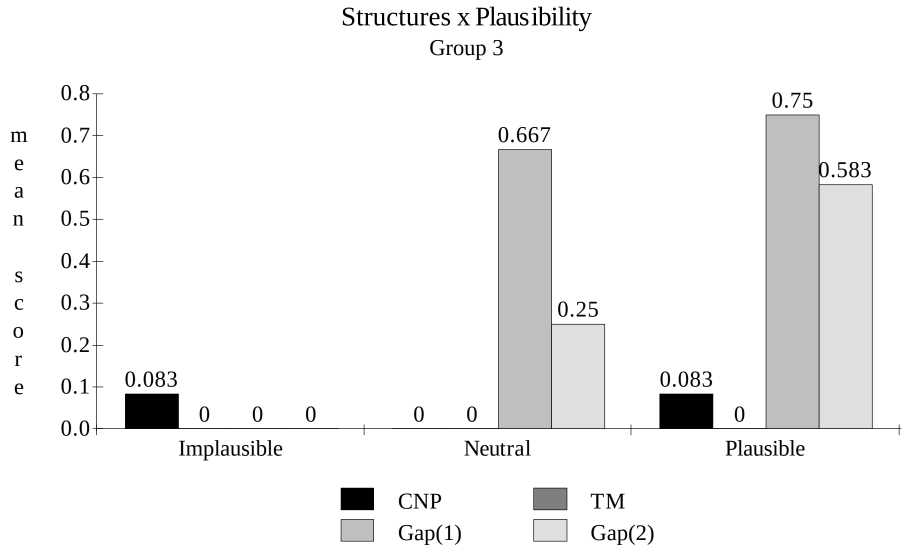
There was a non significant interaction between Plausibility and Groups for the key question, $F_{4,54}$ = 2.4188, p<0.0596. However, this interaction is significant when scores for both questions 1 and 2 (the key question and the second most difficult question, respectively) were analysed, $F_{4,54}$ = 4.2906, p<0.0044 (fig. 3.2). There was no significant effect of Plausibility for Group 1, $F_{2,18}$ = 1.3676, p<0.2799 or for Group 2, $F_{2,18}$ = 1.5, p<0.2497 whereas the effect of Plausibility on Group 3 was highly significant, $F_{2,18}$ = 16.333, p<0.0001.

fig. 3.2

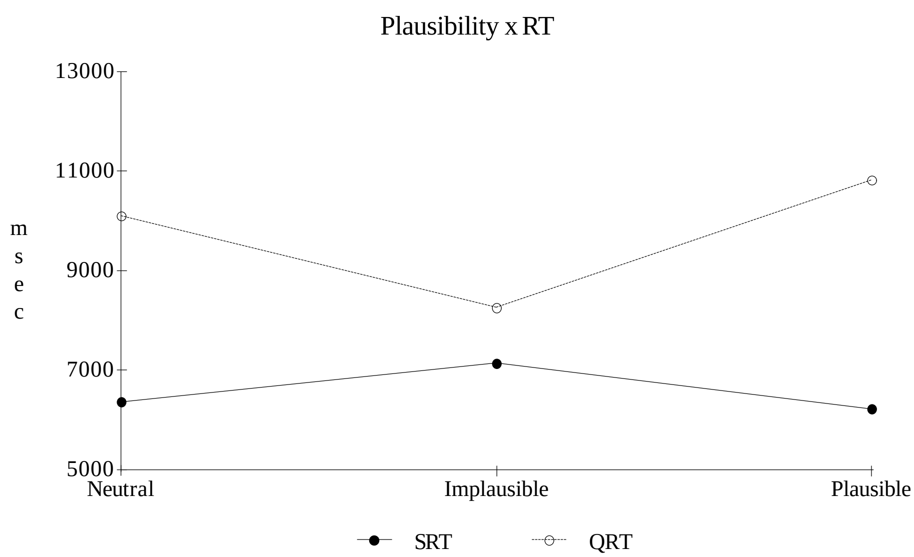


There was an interaction between Structures and Plausibility (key questions only), $F_{6,162}$ = 2.7784, p<0.0134 and a three-way interaction between Groups, Structures and Plausibility, $F_{12,162}$ = 2.3011, p<0.0098. A *post hoc* test shows that the interaction between Structures and Plausibility is not significant for Group 1, $F_{6,54}$ = 1.2917, p<0.2769 or for Group 2, $F_{6,54}$ = 1.1695, p<0.3363 but that it is highly significant for Group 3, $F_{6,54}$ = 9.7895, p<0.0000 (fig 3.3).

fig 3.3

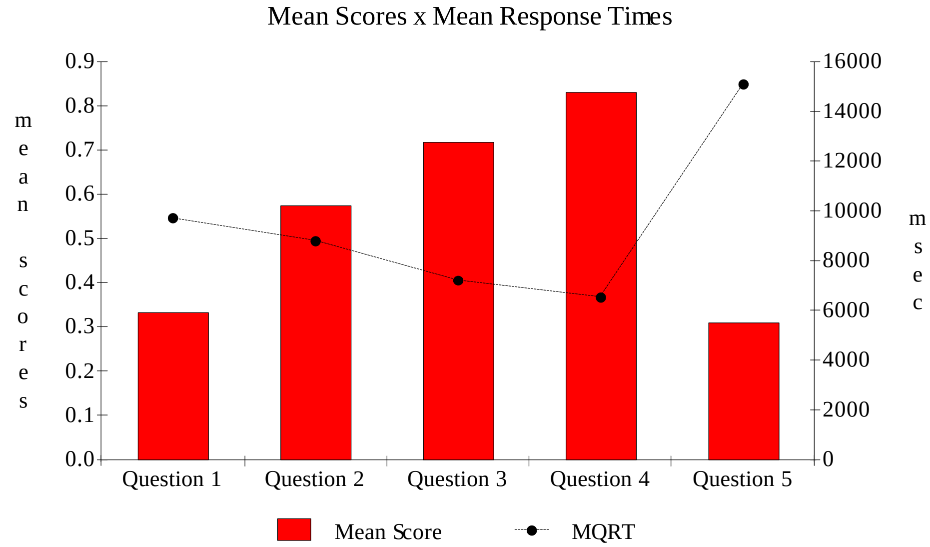


There was no main effect of Plausibility on Response Times. However, there was an interaction between Plausibility and sentence response times vs. question response times for question 1, $F_{4,108} = 6.0013$, p<0.0002 (fig. 3.4).



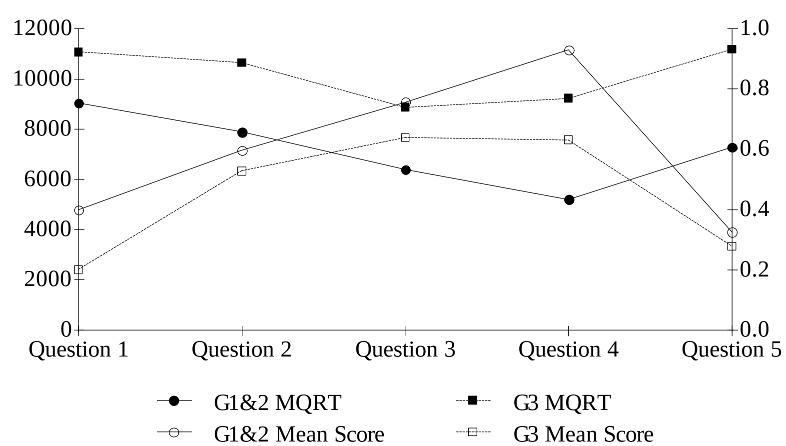
Automaticity. There was a main effect of Questions on scores, $F_{4,108} = 86.176$, p<0.0000. It may be recalled that question 1 is the most difficult or key question; question 2 is the second most difficult question; question 3 is the third most difficult; question 4 is the give-away question and question 5 is the grammaticality judgement. The level of significance is not affected if the grammaticality judgement (question 5) is excluded from the analysis (the scores for the key question and the grammaticality judgement are virtually identical). There was also a main effect of Questions on question response times, $F_{4,108} = 36.5672$, p<0.0000. Yet again, the level of significance is not affected if response times to the grammaticality judgement are excluded from the analysis. Fig. 4.1 shows that there was an inverse relationship between scores and response times. The same graph is obtained if only the response times for correct responses are plotted against scores.

fig. 4.1



There was an interaction between Groups and Questions for the comprehension data, $F_{8,108}$ = 2.6803, p<0.010 as well as for the Question Response Time, $F_{8,108}$ = 30.6911, p<0.0000. The interaction is between Group 3 on the one hand, and Groups 1 and 2 on the other hand, and Questions 3 and 4 (fig. 4.2. NB. Reading time values are shown on the left and comprehension scores on the left. The mean response time value for Group 3 Question 5 is actually 30741 msec and was adjusted downwards for the purpose of making it fit into the same graph with the other values).

fig. 4.2

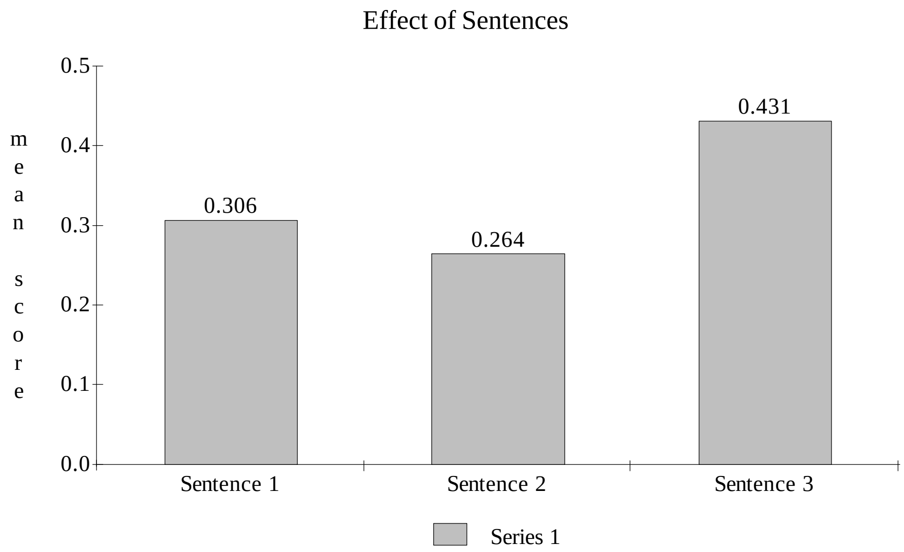


It may be recalled that Question 4 is the give-away question. It appears from fig. 4.2, however, that Group 3 found Question 4 just as difficult as Question 3 and spent equal amounts of time answering both questions.

Constancy. It may be recalled that there were three sentences for each of the four structures. Data was analysed to see if there were any differences between the first, second and third sentence of each structure. It should be noted, however, that the item/order confound described in the Design section requires a cautious interpretation of any seeming effects of sentence order. A main effect of Sentences was found, $F_{2,54}$ = 9.298, p<0.0003. fig.

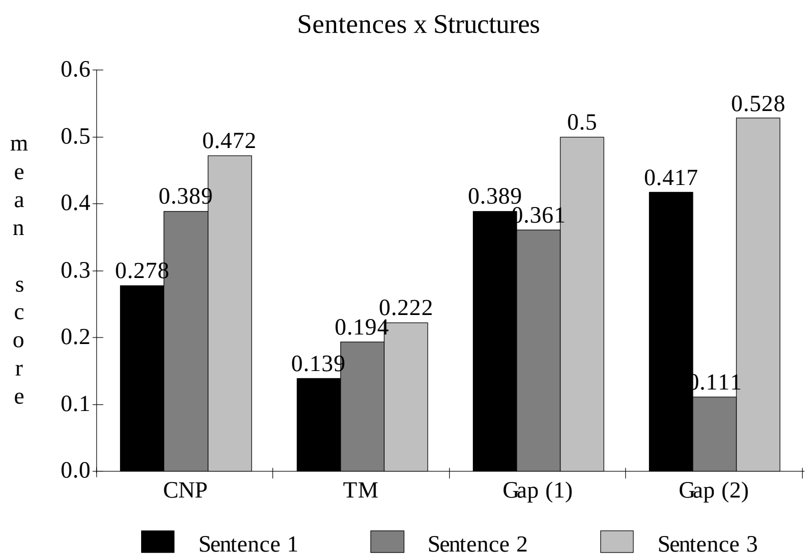
5.1 shows mean scores for sentences collapsed over Plausibility, Structures and Groups.

fig 5.1



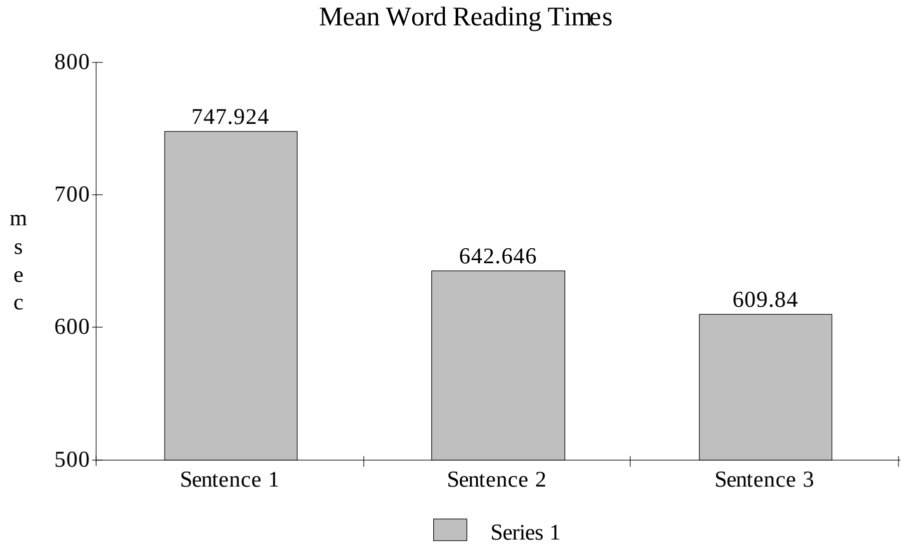
There was an interaction between Sentences and Structures, $F_{6,162} = 2.88$, $p<0.0107$ (fig. 5.2)

fig 5.2



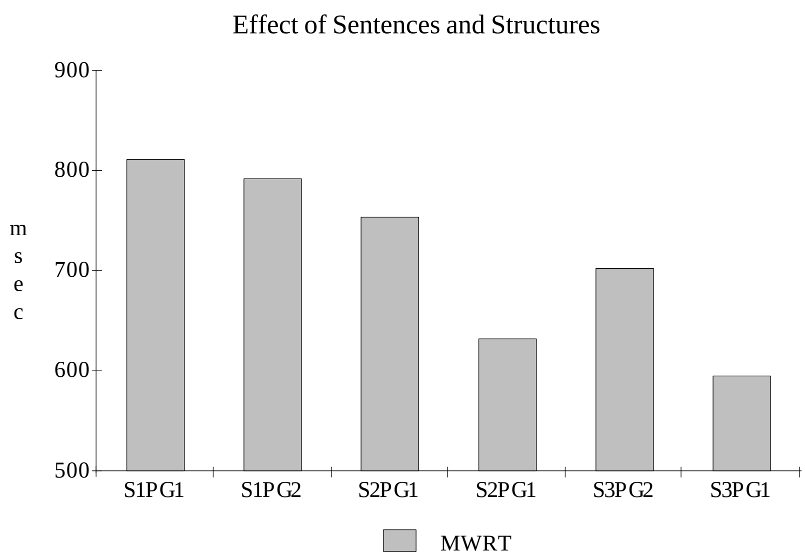
For all structures, performance on the third sentence is higher than performance on the first sentence. The markedly depressed performance on sentence 2 of PG2 appears to be an item effect. More suggestive evidence for order effects comes from the response time data. There was a main effect of Sentences on mean word reading time, $F_{2,54} = 13.9555$, $p<0.0000$ but not on sentence response time or question response time. Fig. 5.3 shows mean word reading times collapsed over Plausibility, Structures and Groups.

fig. 5.3



It may be recalled from section 1.2 (Generativity) that mean word, sentence and question response times were significantly less for the PG2 sentences compared to the PG1 sentences. Since the two types of sentences are structurally identical, and each PG2 sentence was presented after a PG1 sentence, the facilitation of PG2 sentences is a possible order effect. Fig. 5.6 shows that there is a jagged pattern of decrease in reading times for the last four PG1 and PG2 sentences.

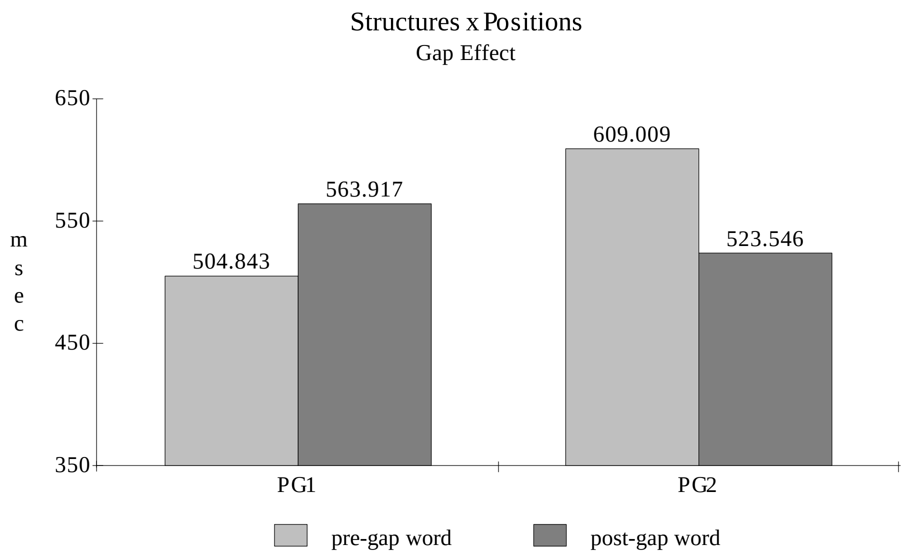
fig. 5.4



The effect of Sentences (i.e. Sentence 1, 2 and 3 of both PG1 and PG2) is significant, $F_{2,54}$ = 7.3981, p<0.0014 and the effect of Structures (i.e. PG1 vs. PG2) is also significant, $F_{1.27}$ = 10.2902, p<0.0034.

There also appear to be order effects on the presence of the gap effect. PG1 and PG2 sentences had two gap positions: The *girl who Peter saw* $_{(gap\ 1)}$ *after discovering Alex proposing to dismiss* $_{(gap\ 2)}$ *had lunch in a cafe*. There is no main effect of Positions (i.e. differences between pre-gap and post gap word reading times) for the first gap, $F_{1,27}$ = 0.2613, p<0.6134, but there is a significant effect for the second gap, $F_{1,27}$ = 6.8733, p<0.0142. However, there is an interaction between Structures and Positions for the first gap, $F_{1,27}$ = 10.7967, p<0.0028. In other words, there is what appears to be a gap effect in PG1 which is reversed in PG2 (fig. 5.6)

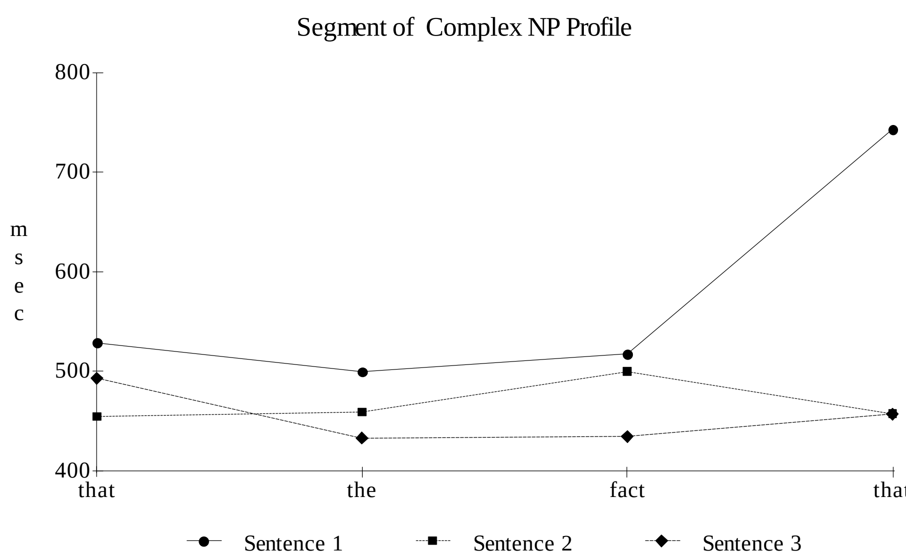
fig *5.6*



For the second gap, however, there is a weak three way interaction between Groups, Structures and Positions, $F_{1,18}$ = 4.5649, p<0.0466. For Group 1, the gap effect is present only in PG1 but for Group 3 it is present only in PG2 (see fig 1.4).

Moving on to the CNP sentences, there is a significant interaction between Sentences and Word Positions in one region of the reading time profile, $F_{6,162}$ = 4.4907, p<0.0003 (fig. 5.7). It looks as if that the reading time peak decreases in magnitude and migrates backwards with each presentation.

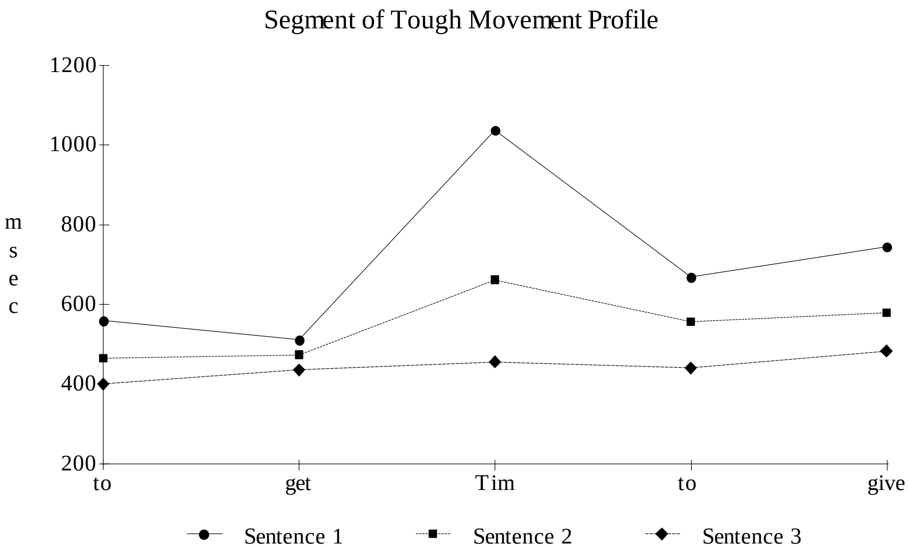
fig 5.7



All CNP sentences began with "NP knows that the fact that.." , so it is unlikely that the observed changes are item specific. Analysis of responses to questions on the CNP sentences (not presented here) suggests that this region is involved in comprehension difficulties.

A similar interaction between Sentences and Word Position was also found for the Tough Movement construction, $F_{8.216} = 5.9208$, p<0.0000. As with the CNP construction, the region shown on the graph is invariant for all TM sentences, except for the choice of noun and the verb, so the change in the profile is also unlikely to be an item effect. Simpler tough movement sentences read, for example "Mary is difficult to see". The peak observed in sentences 1 and 2 therefore occurs at the position where a simple TM sentence might have ended (fig. 5.8).

fig. 5.8



DISCUSSION

The findings are first summarised under separate headings then discussed together. Starting with *Uniform Competence,* between-group differences were found: graduate natives outperformed non-graduate natives. This result echoes Geer, Gleitman and Gleitman’s finding that subjects “from different educational levels differ in grammatical organisation” (1972:354). Surprisingly, non-graduate natives outperformed both native speaker groups. However, the difference between non-native and native graduates was not significant in terms of scores and was only marginally significant in terms of consistency. A number of specific differences between the groups were found. Firstly, with regards to object relative structures (PG1&2) Group 3 but not Group 1 seemed to experience difficulties at the noun phrase immediately following the complementiser, eg. “The man who (=COMP) Peter (=NP) saw (=VERB)...”. In previous studies, difficulty has only been thought to occur at or just after the embedded verb (see Just and Carpenter, 1992). A second difference is that Group 1 and Group 3 interpreted control structures differently. Given "The man who Paul met after discovering ..." Group 1 were more likely to name 'Paul' as the subject of 'discovering', while Group 3 were more likely to name 'The man'. Thirdly, there was an inexplicable interaction between Groups and Structures in the gap effect. *Generativity.* Performance was not consistent across structures. The greatest inconsistencies were found in the performance of native non-graduates followed by native graduates. Only two members of Group 3 answered one key question on the CNP sentences correctly, and none answered the key question on the TM sentences correctly.

These results accord with Sanders' (1972) finding that

> *At least one structure thought to be commonly known to adults was not consistently understood by almost half the subjects in the study. It is not unreasonable to speculate that there may well be other syntactic structures which have been assumed to be part of ordinary adult language, but which most adults may not understand.* (744).

Unexpectedly, the non-native graduates performed most consistently across structures and conditions. *Autonomy*. Group 3 was significantly affected by plausibility and failed almost completely to answer questions on the implausible sentences correctly. This finding conflicts with the claim that

> *Readers and listeners can arrive at an interpretation that violates their beliefs about the message being conveyed, their feelings of plausibility or contextual felicity, and their wishes and expectations. They do not arrive at interpretations that violate the grammar.* (Frazier and Clifton, 1996:3).

The fact that the native and non-native graduates were not affected by plausibility indicates that they were using purely syntactic information to decode sentences in the neutral and implausible conditions. On the other hand, the fact that native non-graduates were affected by plausibility indicates that they were relying on semantic information to decode the test sentences - a strategy which failed them on the neutral and implausible sentences. *Automaticity*. The experiment produced evidence for incomplete syntactic comprehension. Subjects were not able to answer all questions about the same sentence equally well, suggesting that they had only partially understood the test sentences. Even when subjects answered questions correctly, however, they still took longer to answer the difficult questions compared to the easier ones. Response times ranged from 6 to 16 seconds and subjects often made false starts, suggesting that they had to consciously work out the answers. This finding indicates where novel structures are concerned, parsing is not necessarily fast and grammatically controlled as claimed by Frazier (1998:126). *Constancy*. The experiment found what seem to be order effects in comprehension scores and reading times. Comprehension scores increased and reading times decreased on successive presentations of each structural type. However, an item/order confound makes it impossible to be certain. Nevertheless, the changes in reading time profiles are particularly suggestive of order effects since they occur in lexically invariant regions of each structural type. In the existing literature, practice effects largely take the form of a reduction in reading times (eg. Branigan, Pickering and Stewart, submitted). In the present experiment, practice effects seemed to have the effect of actually altering the shape of reading time profiles suggesting changes in patterns of expectancy.

There are a number of possible explanations for the results which are compatible with the classical language user. The conventional interpretation of the results would be that subjects were equally competent but differed in the availability of computational resources. Before discussing this view in more detail, some other explanations which are compatible with the classical language user will be discussed briefly. Firstly, it could be said that all experimental groups carry out complete and accurate syntactic processing, but that there are group differences in the ability to map the syntactic analysis to a semantic interpretation. In the absence of a reliable method for determining a stage of purely syntactic processing, however, this explanation cannot be tested. Another account suggestion might be that the present results do not reflect individual differences in linguistic competence. Rather, the structures used in the study are inherently unparsable to the parsing mechanisms of all native speakers of English in the same way that a finite state automaton is incapable of parsing structures of arbitrary complexity. It could then be argued that all groups performed equally badly in terms of syntactic processing, except that the highly educated groups employed more sophisticated repair strategies. The availability of these strategies would then be the source of the observed individual differences. The challenge for such an account is to provide a principled distinction between normal processing and repair strategies, particularly if the repair strategies use purely syntactic information. A related explanation might be that the structures used in the experiment lie outside the competence of native speakers of English. That is to say that the structures used in the study are ungrammatical. It should be borne in mind that a generative grammar is capable of generating both simple and complex sentences equally well and that it cannot distinguish between familiar and unfamiliar sentences. Issues of complexity or familiarity lie arise outside the grammar. Hence the unnaturalness of the test items is irrelevant to *grammatical competence.* Thus the challenge for this account would be to devise a grammar of English which is generative yet excludes the structures used in the study. Such an explanation would also need to explain the effect of education on the ability to understand the structures used in the experiment. Another explanation along these lines would be to suggest that the structures used in the test are indeed grammatical, but they are unacceptable to native speakers. On this view, subjects fail to answer questions correctly because they refuse to answer questions on sentences which they find unacceptable. However, subjects were able to answer sentences in the plausible condition, so unacceptability cannot be used to explain their failure to answer sentences in the implausible condition. A fifth explanation for the results could be that the differences in performance between Group 1 and Group 3 reflect differences in literacy rather than differences in syntactic competence. However, all groups performed equally well on the filler items, so literacy cannot be held accountable for the group differences on the test items. A fifth explanation for the findings might be made in terms of a distinction between core and periphery grammar (see Shacter, 1996 for a discussion of this distinction). In this account, the individual differences reported here might be said to belong to the periphery rather than to the core grammar. However, it is not clear what the core might consist of if it excludes such a key notion as phrase structure. It is possible to generate endless post hoc explanations for the results which are compatible with the classical language user. What should be borne in mind in considering such explanations is that the classical language user is only a set of hypothesis with no empirical underpinning. The burden of proof therefore lies with the proponents rather than the detractors of the model.

In the current climate of opinion, the overwhelming tendency is to explain the results of the present study in terms of limitations in computational resources. This approach assumes that language users know their language perfectly and that their computational resources are severely limited. Both of these assumptions are highly questionable. It has never been shown that native speakers of a language do, in fact, have complete grammatical mastery of their language. Quite the contrary, past research indicates considerable differences in grammatical knowledge between native speakers of a language. In the present study, a number of such differences were found under conditions in which possible effects of resource limitations were neutralised. Additionally, the group differences observed with respect to object relatives; control structures and gap effects do not yield to any existing resource limitation account. The notion of limited computational resources is, in any case, confounded with knowledge. This confound is evident in two papers which have played a key role in developing the notion of the perfectly competent but resource limited language user - Miller (1956) and Miller & Isard (1964). In his discussion of channel capacity, Miller (1956) reports that subjects have difficulty in discriminating between more than six one-dimensional stimuli. He states, *Most people are surprised that the number is as small as six. Of course, there is evidence that a musically sophisticated person with absolute pitch can identify accurately any one of 50 to 60 different pitches. Fortunately, I do not have the time to discuss these remarkable exceptions. I say it is fortunate because I do not know how to explain their superior performance.* (1956:84).

Here, Miller avoids a key question - the question of whether capacity limitations reflect intrinsic limitations of mental hardware or whether they reflect the limits of acquired knowledge. This question re-emerges in Miller's discussion of immediate memory. Miller reports that the span of immediate memory is restricted to 7 plus or minus 2 items. However, he also reports that, Sidney Smith, a fellow psychologist, increased his own memory span by 400% through hierarchical recoding. Thus, although the title of Miller's paper suggests that capacity limitations are intrinsic, the paper itself provides evidence that these limits are somehow related to domain knowledge. Miller is equivocal on this point: "There seems to be some limitation built into us either by learning or by the design of our nervous systems" (86). It should be noted that, in any case, even in the case of untrained subjects, performance in memory span tests depends on ltm access. The reason for this is that "incoming information must make contact with the long term memory store in order for it to be categorically coded" (Cowan, 1995:113). In other words, stimulus categories such digits, letters and words must first be retrieved from long term memory before they can be held in short term memory. Chi (1976) reviews studies of the development of short term memory capacity and argues that children's stm may be smaller than that of adults because children do not have the same knowledge base in ltm as adults do. For instance, a written word may be a single chunk to an adult, but it may be set of separate items to a child still learning to read.

Notwithstanding the important and unresolved question of the role of knowledge in capacity limitations, Miller later attributes the difficulty of centre embedded sentences to limits in intrinsic memory capacity, rather than linguistic knowledge: "The fact that an indefinite number of self-embeddings is grammatically acceptable, yet at the same time psychologically unacceptable would seem to imply that a clear distinction is necessary between our theory of language and our theory of the language user." This distinction is made in terms of "the distinction between knowing a rule and obeying it" (Miller and Isard, 1964:294). Miller and Isard therefore assume that language users possess a generative grammar but are unable to obey its rules on account of memory limitations. However, no hard evidence is offered to support the idea that language users do in fact possess generative grammars. In fact, the results of Miller and Isard's famous experiment support an alternative interpretation. In their study, they presented subjects with 22 word sentences in various syntactic structures. Subjects heard the recorded sentences over earphones and attempted to repeat each sentence verbatim as soon as it was played out. The procedure was carried out five times in succession for each sentence. It was found that subjects had difficulties in recalling sentences with more than one self-embedding but had no such difficulties with non-self-embedded sentences. Miller and Isard interpret this result "in terms of an analogy with information processing by computer systems" whereby a subroutine which calls itself erases its return address in program memory, resulting in comprehension failure. In other words, self-embedding is said to cause interference (see Stabler, 1994 and Lewis, 1996 for recent developments of this idea). However, Miller and Isard also report two results which are incompatible with this account and which are never mentioned in numerous citations of this study. Firstly, "learning occurred; for all types of sentences the average percentage of the words recalled and in correct order increased on successive repetitions" (297). It is not clear how learning can be accommodated by the subroutine hypothesis. Recall that subjects showed effects learning even though they had to repeat the sentences *immediately* after they had been presented. This suggests subjects must have been storing the sentences in long term memory as they heard them. Poor recall of self-embedded sentences might therefore not be purely a question of short term memory architecture but also of ltm storage and retrieval. Incidentally, Ericsson and Kintsch (1995) review evidence that performance in the Daneman and Carpenter test is correlated with recognition memory for the test sentences. This finding suggests that superior performance in the test relies on superior access to ltm. Secondly, Miller and Isard report individual differences in the recall of centre embedded sentences: "some subjects handled two self-embeddings as well as they handled one, which suggests that the subroutine analogy needs to modified slightly to some subjects can handle two self-embeddings" (301). It is not easy to see how the subroutine analogy can be modified in a non-arbitrary way to account for individual differences in the ability to handle multiple self embedding and no such modifications appear to have been proposed. However, since Miller and Isard do not appear to have controlled for grammatical skill, it is quite possible that the individual differences they found simply reflect different levels of linguistic skill. Evidence for this possibility is provided by Blumenthal (1966) and Stolz (1967). Thus, two empirical cornerstones of the competent but resource limited language user are open to alternative interpretations, namely that capacity limitations in general are due to a lack of knowledge and in particular, that breakdowns in sentence processing are due to a lack of the requisite grammatical knowledge. Subsequent research (reviewed in Ericsson and Kintch, 1995) has established that increases in domain knowledge are accompanied by increases in domain specific working memory capacity and that the interference which normally results when two similar tasks are performed at the same time can be reduced through training (Spelke, Hirst & Neisser, 1976). Resource limitation accounts of the results of the present study are therefore far from compelling. It should also be noted, in any case, that it is contradictory to blame working memory capacity limitations for a failure to apply phrase structure rules, since phrase structure is a form of recoding and should increase rather than deplete memory (see proposals in Miller, 1962; Johnson, 1965 and Levelt, 1970).

The present results are generally more compatible with key findings in skilled processing research. Some systematic differences between the performance of experts and novices have been established which correspond to some of the present findings. Compared to novices, skilled performers are a) more accurate; b) generally faster; c) reliant on less information in order to carry out a task; d) more consistent and e) in possession of more hierarchically developed knowledge structures (The first four characteristics of skilled performers are discussed in Johnson, 1996 and the fifth in Chi, Glaser and Reese, 1982). This summary of expert novice differences fits the results of this study quite well if graduates are considered as grammatical experts and non-graduates as grammatical novices (recall that the graduates have studied linguistics and many of them have taught English professionally). The graduates are a) more accurate at the parsing task; b) faster; c) able to use syntactic information in the absence of semantic support; d) more selective and therefore more consistent and e) better able to handle complex hierarchical structures. This application of the expert-novice distinction to linguistic skill is consistent with earlier findings in reading research. Compared to less skilled readers, skilled readers are a) more accurate; b) are faster; c) able to use less information to accomplish a task; d) better at suppressing irrelevant information and e) better able to organise textual information into chunks (see Cromer, 1970; Berger & Perfetti, 1977; Gernsbacher, 1990 and Meiran, 1996 for characteristics of skilled readers). The fact that most non-native graduates were taught grammar explicitly may explain their superior performance and consistency, since this training may have helped them to assign greater salience to grammatical structure than native speakers.

At a finer level of analysis, the results are consistent with one particular theory of skill acquisition. According to Logan's (1988) instance theory of automatisation:

> *... novices begin with a general algorithm that is sufficient to perform the task. As they gain experience, they learn specific solutions to specific problems, which they retrieve when they encounter the same problems again. Then, they can respond with the solution retrieved from memory or the one computed by the algorithm. At some point, they may gain enough experience to respond with a solution from memory on every trial and abandon the algorithm entirely. At that point, their performance is automatic. Automatisation reflects a transition from algorithm-based performance to memory-based performance*. (1988:493).

This theory stands in direct contrast to the classical model with respect to the role played by long term memory during processing. The classical language user executes exactly the same computations each time it encounters the same input regardless of prior experience. This amnesia results from the classical language user's inability to make use of its prior linguistic experiences stored long term memory. However, it in not clear that real language users suffer from the same handicap. According to Ericsson and Kintsch, 1995, skill acquisition depends on making efficient use of long term memory. Efficient use of long term memory, in turn, depends on the development of retrieval structures which provide rapid and reliable access to ltm. Elaborating and streamlining the retrieval structures to meet task demands increases the accessibility of ltm, hence the superior task-specific working memory capacities of experts over novices. This approach makes it easier to understand why practice makes perfect in human performance. The effect of practice may be to develop the knowledge structures needed to perform a given task. For instance, performance in rule-governed domains such as mathematics and chess is typically slow and error prone when it is based purely on knowledge of the rules but become faster and more accurate with practice. Children who are learning basic mathematics have to apply an algorithm slowly and laboriously in order to multiply 2 by 5. More skilled children and adults obtain the answer simply and rapidly through memory retrieval. In solving a more complex mathematical problem, skilled individuals break the problem down into previously encountered problems which can be solved directly through memory retrieval (Logan, 1988:493). The fact that rule governed tasks like mathematics are heavily dependent on long term memory for efficiency is particularly significant in view of the fact that arguments for infinite generativity have often been illustrated in terms of an analogy to mathematics (eg. Chomsky, 1963:327 and Fodor & Pylyshyn, 1988:34-5).

Key findings from the experiment are compatible with Logan's theory of automatisation. To begin with, the theory predicts between and within subject differences in the performance depending on the choice of task. This prediction is borne out by the between-group and within-subject differences observed in the experiment. While the highly educated subjects outperformed the less educated subjects, there were also large within subject differences on performance on filler vs. test items as well as for key vs. non-key questions. The large effect of plausibility which is also consistent with the theory. The explanation for this result in terms of the model is simple: just as we can arrive at the answer to 2 X 5 by memory retrieval rather than by application of a multiplication algorithm, so the interpretation of plausible sentences can be retrieved from memory with little or no syntactic analysis. Given, "The convict will be difficult to get the bank manager to give a loan to", the mere registration of the concepts 'bank manager'; 'convict' and 'difficult to give loan to' is sufficient to trigger event related knowledge which favours the interpretation that it is the bank manager who might a give a loan to the convict. This knowledge is of course misleading when it comes to "The bank manger will be difficult to get the convict to give a loan to". The fact that subjects were easily misled in this way is an indication that they were not paying close attention to the syntax. A related finding was reported by Stolz:

> *When a sentence involving highly restrictive semantic constraints is encountered, the average S[ubject] seemed to do very little syntactic processing. It may generally be the case that syntactic analysis is only carried out insofar as an analysis is absolutely required to produce a semantic interpretation for a sentence. If there is only a single reasonable semantic interpretation due to semantic (or pragmatic) constraints, then detailed syntactic processing may be ignored.* (1967:872)

The pattern of errors is suggestive of memory retrieval effects. In the object relative regions of the PG structures, eg. “The man who Peter saw...” Group 3 displayed a significant increase in reading times at ‘Peter’. A possible explanation is that members of Group 3 may have been anticipating a subject rather than an object relative structure and were therefore expecting to see a verb rather than an NP after the complementiser (i.e. they were expecting to see “The man who VERB” rather than “The man who NP”). It is also possible that the use of ‘who’ rather ‘whom’ may have garden pathed Group 3. In either case, there is a group difference which is theoretically attributable to frequencies in patterns of usage. Secondly, a gap effect (i.e. an increase in reading times at the gap region) was found even though most subjects did not fill the gap correctly. A possible explanation for this result is that subjects were expecting an object at the gap location and this expectation was violated. Subjects did, in fact, complain that “there is an object missing”. Thirdly, subjects often omitted an important constituent in the CNP sentences, interpreting “The doctor knows that the fact that taking good care of himself is essential surprises Tom” as either “The doctor knows that the fact that taking good care of himself is essential” or “The doctor knows that the fact that taking good care of himself surprises Tom”. The actual test sentence has the structure: NP knows that the fact that NP is VP VP. The subjects’ responses had the structure: X knows that NP VP. It may be that this simpler structure is the closest approximation to the input that subjects could come up with and it is this which they used to represent the sentence. This possibility is consistent with Blumenthal’s (1966) finding that subjects often misinterpreted centre embedded sentences as simpler, more familiar conjoined relative clause sentences. The idea that comprehension depends on long term memory representations has been developed by MacDonald & colleagues, 1994 and 1995; & Christiansen (in press). This view is clearly at odds with the notion that syntactic structures are generated by rule for each sentence. The key objection to the use of precompiled linguistic representations is that they do not account for creative uses of language (Frazier, 1995). However, it should be considered that, while a knowledge of mathematics imparts the knowledge necessary to carry out a range of mathematical tasks, familiar mathematical problems are easier to solve than unfamiliar ones. In the case of language, it needs to be seen whether unfamiliar structures are understood as effortlessly as familiar ones. The present experiment has shown that native speakers have difficulties and may even fail to comprehend novel structures. When successful, they take longer to comprehend unfamiliar structures relative to familiar ones. It appears from the experiment that understanding novel structures may require conscious problem solving. There is no guarantee, however, that conscious problem solving will help subjects arrive at correct interpretations.

The study has raised a number of specific questions for further research. Firstly, there are indications that the locus of difficulty in object relative constructions may occur earlier in the sentence than was previously thought. Secondly, there may be individual differences associated with this difficulty which are mediated by patterns of usage. Thirdly, individual differences were found the interpretation of control structures. Given "The man who Paul met after discovering ..." Group 1 were more likely to name 'Paul' as the subject of 'discovering', while Group 3 were more likely to name 'The man'. It is not clear why this should be so. Fourth, there was an interaction between Groups and Structures in the gap effect which is also not easy to explain. Fifth, it would be interesting to find out how much lexical knowledge contributes to the comprehension of simple declarative sentences. Are syntactically more skilled individuals better at parsing sentences like “Sleep stars radio green” than less skilled individuals? (see, for instance, Geer at al, 1972). Sixth, to what extent can syntactic skill be acquired in an experimental context? This question has been addressed by the training studies cited earlier. However, it is important to ask this question with the aim of finding out how many of the processing difficulties attributed in the past to resource limitations can be eliminated through training. Additionally, can practice reduce reading times in accordance with the power law of practice? A positive answer would bring syntactic processing even closer to other kinds of acquired skill. Seventh, can ungrammatical sentences prime each other? If they do, this would be evidence in favour of the long term memory based approach. Eighth, how reliably do subjects interpret novel structures in terms of familiar ones? Ninth, can semantic processing outrun syntactic processing so as to bring about an early termination of syntactic processing? Tenth, to what extent is parsing affected by task demands? Finally, it necessary to measure on-line differences in the way familiar vs. unfamiliar structures are processed and to try and find out to what extent conscious sentence comprehension resembles conscious problem solving.

CONCLUSION

In line with previous studies, this study found negative evidence for the idea that native speakers of a language know their language perfectly. It is important to bear in mind that, for all its influence in psycholinguistics, this idea rests purely on an idealisation made by Chomsky for descriptive purposes. Chomsky has already been cited in the introduction of this paper acknowledging the possibility of individual differences in linguistic competence. A number of consequences attend the possibility of individual differences in grammatical competence between native users of a language. Firstly, such differences open the logical possibility that linguistic knowledge is not fully productive. While it is also possible that individuals may possess different but still generative grammars, the evidence from this and other studies indicates that there are normal native speakers of English who are not fully syntactically productive. On the contrary, this study and that of Sasaki (1997, with respect to Japanese) indicates that non-native speakers of a language may be more productive than native speakers. If native speakers are not uniformly competent, questions arise as to the source of such differences. This study and others have indicated that the level of formal education appears to correlate with linguistic ability. The superior performance of non-native speakers of English who acquired English through explicit instruction provides further evidence for this argument. Additional evidence is provided by studies, cited earlier, which report improvements in the grammatical performance of native speakers who undergo training in grammar. Generally speaking, experience appears to be a key source of individual variation. Thirdly, if grammatical competence is not uniform, it may be expected that native speakers will attempt to make up for insufficient grammatical knowledge by using world knowledge. The large effect of plausibility on the native non-graduates indicates as much. Finally, differences in competence suggest that native speakers will not be able to automatically comprehend sentences which lie outside their grammatical ability. Differences in scores on questions on the same sentence in this study provide evidence for this argument. In brief, once differences in linguistic competence are acknowledged, it becomes at least logically possible that native speaker competence is not fully generative, autonomous, automatic and constant. This study and others before it have provided empirical evidence for these logical possibilities. It is noteworthy that the two key architects of the classical language user are ambivalent towards it. On the hand, Chomsky uses the term ‘grammar’ with “a systematic ambiguity” to refer, “first, to the native speaker’s internally represented ‘theory of his language’ and, second, to the linguist’s account of this” (1965:25). This ambiguity, according to Ney (1993), has misled many of Chomsky’s followers (see also the discussion in Moore and Carling, 1982:104-117). On the other hand, Miller (1956:86) leaves unresolved the question of whether capacity limitations are “built into us either by learning or by the design of our nervous systems”. Psycholinguistic research has so far explored one of the possibilities suggested by Chomsky and Miller - the notion of an infinitely competent but capacity limited language user. There is another possibility still to be explored - that of a finitely competent language user with an infinite capacity to learn.

References

Atkinson, R.C. & R.M. Schiffrin. (1968). Human memory: A proposed system and its control processes. In K.W. Spence and J.T. Spence (eds), *Advances in the psychology and learning and motivation research and theory. Vol 2*. New York: Academic Press.

Baruzzi, Anna (1982). Effects of Degree of Education on the Comprehension of Syntactic Structures in Normal and Aphasic Populations. *McGill Working Papers in Linguistics,* 2, 56-74.

Baird, R. & J.D. Koslick. (1974). Recall of grammatical relations within clause-containing sentences. *Journal of Psycholinguistic Research, 3,* 165-171. cited in Levelt, W.J.M and G.B. Flores D’Arcais. (1978)

Bates, E. & J.C. Goodman. (1997). On the Inseperability of Grammar and the Lexicon: Evidence from Acquisition, Aphasia and Real-time Processing. *Language and Cognitive Processing, 12, 5*, 505-892.

Bates, E., McNew, S., MacWhinney, B.,Devescovi, A. & S. Smith. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11:245-99. Cited in Kilborn and Ito (1989).

Bauer, B., & Besner, D. (1997). Processing in the Stroop task: Mental set as a determinant of performance. *Canadian Journal of Experimental Psychology, 51*, 61-68.

Berger, N.S. and C.A. Perfetti (1978). Reading skills and memory for spoken and written discourse. *Journal of Reading Behaviour*, 9, 1, 7-16.

Bever, T.,Sanz, M. and D.T. Townsend. (1998). The Emperor’s Psycholinguistics. *Journal of Psycholinguistic Research, 27, 2,* 261-283.

Blank, G.D. (1989). A finite and real-time processor for natural language. *Communications of the ACM, 32,* 1174-1189. cited in Lewis, R.L. (1996).

Blaubergs, M.S. & Braine, M.D.S. (1974). Short term memory limitations on decoding self-embedded sentences. *Journal of Experimental Psychology 102*, 745-748.

Blumenthal, A. (1966). Observations with self-embedded sentences. *Psychonomic Science 6* (10).

Boyland, J.T. (1996). Morphosyntactic Change in Progress: A Psycholinguistic Treatment. Doctoral dissertation. University of California, Berkeley.

Branigan, H., Pickering, M. and A.J. Stewart (submitted). Syntactic Priming in Comprehension.

Broadbent, D.E. (1954). The role of auditory localisation in attention and memory span. *Journal of Experimental Psychology, 47*, 191-196.

Bybee, J. And S. Thompson. (1997). Three frequency effects in syntax. Paper presented at Twenty-Third Annual Meeting of the Berkeley Linguistics Society.

Caplan, D. and G. Waters. (in press). Verbal Working Memory and Sentence Comprehension. *Behavioral and Brain Sciences*.

Chi, M.T.H. (1976). Short-term memory limitations in children. Capacity or processing deficits? *Memory and Cognition, 4, 5*, 559-572.

Chi, M.T.H., Feltovich,P.J and R. Glaser. (1980). Categorisation and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5, 121-152

Chi, M.T.H., Glaser, R. and E. Rees. (1983). Expertise in problem solving. In R.J. Sternberg (ed.), *Advances in the psychology of human intelligence*, Volume 2. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Chomsky, C (1969) *The Acquisition of Syntax in Children from 5 to 10.* MIT Press.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1963). Formal Properties of Grammars. In Luce, R.D., Bush, R. and E. Galanter. (eds*), Handbook of Mathematical Psychology, Vol II*, 323-418. New York: Wiley.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA. MIT Press.

Chomsky, N.(1980). Comments in M. Piatelli-Palmarini (ed.) Language and Learning. London: Routledge & Kegan Paul.

Connell, P. (1982). On Training Language Rules. *Language Speech and Hearing Services in Schools, 13, 4*, 231-40.

Cowper, E.A. (1976). *Constraints on sentence complexity: A model for syntactic processing.* Unpublished doctoral dissertation, Brown University. cited in Lewis, R.L. (1996).

Cowan, N. (1995). *Attention and memory: an integrated framework*. Oxford University Press.

Cox, J. (1976). A study of the Syntactic Competence of Adult Beginning Readers. Paper presented at the Annual Meeting of the International Reading Association, based on University of North Carolina Ph.D thesis.

Craik, F.I.M. & R.S. Lockhart. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

Cromer, Ward. (1970). The Difference Model: A New Explanation For Some Reading Difficulties. *Journal of Educational Psychology*, 61, 6, 471-483.

Cupples, L. and V.M. Holmes (1987). Reading skill and interpretation of temporary structural ambiguity. *Language and Cognitive Processes*, 2, 179-203.

Cupples, L. and V.M. Holmes (1992). Evidence for a Difference in Syntactic Knowledge Between Skilled and Less Skilled Adult Readers. *Journal of Psycholinguistic Research*, 21, 4, 249-275

Dabrowska, E. (1997). The LAD goes to school: A cautionary tale for nativists. *Linguistics*, 35, 4, 735-766.

Dearborn, W.F. and Anderson, I.H. (1973). A new method for teaching phrasing and increasing the size of reading fixations. The Psychological Record 1, 459-475. cited in Cupples, L. and V.M. Holmes (1987).

Elardo, R. (1972). The experimental facilitation of children's comprehension and production of four syntactic structures. *Child Development*, *42*, *6*, 2101-2104.

Ericsson, K.A. and W. Kintsch. (1995). Long-Term Working Memory. *Psychological Review*, 102, 211-245.

Eysenck, M.W. and M.T. Keane. (1995). *Cognitive Psychology: A Student's Handbook.* Hove, Psychology Press.

Fodor, J. (1983). *Modularity of Mind*. MIT.

Fodor, J.A. and Z. W. Pylyshyn.(1988). Connectionism and Cognitive Architecture: A

Critical Analysis. *Cognition 29*, 3-17

Fodor, J.D. (1998). Learning to Parse? *Journal of Psycholinguistic Research, 27, 2*, 285-319.

Forster, K.I. and Olbrei, I. (1973). Semantic heuristics and syntactic analysis. Cognition 6, 291-325. cited in Cupples, L. and V.M. Holmes (1987). Reading skill and interpretation of temporary structural ambiguity. *Language and Cognitive Processes*, 2, 179-203.

Frazier, l. (1998). Getting There (Slowly). *Journal of Psycholinguistic Research, 27, 2*, 123-146.

Frazier, L. (1995). Constraint Satisfaction as a Theory of Sentence Processing. *Journal of Psycholinguistics, 34,6*, 437-476.

Frazier, L. and C. Clifton.(1996). *Construal*. MIT.

Frazier, L. and J.D. Fodor. (1978). The sausage machine: A new two-stage parsing model. Cognition, 6, 291-210. cited in Lewis, R.L. (1996).

Geer, S.E., Gleitman, H. and Gleitman, L. (1972). Paraphrasing and remembering compound words. *Journal of Verbal Learning and Verbal Behaviour*, 11, 348-355.

Gernsbacher, M.A. (1990). *Language Comprehension as Structure Building*. Hillsdale NJ: Lawrence Erlbaum.

Gibson, E.A. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown.* Tech Rep. CMU-CMT-91-125, Centre for Machine Translation, Carnegie Mellon University, Pittsburgh, PA 15213. cited in Lewis, R.L. (1996).

Gillon, G. and B. Dodd. (1996). The effects of training phonological, semantic and syntactic processing skills in spoken language on written language ability. Language, *Speech and Hearing Services in Schools*, *26,1*, 58-68.

Gordon, P.C., Eberhardt, J.L. & J.G. Rueckl. (1993). Attentional Modulation in the Phonetic Significance of Acoustic Cues. *Cognitive Psychology, 25*, 1-42.

Graesser, A.C., Hoffman, N.L. and L.F. Clark. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior 19*, 135-151.

Greenbaum, S. & R. Quirk. (1970). *Elicitation Experiments in English: Linguistics Studies in Use and Attitude.* Longman.

Gross, M. (1979). On The Failure of Generative Grammar. *Language 55,* 4, pp. 859-885.

Hakes, D.T., Evans, J.S. & L.L. Brannon. (1976). Understanding sentences with relative clauses. *Memory and Cognition, 4*, 283-290.

Harrington, M. (1987). Processing transfer: language specific strategies as a source of interlanguage variation. *Applied Psycholinguistics, 8*:351-78. Cited in Kilborn and Ito (1989).

Henderson, J. (1994). Description based parsing in a connectionist network. Tech. Rep. IRCS-94-12, Inst. for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut St. Suite 400C, Philadelphia, PA 19104. cited in Lewis, R.L. (1996).

Henik, A., Friedrich, F.J. and W.A. Kelogg, (1983). The dependence of semantic relatedness on prime processing. *Memory and Cognition 11*, 366-373.

Hill, A.A. (1961). Grammaticality. *Word*, 17, 1-10.

Holmes, V.M. (1979). Some hypotheses about syntactic processing in sentence comprehension. In W.E Cooper and E.C.T. Walker (eds.) Sentence Processing: Psycholinguistic Studies Presented to Merril Garrett. Hillsdale, NJ: Lawrence Erlbaum pp 221-247. cited in Cupples, L. and V.M. Holmes (1987).

Hudson, R. (1996). The difficulty of (so-called) self-embedded structures. http://www.phon.ucl.ac.uk/home/dick/papers.htm

Huey, E.B. (1968). *The Psychology and Pedagogy of Reading*. MIT.

Johnson, F. (1965). The Psychological Reality of Phrase Structure Rules. *Journal of Verbal*

Learning and Verbal Behavior, 4, 469-475.

Johnson, K. (1996). Language Teaching and Skill Learning. Oxford: Blackwell.

Just, M. A. and P. A. Carpenter (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149

Karanth, P., Kudva, A. and V. Aparna (1995). Literacy and Linguistic Awareness. In de Gelder, B. and J. Morais (eds.) *Speech and Reading*. UK: Lawrence Erlbaum. 303-316.

Kail, M. (1989). Cue Validity, cue Cost, and Processing Types in Sentence Comprehension in French and Spanish. In MacWhinney, B. and E. Bates. (eds) *The Crosslinguistic Study of Sentence Processing*. CUP, 77-117.

Kilborn, K. & A. Cooreman. (1987). Sentence interpretation strategies in adult Dutch-English bilinguals. *Applied Psycholinguistics, 8:41* 5-31. cited in Kilborn and Ito (1989).

Kilbon, K. & T. Ito. (1989). Sentence Processing Strategies in Adult Bilinguals. In B. MacWhinney and E. Bates (eds). *The Crosslinguistic Study of Sentence Processing*. CUP, 257-291.

Kramer,P.E., Koff, E. and Z. Luria. (1972). The Development of Competence In An Exceptional Language Structure In Older Children And Young Adults. *Child Development, 43,* 121-130.

Kren, J. et al. (1989). Training students to recognize Intrasentence Phrases to Enhance Reading Comprehension. *Education and Treatment of Children, 12, 3,* 232-42.

Lachman, R., Lachman, J.L. and E.C. Butterfield (1979). Cognitive psychology and information processing. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Leonard, L. (1981). Facilitating linguistic skills in children with specific language impairment. *Applied Psycholinguistics*, 2, 2, 89-118.

Lewis, R.L. (1996). The magical Number Two (or Three) in Sentence Processing. *Journal of Psycholinguistic Research, 25, 1*, 93-115.

Levelt, W.J.M. (1970). Hierarchical chunking in sentence processing. *Perception and Psychophysics 8*, 99-103.

Levelt, W.J.M. (1974). Formal grammars in linguistics and psycholinguistics. Vol III: Psycholinguistic applications. The Hague: Mouton.

Levelt, W.J.M. (1978). A Survey of Studies in Sentence Perception: 1970-1976. In Levelt, W.J.M and G.B. Flores D'Arcais (eds.) *Studies in the Perception of Language*. Bath: John Wiley & Sons, 1-74.

Levin, H. and E.L. Kaplan (1970). Grammatical structure and reading. In H. Levin and J.P. Williams (eds.) *Basic Studies in Reading*. Basic Books, New York, pp 119-133. cited in Cupples, L. and V.M. Holmes (1987).

Logan, Gordon. D. (1980). Attention and Automaticity in Stroop and Priming Tasks: Theory and Data. *Cognitive Psychology, 12,* 523-553.

Logan, Gordon, D. (1988). Toward an Instance Theory of Automatisation. *Psychological Review*, 95, 4, 492-527.

McDonald, J. (1989). The Acquisition of Cue-Category Mappings. In MacWhinney, B. and E. Bates. (eds) *The Crosslinguistic Study of Sentence Processing*. CUP, 375-396.

MacDonald, M.C. and M. Christiansen (submitted). Individual Differences without Working Memory: A Reply to Just & Carpenter and Walters & Caplan.

MacDonald, M.C., Pearlmutter, N. and M.S. Seidenberg (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review 101,* 676-703.

Martin, N. (1996). Cognitive approaches to Recovery and Rehabilitation in Aphasia. Special Issue of *Brain and Language 52*.

Meiran, N. (1996). Is Reading-Ability Related To Activation Dumping Speed - Evidence From Immediate Repetition Priming. *Memory and Cognition, 24,1,* 41-59

Miller, G.A., Galanter, E. and K.H. Pribram (1960). Plans and the structure of behavior. New York: Holt, Rinehart and Winston.

Miller, G. and S. Isard (1964). Free Recall of Self-Embedded English Sentences. *Information and Control 7*, 292-303.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97

Mills, J.A. and G.D. Hemsley (1974). The Effect of Level of Education On Judgements of Grammatical Acceptability.

Mistler-Lachman, J.L. (1972). Levels of Comprehension in Processing of Normal and Ambiguous Sentences. *Journal of Verbal Learning and Verbal Behavior*, 11, 614-623.

Mitchell, D.C. (1994). Sentence Parsing. In M.A. Gernsbacher (ed*.). Handbook of Psycholinguistics*. Academic Press Inc.

Moerk, E.L. (1991). *A first language taught and learned*. Baltimore: Paul H. Brookes Publishing.

Moore, T. & C. Carling. (1982). *Understanding Language: Towards a Post-Chomskyian Linguistics*. Bath: MacMillan Press.

Muncer, S.J. and Bever, T.G. (1984). Sensitivity to propositional units in good reading. Journal of Psycholinguistic Research 13, 275-279.

Neville, D. & E. Searles. (1985). The Effect of Sentence Combining and Kernel Identification Training on the Syntactic Component of Reading Comprehension. *Research in the Teaching of English, 19, 1*, 37-61.

Neuman, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. In W. Printz & A. Sanders (Eds.)., *Cognition and motor processes*. Berlin: Springer. Cited in E. Styles (1997). *The Psychology of Attention*. Hove: Psychology Press.

Newell, A. & H. Simon. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall. cited in Lachman, R., Lachman, J.L. and E.C. Butterfield (1979).

Ney, J.W. (1993). On Generativity: The History of a Notion that Never Was. *Histriographica Linguistica 20*, 441-454.

Norman, D.A. & D.G. Bobrow. (1976). On the role of active memory processes in perception and cognition. In C. Cofer (ed.), *The structure of human memory*. San Fransisco: Freeman. cited in Lachman, R., Lachman, J.L. and E.C. Butterfield (1979).

Osterhout, L. (1997). On the Brain Responses to Syntactic Anomalies: Manipulations of Word Position and Word Class Reveal Individual Differences. *Brain and Language, 59*, 494-522.

Pearlmutter, N.J. and M.C. MacDonald (1995). Individual Differences and Probabilistic Constraints in Syntactic Ambiguity Resolution. *Journal of Memory and Language*, 34, 521-542.

Powell, A. and R.G. Peters (1973). Semantic Clues in Comprehension of Novel Sentences. *Psychological Reports*, 32, 1307-1310.

Pulman, S. G. (1986). Grammars, Parsers and Memory Limitations. *Language and Cognitive Processes, 3*, 197-225.

Reder, L. & A. Cleeremans. (1990). The Role of Partial Matches in Comprehension: The Moses Illusion Revisited. *The Psychology of Learning and Motivation, 25,* 233-256.

Remez, R. E. (1994). A guide to research on the perception of speech. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 145-172). New York: Academic Press.

Richgels, D.J. (1987). Grade school children’s listening and reading comprehension of complex sentences. *Reading Research and Instruction, 25, 3,* 201-219.

Roth, F.P. (1984). Accelerating language learning in young children. *Child Language*, 11,

89-107.
Sampson, G. (1997). Educating Eve. London: Cassell.
Sanders, L.J. (1971) The Comprehension of Certain Syntactic Structures By Adults. *Journal of Speech and Hearing Research*, 14, 739-745.
Sanford, A.J. & S.C. Garrod. (1994). Selective Processes in Text Understanding. In M.A. Gernsbacher (ed) *Handbook of Psycholinguistics.* Academic Press Inc, 699-719.
Sasaki, Y. (1997). Individual variation in a Japanese sentence comprehension task: Forms, functions and strategies. *Applied Linguistics 18, 4*, 508-537.
Shacter, J. (1996). Learning and triggering in adult L2 acquisition. In Brown, G., Malmkjaer, K. and J. Williams (eds), *Performance and Competence in Second Language Acquisition*. CUP, 70-88.
Shapiro, L.P., Nagel, N.H. & B.A. Levine. (1993). Preferences for a verb’s complements and their use in sentence processing. *Journal of Memory and Language, 32*, 96-114. cited in Tanenhaus, M.K. & J.C. Trueswell. (1995).
Spelke, E.S., Hirst, W.C. and U. Neisser (1976). Skills of divided attention. *Cognition* 4,215-230.
Spencer, N.J. (1973). Differences between Linguists and Nonlinguists in Intuitions of Grammaticality-Acceptability. *Journal of Psycholinguistic Research, 2*, 83-98.
Stabler, E. (1994). The finite connectivity of linguistic structure. In Clifton,C., Frazier, L. and K. Rayner (eds). *Perspectives on Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum, 303-336.
Stewart. M.F. (1978a). Freshman Sentence Combining: A Canadian Project. *Research in the Teaching of English*, *12, 3*, 257-68.
Stewart. M.F. (1978b). Sentence combining and syntactic maturity in first year university. *Alberta Journal of Educational Research, 24, 4*, 262-74.
Stolz, W. (1967). A Study of the Ability to Decode Grammatically Novel Sentences. *Journal of Verbal Learning and Verbal Behaviour 6*, 867-873.
Swan, M.B. (1979). Sentence Combination in College Composition: Interim Measures and Patterns. *Research in the Teaching of English*, *13, 3*, 217-24.
Tanenhaus, M.K. & J.C. Trueswell. (1995). Sentence Comprehension. *Speech, Language and Communication*, 217-262.
Trask, R.L. (1993). A dictionary of Grammatical Terms in Linguistics. London: Routledge.
Van Wijk, C.H. & G.A. Kempen. (1983). The development of syntactic formulating skills of children from 9 to 16 years old. *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden, 37, 8,* 491-509.
Weaver, K.S. & K.F. Ruder. (1978). The effect of gestural prompt on syntax training. *Journal of Speech and Hearing Disorders, 43, 4*, 513-523.
Weller, C.(1979). The effect of two language training approaches on syntactical skills of language-deviant children*. Journal of Learning Disabilities, 12, 7*, 470-479.
Wulfeck, B., Juarez, L., Bates, E., & Kilborn, K. (1986). Sentence interpretation in health and aphasic bilingual adults. In Vaid, J. (ed), *Language processing in bilinguals: psycholinguistics and neuropsychological perspectives*. Hillsdale, JN: Erlbaum. cited in Kilborn and Ito (1989).

APPENDIX

Tough Movement Construction

Sentence 1
SN Alison will be hard to get Tim to give a loan to.
SP The convict will be hard to get the bank manager to give a loan to.
SI The bank manager will be hard to get the convict to give a loan to .

Questions and Answers for Sentence 1 (SN)

Q1 Who might give a loan to someone? Tim
Q2 Who might be given a loan? Alison
Q3 What will be hard? Getting Tim to give a loan to Alison.
Q4 Who will find it hard to do something? Someone not mentioned in the sentence.
Q5 Do you think the sentence is grammatical? Yes.

Sentence 2
SP The President will be easy to get Sandy to vote for.
SI Sandy will be easy to get the President to vote for.
SN Sarah will be easy to get Tim to vote for .

Sentence 3
SI The little girl will be difficult to persuade the dog to play with.
SN Mary will be difficult to persuade Andrew to play with.
SP The dog will be difficult to persuade the little girl to play with .

Parasitic Gap 1 Construction

Sentence 1
SP The servant who Tim visited before overhearing the lady proposing to dismiss had lunch in a cafe .
SI The lady who Tim visited before overhearing the servant proposing to dismiss had lunch in a cafe .
SN The girl who Peter visited before overhearing Alex proposing to dismiss had lunch in a café

Questions and Answers for Sentence 1 (SP)

Q1 Who might be dismissed? Servant
Q2 Who was proposing to dismiss someone? Lady
Q3 Who had lunch in a cafe? Servant
Q4 Who overheard something? Tim
Q5 Do you think the sentence is grammatical? Yes

Sentence 2
SI The thug who Paul spoke to after seeing the frail old lady rushing to mug got off the bus
SN The boy who Sarah spoke to after seeing Aaron rushing to mug got off the bus .
SP The frail old lady who Paul spoke to after seeing the thug rushing to mug got off the bus

Sentence 3

SN The man who Melissa met before discovering Brian preparing to interrogate lit a cigarette .

SP The prisoner who the guard met before discovering the officer preparing to interrogate lit a cigarette .

SI The officer who the guard met before discovering the prisoner preparing to interrogate lit a cigarette .

Parasitic Gap 2 Construction

Sentence 1

SI The solicitor who the man met after discovering his mother arranging to leave a lot of money for was having coffee.

SN The man who Peter met after discovering his mother arranging to leave a lot of money for was having coffee

SP The man who the solicitor met after discovering his mother arranging to leave a lot of money for was having coffee.

Questions and Answers for Sentence 1 (SI)

Q1 Who was having coffee? Solicitor
Q2 Who discovered something? Man
Q3 Who was to be left a lot of money? Solicitor
Q4 Whose mother was arranging to leave a lot of money for someone? Man or Solicitor.
Q5 Do you think the sentence is grammatical? Yes.

Sentence 2

SN The man who Peter saw after discovering his girlfriend planning to jilt walked away.

SP The man who the counsellor saw after discovering his girlfriend planning to jilt walked away.

SI The counsellor who Peter saw after discovering his girlfriend planning to jilt walked away.

Sentence 3

SP The actress who the nurse visited after noticing her illness beginning to seriously affect sat down.

SI The nurse who the actress visited after noticing her illness beginning to seriously affect sat down.

SN The actress who the ballerina visited after noticing her illness beginning to seriously affect sat down.

Complex Noun Phrase Construction

Sentence 1

SN Peter knows that the fact that taking good care of himself is essential surprises Tom.

SI Tom knows that the fact that taking good care of himself is essential surprises the doctor.

SP The doctor knows that the fact that taking good care of himself is essential surprises Tom.

Questions and Answers for Sentence 1 (SN)

Q1 What does Peter know? That the fact that taking good care of himself is essential surprises Tom.

Q2 What is essential? Taking good care of himself.

Q3 For whom is something essential? Peter or Tom.

Q4 What surprises Tom? The fact that taking good care of himself is essential.

Q5 Do you think the sentence is grammatical? Yes.

Sentence 2

SP The detective realised that the fact that letting himself be seen by many witnesses was dangerous would not be overlooked by the criminal.

SN Peter realised that the fact that letting himself be seen by many witnesses was dangerous would not be overlooked by Mike .

SI The criminal realised that the fact that letting himself be seen by many witnesses was dangerous would not be overlooked by the detective.

Q2 What was dangerous?

Q2 What would not be overlooked by the criminal?

Q3 What did the detective realise?

Q4 For who might find something be dangerous?

Q5 Do you think the sentence is grammatical?

Sentence 3

SI The child remembered that the fact that washing herself was difficult annoyed her mother.

SP The mother remembered that the fact that washing herself was difficult annoyed the child.

SN Sarah remembered that the fact that washing herself was difficult annoyed Anna .

Q1 What annoyed the mother?

Q3 What was difficult?

Q3 Who found something difficult?

Q4 What did the child remember?

Q5 Do you think the sentence is grammatical?

Filler Items

1. Peter knew that Jim would get the job even if he was not really qualified for it.

Questions and Answers
1. What did Peter know? That Jim would get the job even if he was not really qualified for it.
2. Who might get something? Jim.
3. What might someone get? A job
4. Who was not qualified for something? Jim (or Peter).
5. Do you think this sentence is grammatical? Yes.

2. Alex knew that the best way to find out whether or not the plan would work was to ask the man who played the guitar at the party.

Questions and Answers
1. Who wanted to find out something? Alex
2. What was played at the party? A guitar.
3. What did Alex want to find out? Whether or not the plan would work.
4. Who played something at the party? The man.
5. Do you think this sentence is grammatical?

3. Elaine was well aware that, although the building had been designed by John and herself, only he would get the recognition which would ensure his future in the field of architecture.

Questions and Answers
1. Who was aware of something? Elaine.
2. Who had designed the building? Elaine and John.
3. Who would get recognition? John.
4. What was Elaine aware of? That, although the building had been designed by John and herself, only he would get recognition.
5. Do you think this sentence is grammatical? Yes.