

Constraints on Movement Phenomena in Sentence Processing: Evidence from Event-related Brain Potentials

Richard McKinnon and Lee Osterhout

*Departments of Psychology and Linguistics, University of Washington,
Seattle, WA, USA*

Event-related potentials (ERPs) were recorded from 13 scalp electrodes while subjects read sentences, some of which contained a violation of constraints on constituent movement (subjacency or the empty category principle). In two experiments, the first word that indicated a violation elicited a widely distributed positive-going wave with an onset between 250 and 300 msec and a duration of several hundred milliseconds. This effect was similar to the positive shifts that have been observed in response to other types of syntactic violations (P600), but qualitatively distinct from the N400 response elicited by semantically/pragmatically anomalous words. These findings are taken as an indication that movement constraints can be applied during the earliest stages of sentence processing, perhaps in conjunction with the creation of phrase structure. We evaluate the hypothesis that the P600 is a general electrophysiological marker of syntactic anomaly.

INTRODUCTION

Beginning with the pioneering work of Kutas and Hillyard, a growing body of evidence has demonstrated that scalp-recorded event-related brain potentials (ERPs) are sensitive to particular aspects of sentence processing. ERPs are negative and positive voltage changes (or *components*) in the ongoing electroencephalogram that are time-locked to the onset of a sensory, motor or cognitive event (Hillyard & Picton, 1987). Kutas and

Requests for reprints should be addressed to Richard McKinnon, Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA. E-mail: mckinnon@psych.umass.edu.

We thank Janet Fodor for her assistance and guidance at every stage of this project, Julie Boland, Rob Kluender, Cecile McKee, Fritz Newmeyer and two anonymous reviewers for very useful comments, Dierdre Quinn for help with the materials used in Experiment 1, and Dawn Aiken, Linda Mobley and Sherri Sipe for help in running subjects. Portions of these data were presented at the Fifth Annual (1992) and Sixth Annual (1993) CUNY Conferences on Human Sentence Processing. This research was supported by research Grant No. 5 R29 DC 01947 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health, awarded to L.O.

Hillyard (1980a, b, c) reported that contextually inappropriate words (e.g. “I take my coffee with cream and *dog*”) elicit a large-amplitude, negative-going component with a peak around 400 msec (the N400 component). Subsequent research has indicated that N400 amplitude is a function of the semantic or pragmatic fit between the target word and context, even when the target word is contextually appropriate (Kutas & Hillyard, 1984). Although the precise cognitive events underlying the N400 are not known, a consensus exists that N400 amplitude is sensitive to semantic/pragmatic aspects of sentence processing.

Recent attempts to identify ERP effects that are sensitive to syntactic aspects of sentence processing have produced a greater variety of effects (for a review, see Osterhout & Holcomb, 1995). A number of studies have shown that a disparate set of syntactic anomalies, including violations of verb subcategorisation, phrase structure and agreement constraints, elicits a widely or posteriorly distributed positive wave with an onset around 500 msec and a duration of several hundred milliseconds (labelled the P600 by Osterhout & Holcomb, 1992; see also Hagoort, Brown, & Groothusen, 1993; Neville et al., 1991; Osterhout & Holcomb, 1993; Osterhout, Holcomb, & Swinney, 1994; Osterhout & Mobley, 1995). However, other researchers have reported that a variety of syntactic anomalies, including some of those listed above (e.g. subcategorisation errors), elicit an anterior negativity within the window associated with the N400 component (Münte, Heinze, & Mangun, 1993; Rösler, Friederici, Pütz, & Hahne, 1993). Furthermore, in some reports, syntactic anomalies have elicited both an anterior negativity and a P600-like positivity within the same epoch. For example, phrase structure violations have sometimes been reported to elicit both an increase in negativity over left anterior sites beginning at about 300 msec and a posteriorly distributed positive wave with an onset around 500 msec (Neville et al., 1991; Osterhout & Holcomb, 1992, 1993).

The focus of the present study is on the brain’s response to one aspect of syntactic processing, namely the processing of sentences that violate constraints on constituent movement. These constraints are typically considered to be syntactic in nature. For example, one standard theory of grammar, Government and Binding (GB) Theory (Chomsky, 1981, 1986), posits two sets of constraints on constituent movement, labelled “subjacency” and the “empty category principle” (ECP). Consider sentences (1) and (2):¹

1. Who_i do the police believe that John shot ——— $i?$
2. * Who_i do the police believe the claim that John shot ——— $i?$

According to GB Theory, the word *who* in sentences (1) and (2) has moved from its position as object of the verb *shot*. The moved constituent is

¹Asterisks are used to identify sentences that are judged to be ungrammatical.

claimed to leave behind an empty position (or *gap*, indicated by — in the above examples) within syntactic structure. The grammatical constraint that can account for the well-formedness of (1) and the ungrammaticality of (2) has been subsumed under subjacency (Chomsky, 1973, 1981). This constraint results in the creation of a “syntactic island” preventing further extraction of constituents. The movement in (1) is allowed, but in (2) the added noun phrase *the claim* creates a syntactic island preventing movement of *who* to the front of the sentence.²

Another constraint on movement has been subsumed under the empty category principle (Chomsky, 1981, 1986; Epstein, 1991; Huang, 1982; Lasnik & Saito, 1992). This constraint specifies that a moved constituent must maintain a particular structural relationship to the position from which it has been moved (Chomsky, 1981; Lasnik & Saito, 1984, 1992). Consider sentences (3) and (4):

3. John_{*i*} is likely — _{*i*} to win.
4. *John_{*i*} seems that it is likely — _{*i*} to win.

Within GB Theory, the noun phrase *John* in sentences (3) and (4) is assumed to have moved from the subject position of the clausal complement *to win*. This movement is allowed in (3) as the proper structural relationship is maintained between the moved constituent and the gap. However, in (4), the complementiser *that* intervening between the moved constituent and the gap prevents the proper structural relationship, thereby disallowing movement.

The primary question addressed here concerns the ERP response to violations of constraints on constituent movement. Previous experiments examining the response to such anomalies have produced conflicting results. Neville et al. (1991) contrasted the ERP response to subjacency violations (e.g. “What was a sketch of *admired* by the man?”) and well-formed controls (e.g. “Was a sketch of the landscape *admired* by the man?”). ERPs to the first word that indicated a deviation from well-formedness (italicised in the above example) elicited an increase in the amplitude of the P2 component and a subsequent P600-like positive-going wave. In contrast, Kluender and Kutas (1993a) failed to find any consistent differences between the ERPs to subjacency-violating words (e.g. “What do you wonder *who* ...”) and controls (“What do you suppose *that* ...”). One possible explanation for this discrepancy notes the problems engendered when the critical word is immediately preceded by different words across constraint-violating and

²The subjacency constraint has recently been reformulated by Chomsky as part of a more general constraint. In an attempt to unify the barriers to movement and government, Chomsky (1986) argues that certain configurations are crucial to each. On this view, cyclic nodes are no longer stipulated, but arise out of particular combinations of phrasal nodes. We adopt the account of subjacency proposed by Chomsky (1973, 1981) for reasons of ease of exposition; for present purposes, the choice of formalism makes little or no difference. Furthermore, subjacency and the ECP are maintained as distinct modules under the newer barriers approach.

well-formed control versions of each sentence. In order to compare two epochs of ERP, one must assume equality in voltages across conditions at some point before or during the epoch of interest (cf. Garnsey, 1993). The standard "baseline" of activity is the 100 msec preceding critical word onset. This baseline procedure is reasonable when the context immediately preceding the critical words is constant across conditions. However, in both the Neville et al. and Kluender and Kutas experiments, the critical words (at least on some trials) were immediately preceded by different words across conditions. As Neville et al. themselves note, this problem is particularly severe in their study because the critical words in the subadjacency-violating and control sentences were preceded by words from different classes. A function word (preposition) immediately preceded the critical word in the subadjacency-violating sentences, whereas a content word (noun) immediately preceded the critical word in the control sentences. Function and content words are known to elicit distinctly different brain responses (Neville, Mills, & Lawson, 1992). Hence, it is possible that the ERP differences in these two conditions were partially or largely due to artificial remnants of a differential response to the immediately preceding context, rather than being reflections of the brain response to the critical words themselves. Another concern with materials relates only to the Kluender and Kutas study. In the Neville et al. study, violating and non-violating sentences were matched such that the critical word across conditions was always identical. By contrast, in the Kluender and Kutas study, different critical words were presented in the violating and non-violating sentences. This introduces a potential confound between the effects of the grammatical status of the sentences and the lexical content of the critical words.

A second set of questions concerns the temporal course over which constraints on constituent movement are applied during sentence processing. There is considerable evidence that readers rapidly assign syntactic structure in real time, that is on a word-by-word basis (Frazier, 1987; Frazier & Rayner, 1982; Marslen-Wilson & Tyler, 1980). Correspondingly, words that violate phrase structure rules or verb subcategorisation constraints are rapidly perceived to be anomalous (Neville et al., 1991; Osterhout & Holcomb, 1992; Osterhout et al., 1994). Whether or not this is also true of constraints that apply to constituent movement has been a matter for debate. The distinction between *structure-building* and *structure-checking* processes provides a useful framework for discussing this question (Mitchell, 1989; see also Pickering, Barton, & Shillcock, 1994). At least some evidence appears to indicate that movement constraints are applied during the initial structure-building stage of sentence processing. For example, Stowe (1986) reported evidence indicating that readers do not posit a gap in cases where the gap would violate a movement constraint. However, exactly the opposite conclusion was reached by Pickering et al. (1994), who interpreted their data

as indicating that the parser attempts to form dependencies between posited gaps and their antecedents even when the dependency violates constraints on constituent movement. Pickering et al. suggest that constraints on movement are not part of the core syntactic operations of the parser (the construction of phrase structure), but instead are applied during a subsequent structure-checking stage in which the computed structure is evaluated for violations of movement (and possibly other) constraints. Consistent with such a possibility, evidence from other studies has been taken to indicate that the application of movement constraints can be dissociated from the construction of phrase structure (Forster, 1987; Freedman & Forster, 1985; but see Crain & Fodor, 1987).

Given their on-line, continuous nature, ERPs promise, in principle, to provide an initial indication as to how quickly anomalies involving constituent movement are detected during processing. Unfortunately, however, inferences about the timing of cognitive processes based on ERP data are often problematic (Garnsey, 1993; Osterhout, 1994). This is particularly true in cases where the cognitive events underlying an effect are not known. For example, the P600 effect elicited by syntactic violations often has an onset around 500 msec after presentation of the anomalous word. This by no means licenses the inference that the assignment of syntactic roles occurs only 500 msec after word presentation, because the P600 might reflect syntactic processes only indirectly. Nonetheless, under certain conditions, fairly strong inferences seem warranted. With respect to the current study, evidence that the waveforms elicited by non-violating and violating sentences do not diverge at the critical words (i.e. the first word indicating a movement violation and the control word in the grammatical sentences), but instead diverge only at a later point in the sentence, might be taken as an initial indication that movement constraints are not immediately applied. Conversely, evidence that the waveforms diverge very early in the epoch elicited by the critical words would suggest that the movement constraints are in fact applied at an early stage of processing.

To summarise, we examined the brain response to violations of two constraints on constituent movement: subjacency (Experiment 1) and the empty category principle (Experiment 2). Critical words were always identical across violating and control sentences. In order to eliminate the potentially confounding effects of variable contexts preceding the critical words, we carefully matched the material preceding the critical regions of constraint-violating and control sentences.

EXPERIMENT 1

In Experiment 1, we presented sentences similar to (1)–(4) in Table 1. Sentence (1) contains a constituent (*which of his staff members*) that has moved from its canonical position (indicated by —) as part of a

prepositional phrase within the complement clause. The presence of the word *when* in complementiser position prevents movement of the moved constituent, resulting in ungrammaticality (i.e. a violation of subadjacency). Sentence (2) provides a matched control sentence that does not contain a moved constituent. The contrast of interest concerns the response to the word in complementiser position (*when*) in each sentence type. This word in (1) represents the earliest indication that the sentence violates subadjacency, whereas the same word in (2) is not anomalous in any sense. Note that the immediate context preceding the critical words was constant across conditions, eliminating the possibly confounding effects of variable context.

However, two other potential confounds exist in the comparison of critical words in sentences (1) and (2). First, the sentence structure preceding the critical word is more complex in (1) than in (2); hence, any ERP differences in the response to the critical word across conditions might be due to a syntactic complexity effect. Second, Kluender and Kutas (1993a, b) have recently reported an ERP effect which they claim is associated with holding a moved constituent in memory. Words immediately following a moved constituent in sentences similar to that in sentence (1) in Table 1 elicited a negative-going wave over anterior portions of the left hemisphere. In order to determine whether any differences we observed in the ERPs to (1) and (2) were due to these factors rather than to the movement violation, we presented sentences similar to (3) and (4) in Table 1. Sentence (3) is identical to (1) except that the complementiser *that* introduces the subordinate clause. While (3) entails some awkwardness, standard GB Theory deems such sentences to be well-formed. Sentence (4) represents the corresponding control sentence that does not contain a moved constituent.

We were therefore primarily interested in the response to the words in complementiser position. If the parser recognises the subadjacency violation at

TABLE 1
Examples of Sentences Presented During Experiment 1

-
1. *I wonder which of his staff members_i the candidate was annoyed *when* his son was questioned by ______i.
(ungrammatical, movement, "when" sentences)
 2. I wonder whether the candidate was annoyed *when* his son was questioned by his staff member.
(grammatical, no movement, "when" sentences)
 3. I wonder which of his staff members_i the candidate was annoyed *that* his son was questioned by ______i.
(grammatical, movement, "that" sentences)
 4. I wonder whether the candidate was annoyed *that* his son was questioned by his staff member.
(grammatical, no movement, "that" sentences)
-

the earliest possible position within the sentence, and if such violations elicit a measurable brain response, then the ERP to *when* in sentences like (1) should be distinct from the ERP to the same word in (2). Such differences should not be evident when comparing the responses to the complementisers in sentences like (3) and (4), neither of which violates subadjacency.

Method

Subjects. Sixteen undergraduate and graduate students from the University of Washington (9 males, 7 females) participated for course credit or for a small compensation. The subjects' age ranged from 18 to 32 (mean = 23) years. All subjects were right-handed and native speakers of English with normal or corrected-to-normal vision.

Materials. Four versions of 120 "root" sentences were prepared (Table 1). Each root sentence contained a main clause followed by an embedded clause. The four versions of each sentence were formed by factorially manipulating "movement" (movement *vs* non-movement) and complementiser (*when vs that*). In all sentences, the embedded clause contained a direct object or a prepositional phrase. In the movement sentences, the direct object or the object of the preposition was moved in front of the main clause for question formation. A sample of the experimental sentences presented is provided in Appendix 1.

These materials were then used to create four stimulus lists. Each list contained 30 exemplars of each sentence type. The items were counterbalanced such that only one version of each root sentence was presented on a given list. Additionally, 60 filler sentences were constructed which were similar to the experimental sentences. Half of the filler sentences contained subadjacency violations, phrase structure violations or subcategorisation violations. Half of the fillers contained legal movement constructions. Thus, each subject saw a total of 180 sentences, 60 of which were ungrammatical.

Procedure. Each trial consisted of the following events: A fixation cross appeared for 500 msec, after which a sentence was presented in a word-by-word manner, with each word appearing at the centre of the screen for 300 msec. A blank-screen interval of 100 msec separated words. Sentence-ending words appeared with a period. A 1450 msec blank-screen interval followed each sentence, after which a prompt appeared asking subjects to decide if the preceding sentence was "acceptable" or "unacceptable". Acceptable sentences were defined as those which were semantically coherent and grammatically well-formed. The subjects responded by pressing one of two buttons, which were counterbalanced (left and right)

across subjects. The subjects were tested in one session which lasted approximately 2 h, during which they were seated in a comfortable chair located in an isolated room.

Data Acquisition and Analysis. Continuous EEG was recorded from 13 scalp sites using tin electrodes attached to an elastic cap (Electrocap International). Electrode placement included the International 10–20 system locations (Jasper, 1958) over homologous positions over the left and right occipital (O1, O2) and frontal (F7, F8) regions and from the frontal (Fz), central (Cz) and parietal (Pz) midline locations. In addition, several non-standard sites over posited language centres were used, including Wernicke's area and its right hemisphere homologue (WL, WR: 30% of the interaural distance lateral to a point 13% of the nasion–inion distance posterior to Cz), posterior temporal (TL, TR: 33% of the interaural distance lateral to Cz) and anterior temporal (ATL, ATR: one-half the distance between F7/F8 and T3/T4). Vertical eye movements and blinks were monitored by means of two electrodes, one placed beneath the left eye and the other placed to the right of the right eye. The above 15 channels were referenced to an electrode placed over the left mastoid bone and were amplified with a bandpass of 0.01–100 Hz (3 dB cut-off) by a Grass Model 12 amplifier system. Activity over the right mastoid was actively recorded on a sixteenth channel to determine if there were any differences in the effects of the experimental variables on the mastoid recordings. No such effects were observed.

Continuous analogue-to-digital conversion of the EEG and stimulus trigger codes was performed by a Data Translation 2801-A board and a 486-based computer at a sampling frequency of 200 Hz. Epochs were comprised of the 100 msec preceding and the 1180 msec following presentation of individual words in the sentences. Trials characterised by excessive eye movement or amplifier blocking were removed prior to averaging. The ERPs were quantified as the mean voltage within a latency range following presentation of words of interest relative to a baseline of activity comprised of the 100 msec prior to, and 50 msec subsequent to, presentation of the words of interest (unless otherwise noted).^{3,4} Analyses of

³The convention is to use the 100 msec of activity preceding critical word onset as the baseline. We chose to use an unconventional baseline because, in our judgement, such a baseline most adequately eliminated differences between conditions that existed prior to onset of the critical words; for example, we wanted to minimise the influence of ERPs elicited by the words preceding the critical region on the response to the critical words themselves. Such effects were exacerbated by the relatively rapid rate of word presentation used here. However, all analyses reported here were also conducted with the standard baseline composed of the 100 msec preceding critical word onset. The results of these analyses were qualitatively similar to those reported.

⁴Unless otherwise noted, the averages reported here were not response-contingent; that is, we report averages formed over grammatical and ungrammatical sentences with no regard to

variance were performed on mean amplitude within three time windows: 150–300, 300–500 and 500–800 msec. These windows were chosen because they roughly correspond to the latency ranges of the N1–P2 complex, N400 component and P3/P600 positivities often reported in cognitive ERP studies. Data acquired at midline and lateral sites were treated separately to allow for quantitative analysis of hemispheric differences. To protect against Type I error due to violations of the assumption of equal variances of differences between conditions of within-subject factors, the Greenhouse and Geisser (1959) correction was applied when evaluating effects with more than one degree of freedom (reported as a second probability value in such analyses).

Results and Discussion

Acceptability Judgements

Subjacency-violating sentences [e.g. sentence (1) in Table 1] and their well-formed controls [e.g. sentence (2)] were judged to be acceptable on 25 and 85% of the trials [$F(1,15) = 59.23, P < 0.01$]. The low accuracy rate for subjacency violations fits with previous findings by Neville et al. (1991), where subjects were 79% accurate. Well-formed movement sentences and controls were judged to be acceptable on 48 and 82% of the trials [$F(1,15) = 25.38, P < 0.01$]. Hence, although most subjects on most trials judged the subjacency-violating sentences to be unacceptable, the “well-formed” movement sentences were judged to be acceptable on only half the trials. This corresponds to the observations of Kluender and Kutas (1993a), who found that grammatical movement constructions were judged to be significantly less acceptable than non-movement constructions.

Event-related Potentials

ERPs to Critical Words. Fifteen percent of the trials involving the presentation of complementisers (roughly equally distributed across conditions) were rejected due to artifact. Grand-average ERPs elicited by the *when* in sentences exemplified by (1) and (2) in Table 1 are shown in Fig. 1. In this and all subsequent plots, the general shapes of the waveforms were similar to those reported in other studies involving visually presented language stimuli (Neville, Kutas, Chesney, & Schmidt, 1985; Osterhout & Holcomb, 1992). A negative–positive complex was visible in the initial 300 msec after presentation of most words. The negative-going component (“N1”) peaked at about 150 msec at most sites; the positive-going wave (“P2”) peaked at about 250 msec and was largest over midline sites. (The

the subjects' end-of-sentence acceptability judgements. Response-contingent averages would have reduced the number of epochs entering into the grand average to unacceptable levels for signal-averaging purposes. Also, the relationship between subjects' on-line processing of the sentence and their eventual acceptability judgements is unclear.

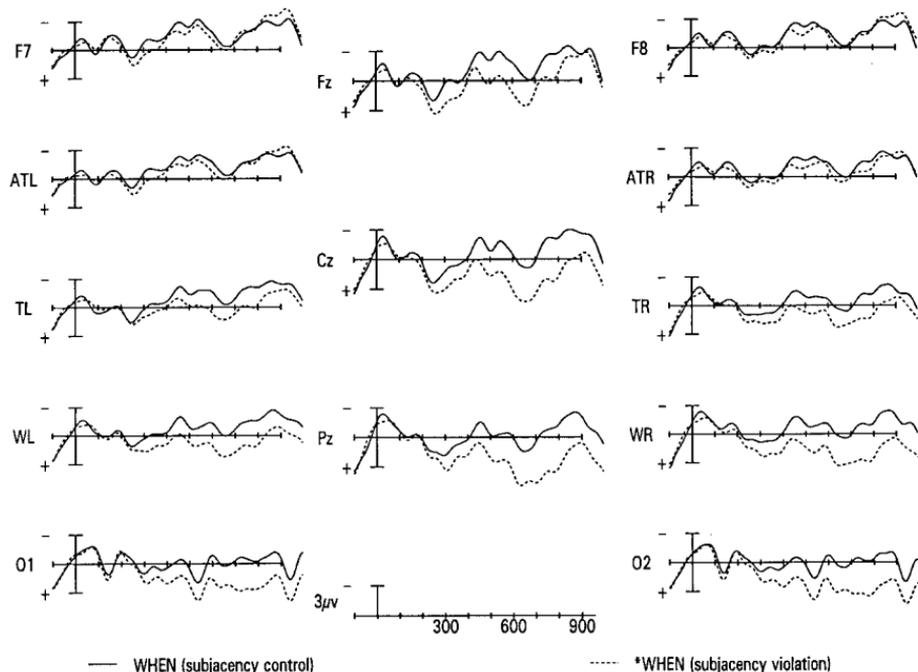


FIG. 1. Grand-average ERPs to *when* in the movement and non-movement sentence conditions, recorded over 13 scalp electrode sites. The vertical bar indicates stimulus onset; subsequent marks set at 100 msec intervals. ERPs plotted relative to a baseline of activity comprised of 100 msec prior to, and 50 msec subsequent to, presentation of critical words.

amplitude of these components was somewhat reduced relative to previous reports, possibly due to the relatively rapid presentation rate employed here; i.e. these components might have been refractory.) Another negative-going component with a peak around 400 msec (N400) was evident in the response to most open-class words. Sentence-final words were followed by a large-amplitude positivity, an effect often seen following well-formed sentences (Friedman, Simson, Ritter, & Rapin, 1975).

The most notable difference across conditions was the presence of a large, widely distributed positive-going wave in the response to *when* in subjacency-violating sentences, relative to controls. Differences between conditions had an onset about 200 msec subsequent to onset of the critical words. Statistical analysis found this effect to be highly reliable within all three time windows. Within the 150–300 msec window, ERPs to *when* in subjacency-violating sentences were more positive than controls [midline: $F(1,15) = 7.35$, $P < 0.02$; lateral: $F(1,15) = 4.42$, $P < 0.05$]. Similar effects were observed in the 300–500 msec window [midline: $F(1,15) = 6.04$, $P < 0.05$; lateral: $F(1,15) = 6.23$, $P < 0.05$]. At lateral sites, differences between conditions were largest over posterior sites [sentence type \times electrode site:

$F(4,60) = 3.31, P < 0.05, P < 0.05]$. These effects continued into the 500–800 msec window: ERPs to *when* in subjacency-violating sentences were more positive-going than controls [midline: $F(1,15) = 13.97, P < 0.05$; lateral: $F(1,15) = 8.95, P < 0.05]$, and differences between conditions were largest posteriorly both over the midline [$F(2,30) = 4.26, P < 0.05, P < 0.05]$ and lateral [$F(4,60) = 10.30, P < 0.05, P < 0.05]$ sites.

In order to determine more precisely the onset of divergences in the ERPs to movement-violating and control sentences, ANOVAs were performed on successive 50 msec segments beginning at 100 msec. Two sets of analyses were performed, one using a 100 msec baseline comprised of the 100 msec of activity preceding word onset and one using a 150 msec baseline comprised of the 100 msec of activity preceding, and 50 msec of activity following, word onset. In both baseline conditions, reliable differences between conditions did not emerge until the 250–300 msec window [100 msec baseline: $F(1,15) = 8.01, P < 0.02$; 150 msec baseline: $F(1,15) = 5.79, P < 0.05]$.

ERPs to complementisers in the “grammatical movement” and control sentences are shown in Fig. 2. Robust differences were not evident, although ERPs to these words in movement sentences were slightly more positive-going at some sites (e.g. Pz) between roughly 300 and 800 msec. However, no reliable effects were found in any of the analyses ($P > 0.05$).⁵ In order to determine whether the difference in the response to *when* in the subjacency-violating and control sentences was greater than the corresponding difference in the response to the complementisers in the grammatical movement and control sentences, we performed ANOVAs on mean amplitude within the 500–800 msec window, treating sentence pair as a within-subject factor. The interaction between sentence pair and sentence type was reliable both at midline [$F(1,30) = 7.84, P < 0.02]$ and at lateral [$F(1,30) = 6.34, P < 0.03]$ sites.

Hence, ERPs elicited by violations of subjacency resembled those previously observed in the response to a variety of syntactic anomalies. Additionally, the onset of the brain response to violations of subjacency, as indexed by diverging waveforms in the ungrammatical and control constructions, was remarkably rapid, beginning as early as 250 msec following presentation of the first possible indication of a violation. Given that recent estimates of the speed of lexical access are of similar duration (Marslen-Wilson & Tyler, 1980; Sabol & DeRosa, 1976), this would suggest that such violations are almost immediately evident to the processing system.

⁵In order to determine whether the small differences between conditions were due to trials on which subjects judged the movement sentences to be unacceptable, we attempted to form response-contingent averages such that only those trials on which subjects judged the sentence to be acceptable were included in the averaged waveform. However, the small number of such trials and the artifact rejection rate made this unfeasible.

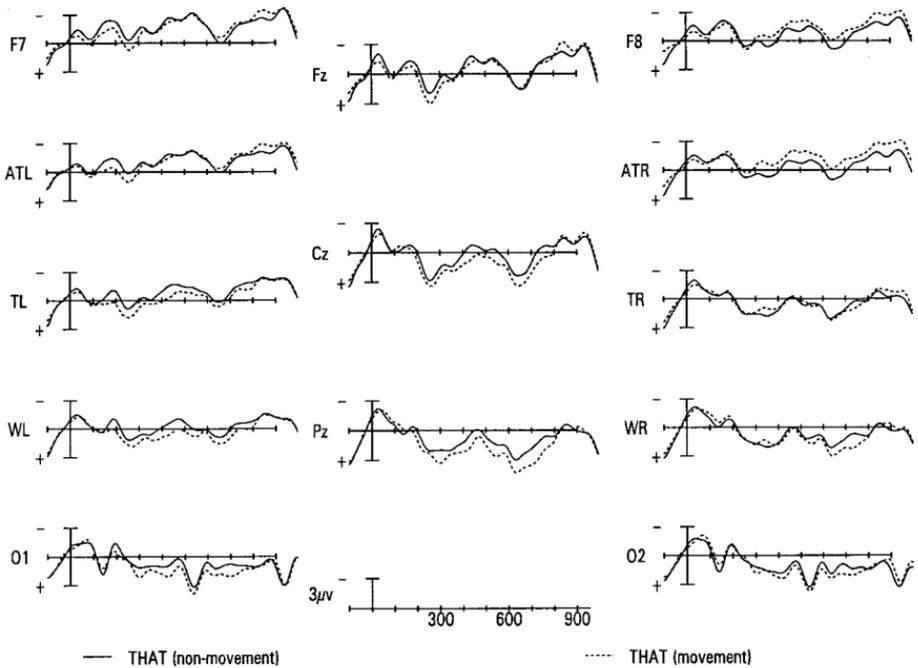


FIG. 2. Grand-average ERPs to *that* in the movement and non-movement sentence conditions.

We should also note that the response observed here is quite similar to that observed by Neville et al. (1991), with one important difference. Neville et al. interpreted the positivity they observed to subjacency violations as consisting of an increase in P2 amplitude followed by a late positive component with an onset around 500 msec. In the current study, such violations elicited an increased positivity concurrent with the peak of the P2 component. However, this positivity did not return to baseline after the P2 but rather extended throughout the recording epoch. At present, it is unclear whether such violations result in the modulation of one or more components.

ERPs to Words Following Moved Constituents. Kluender and Kutas (1993a, b) reported that a left anterior negativity (LAN) between ~ 200 and 500 msec was larger to words that followed a moved constituent than to the same words that did not follow such a constituent. They interpreted this effect as a reflection of the added working memory load occasioned by the moved constituent. Given the large amount of lexical content in the moved constituents used in the present study (e.g. *which of his staff members*), one would expect that there should be a heavy processing load in the movement

sentences presented here, relative to the non-movement sentences. Correspondingly, one would anticipate a large LAN to words that followed the moved constituent (e.g. *the candidate was annoyed*). ERPs to this region in the movement and non-movement sentences, recorded over the left anterior temporal site (the location over which the LAN effect observed by Kluender and Kutas was largest), are shown in Fig. 3, which plots the response to the four words following the moved constituent. Differences between conditions were small in amplitude and statistically unreliable.

ERPs to that and when. Kluender and Kutas (1993a) also reported that ERPs to wh-words (e.g. *who, what*) were more negative-going from about 200 to 800 msec than those to complementisers (e.g. *that*), when these elements appeared in well-formed sentences (fig. 3 of Kluender & Kutas, 1993a). This effect was most notable over right hemisphere and posterior left hemisphere sites, and largely absent from parietal and anterior left hemisphere sites. Kluender and Kutas identified this effect as a change in N400 amplitude and attributed it to differences in the lexical semantics of the various types of complementisers. ERPs to the complementisers in sentences similar to (2) and (4) in Table 1 are shown in Fig. 4. Subjects typically judged these sentences to be acceptable. Consistent with the data reported by Kluender and Kutas, ERPs to *when* were more negative-going than those to *that* between roughly 200 and 800 msec, particularly at the midline, right hemisphere and posterior left hemisphere sites. These differences verged on reliability between 500 and 800 msec [midline: $F(1,15) = 3.74, P = 0.07$]. Although we hesitate to associate this effect with changes in N400 amplitude given its broad distribution over time, this effect clearly indicates that direct comparisons of ERPs to different complementisers in

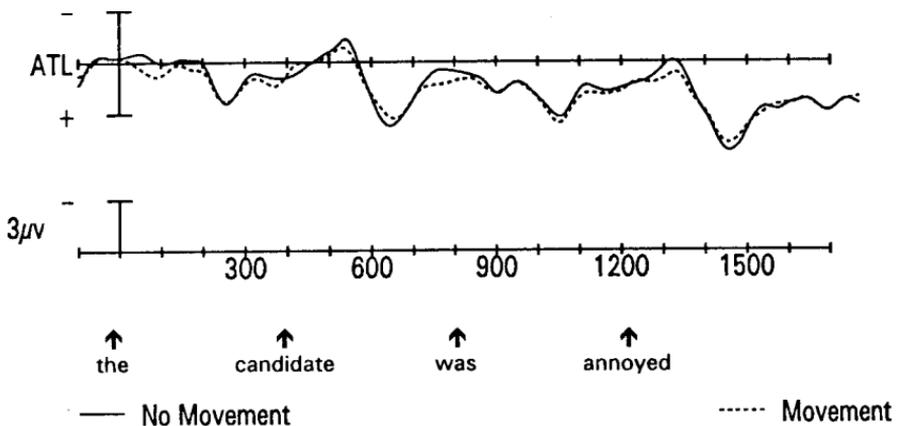


FIG. 3. Grand-average ERPs to the four words immediately preceding the critical words in the movement and non-movement conditions, recorded at the left-hemisphere anterior temporal electrode site.

the current study would potentially confound the effects of anomaly with the effects of word class.⁶

EXPERIMENT 2

In Experiment 2, we examined three brain responses to violations of the empty category principle (ECP). One motivation for including the ECP in our study was to replicate the results of Experiment 1 with different stimuli involving a very different type of movement constraint. As its name implies, the ECP is a constraint on empty categories (i.e. gaps), whereas subadjacency is generally considered a constraint on movement operations. Furthermore, the moved constituent in question constructions, such as sentence (1) in Table 1, was achieved by wh-movement. The moved constituent in sentence

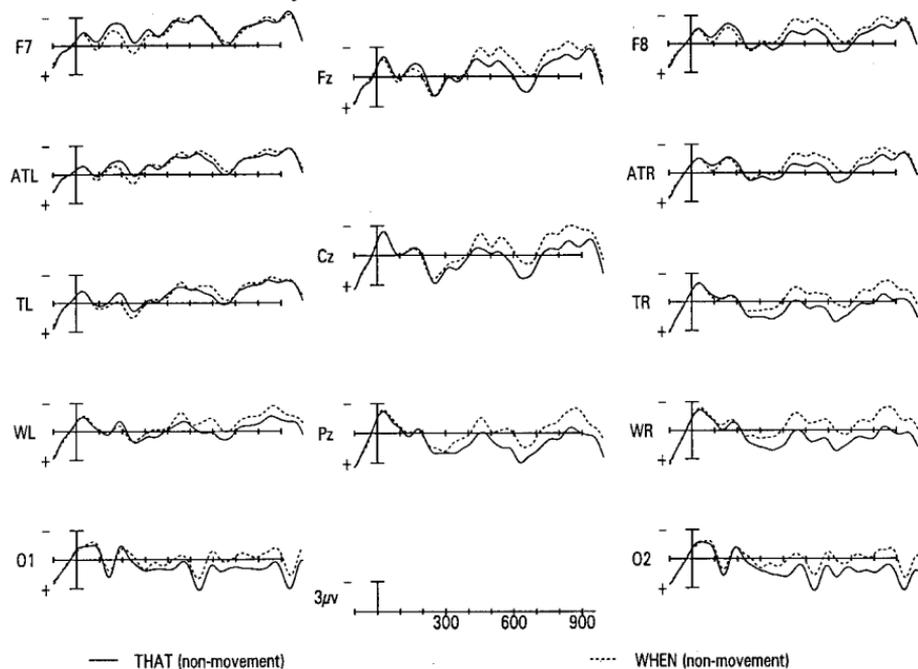


FIG. 4. Grand-average ERPs to *that* and *when* in the non-movement sentences.

⁶A reviewer noted that if one overplots Figs 1 and 2 (overplotting the response to the critical word in all four sentence types), at many electrode sites the response to the subadjacency-violating "when" condition is similar to that elicited by the two "that" conditions. The outlier appears to be the response to the subadjacency control "when" condition, which is more negative-going than the other conditions. Given the apparent word-class differences noted here and by Kluender and Kutas (1993a), it appears that the ERP response to *when* will be more negative-going than the response to *that*, all things being equal. However, the anomaly caused a positive shift in the waveform such that it approximated the response to the two *that* conditions, at least at some electrode sites. Whether or not this is coincidental is a matter for further study.

(2) in Table 2 was displaced via NP-movement (for more on this distinction, see Chomsky, 1981). The two types of movement exhibit characteristic features. For example, *wh*-movement allows a constituent to move an unbounded distance from its canonical position, whereas NP-movement is typically limited to the immediate vicinity of the clause from which the movement originated. The moved constituent in subjacency-violating sentences sits in a *non-argument* position (COMP) within the sentence. This constitutes a salient cue that a gap will be present at a later point in the sentence. Such information potentially could be used by the parser to anticipate a missing element (Nicol, 1988), perhaps initiating processes that actively search for a gap. In contrast, the ECP applies to cases in which the moved constituent sits in an *argument* position (e.g. subject position); in such cases, the parser does not have advance warning that a gap will be encountered.

Finally, the materials in Experiment 1 were somewhat long and complex. Shorter and less complex sentences were presented in Experiment 2.

Method

Subjects. Sixteen undergraduate and graduate students (11 females, 5 males), none of whom had participated in Experiment 1, participated for course credit or for a small compensation. The subjects' age ranged from 18 to 31 (mean = 20) years. All subjects were right-handed native speakers of English.

Materials. Sixty sentence pairs similar to (1) and (2) in Table 2 were constructed. One version of each pair contained two raising verbs, each with a purely grammatical subject *it* [sentence (1)]. The second version contained a double-raising construction which constituted an ECP violation [sentence (2)]. For both versions of each pair, the main clause consisted of *seems* preceded by either *it* (version 1) or a definite noun phrase (version 2). Both versions also contained a clausal complement headed by the complementiser *that*. In ECP-violating sentences, the second clause was followed by an infinitival clause containing a gap in subject position. In these cases, the subject gap was not in a proper structural relationship with its antecedent (or any other licensing element), resulting in a violation of the ECP. A sample of the experimental sentences is provided in Appendix 2. In addition to the ECP-violating sentences, 60 single clause sentence pairs were

TABLE 2
Examples of Sentences Presented During Experiment 2

-
1. It seems *that* it is likely that the man will win.
 2. *The man_i seems *that* it is likely ______i to win.
-

included, 30 of which ended in semantically appropriate words, and 30 of which ended in semantically inappropriate words. Sixty sentences (from another experiment) were used as fillers. The filler sentences contained a matrix clause, followed by an embedded clause introduced by the complementiser *that*, 30 of which were grammatical and 30 of which were ungrammatical. These materials were used to create two stimulus lists in the manner described for Experiment 1. Each subject saw 180 sentences, 90 of which contained an anomaly.

Procedure. All procedures were identical to those used in Experiment 1.

Results and Discussion

Acceptability Judgements

The subjects judged the ECP-violating and grammatical sentences to be acceptable on 18 and 63% of the trials, respectively [$F(1,15) = 61.22, P < 0.0001$]. The low acceptability ratings for the “grammatical” sentences is probably due to the fact that the multiple-raising verb construction is relatively rare.

Event-related Brain Potentials

ERPs to Critical Words. Grand-average ERPs elicited by the complementiser *that* in the two sentence conditions are shown in Fig. 5 (15% of these trials were rejected for artifact). This word represents the earliest possible indication that the ECP-violating sentences were not well formed. Figure 5 reveals that these words elicited a widely distributed positive-going wave with an onset at about 250 msec in the ECP-violating sentences, relative to controls. This positivity was evident throughout the recording epoch. ANOVAs on mean amplitude within the 150–300 msec window found no reliable differences between conditions. However, reliable effects were observed within the 300–500 msec window. Specifically, the response to the critical word in the ECP-violating sentences was more positive-going than the response to controls [midline: $F(1,15) = 7.54, P < 0.05$; lateral: $F(1,15) = 4.65, P < 0.05$]. The results of an ANOVA performed on mean voltage within the 500–800 msec window found reliable effects of sentence type at midline sites [$F(1,15) = 4.84, P < 0.05$], but not at lateral sites [$F(1,15) = 2.32, P > 0.05$].

In order to determine more precisely the onset of divergences in the ERPs to movement-violating and control sentences, ANOVAs were performed on successive 50 msec segments beginning at 100 msec. As in Experiment 1, two sets of analyses were performed using 100 and 150 msec baselines. In both

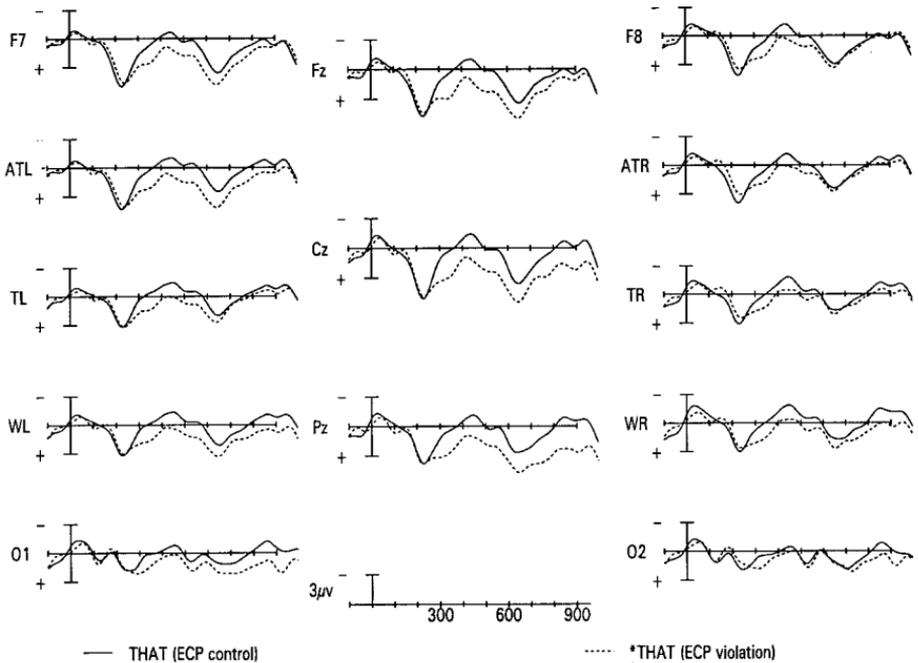


FIG. 5. Grand-average ERPs to *that* in the ECP-violating and non-violating sentence conditions.

baseline conditions, reliable differences between conditions first emerged within the 300–500 msec window [100 msec baseline: $F(1,15) = 8.65$, $P < 0.02$; 150 msec baseline: $F(1,15) = 8.89$, $P < 0.01$].

These results are similar to those reported in Experiment 1, and are consistent with the notion that the parser employs information about the well-formedness of the sentence in a fashion which is remarkably rapid.

ERPs to Sentence-final Words. Previous work has indicated that sentence-final words in sentences typically judged to be unacceptable elicit an enhanced N400-like response, relative to the final words in sentences that are typically judged to be acceptable. Examination of the ERPs to sentence-final words in Experiment 1 was precluded due to the large number of trials rejected for artifact. ERPs to the sentence-final words in the two sentence types presented during Experiment 2 are shown in Fig. 6. ERPs to the sentence-final words in the ECP-violating sentences were indeed more negative-going than those to the same words in well-formed sentences, with the differences between conditions peaking at about 400 msec. ANOVAs on mean amplitude within a 300–500 msec window found reliable differences, both at midline [$F(1,15) = 18.04$, $P < 0.001$] and at lateral [$F(1,15) = 5.36$, $P < 0.05$] sites.

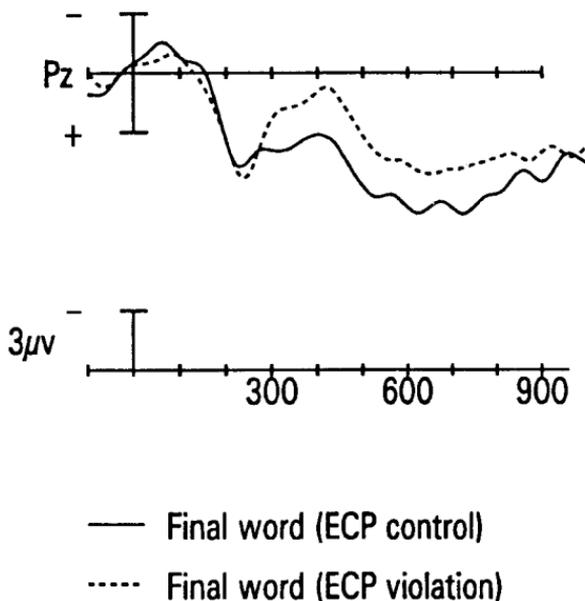


FIG. 6. Grand-average ERPs to sentence-final words in the ECP-violating and non-violating sentences.

GENERAL DISCUSSION

In two experiments, violations of constraints on movement (subjacency in Experiment 1 and the empty category principle in Experiment 2) elicited a widely distributed positive shift. This effect was similar across the two experiments, despite significant differences in the types of movement operations responsible for the anomalous structures (e.g. *wh*-movement vs *NP*-movement). Furthermore, the ERP response was remarkably rapid, beginning within 300 msec following onset of the first word that indicated a violation.

The primary goal of these experiments was to determine whether violations of constraints on constituent movement elicit an ERP response comparable to that elicited by other types of syntactic anomaly. The positive shift elicited by these violations was similar to the P600 effect previously observed in response to a variety of syntactic anomalies (Hagoort et al., 1993; Neville et al., 1991; Osterhout & Holcomb, 1992, 1993), but quite distinct from the N400 effect elicited by semantically anomalous words (Kutas & Hillyard, 1980a). The finding that subjacency violations elicit a positive shift represents an important replication of Neville and co-workers' (1991) observation, but is at odds with the findings of Kluender and Kutas (1993a). Kluender and Kutas failed to observe consistent, reliable differences between the response to subjacency violations and well-formed

control sentences. Although the correct explanation for this discrepancy is uncertain at present, one possibility concerns differences in materials across experiments. In the present study, the critical comparisons were made between identical words, and these words were preceded by identical words across conditions. In the study by Kluender and Kutas, the critical words were different lexical items across conditions, and these critical words were, in some cases, preceded by different words across conditions.

Another explanation for this discrepancy centres on the potentially confounding effects of the sentence-acceptability judgement task. Subjects performed this task in Neville and co-workers' study and in the present study. In the study by Kluender and Kutas, subjects passively read sentences and then made judgements about whether an end-of-sentence target word had occurred in the immediately preceding sentence. This raises the possibility that the positive shift might be a member of the P300 family of positivities known to be elicited by unexpected, task-relevant stimuli (Donchin, 1981; Duncan-Johnson & Donchin, 1977). Clearly, when subjects were performing a sentence-acceptability judgement the constraint-violating words were highly task-relevant, whereas the corresponding well-formed control words were less so. However, prior work has shown that other types of syntactic violations (e.g. agreement and phrase structure violations) elicit a P600-like positive shift even under conditions in which subjects passively read sentences (Hagoort et al., 1993; Osterhout & Mobley, 1995). Furthermore, recent research in our laboratory has provided preliminary indications that the P600-like positivities elicited by syntactic anomalies are distinct from (and at least partly independent of) the P300-like positivities elicited by unexpected, task-relevant stimuli that do not involve a syntactic rule violation (Osterhout, McKinnon, Bersick, & Corey, in press). Thus, it seems unlikely that the effects observed here are entirely or primarily a function of the sentence-acceptability judgement task.

A second goal of the experiments reported here was to determine how quickly violations of movement constraints are detected by the processing system. The present findings indicate that the ERP response to movement violations has a remarkably early onset, as indexed by diverging waveforms in the violating and control conditions. Not only was the electrophysiological response to such violations evident at the first word which introduced the anomaly, but the onset of the effect began at about 250 msec following presentation of the critical word in the subagency violation condition and at about 300 msec in the ECP violation condition. The time-course of this effect is only marginally slower than recent estimates of the speed of lexical access (cf. Marslen-Wilson & Tyler, 1980; Sabol & DeRosa, 1976) and at least as early as previous reports of the ERP response to phrase structure anomalies (Neville et al., 1991; Osterhout & Holcomb, 1992). The remarkable earliness

of this response would seem to suggest that the parser applies movement constraints during the earliest stages of analysis, perhaps during an initial structure-building stage (cf. Mitchell, 1989). Such a conclusion is seemingly at odds with recent claims that the processor posits traces even when the resulting dependency between the trace and the moved constituent violates a constraint on movement (Pickering et al., 1994). It is conceivable that the parser hypothesises the existence of a trace even when information is available indicating that such an hypothesis is ungrammatical; that is, that the processes that posit traces are “modular” with respect to the processes that apply movement constraints.

Another explanation for the earliness of the response to movement violations takes the form of a more general criticism of the current study: Perhaps the claim that movement constraints are detected very rapidly will not generalise beyond the specifics of the experiments reported here. In particular, perhaps the slow rate of presentation (relative to normal language comprehension) and the task demands associated with the sentence-acceptability judgement task are precisely those conditions that encourage the early application of such constraints. Whether or not these results will generalise to other experimental conditions remains an important question for investigation. (Indeed, we are currently replicating this research with continuous natural speech stimuli under conditions in which subjects do not make explicit sentence-acceptability judgements.) However, we believe that the current results are not simply an artifact of the experimental conditions employed here. At the very least, our results indicate that the parser is structured in such a way that information about movement constraints is available at very early stages of processing, at least under some conditions. This result implies that there is no structural impediment within the processing system that would prevent the parser from accessing such information. Of course, it remains to be seen whether this information is consistently exploited in all situations; indeed, the results of Pickering et al. would seem to indicate that they are not. At present, the evidence seems to indicate that, at least under certain conditions, the parser is immediately aware upon entering a syntactic island that a rule governing constituent movement has been violated.

In fact, it turns out to be a considerable challenge to develop a two-stage parser (in which structure-building and gap-filling are carried out independently) that does not generate an intractable number of structures. In an interesting proposal, Fong (1990) explored the problem of controlling the combinatorial explosion of possible structures by assuming that the parser need not wait until phrase structure analysis is completed before beginning to apply certain principles. Fong developed a parser that allows one to selectively interleave independent principles (filters) with the process of building structure. On this view, filters are co-routined with the

structure-building operations in such a way that far fewer illicit structures are generated. A particularly interesting feature of Fong's model is the notion that just those filters that are most effective in reducing the amount of posited candidate structures are interleaved with the creation of phrase structure.

One might argue that the P600-like effect, and in particular its early onset, reflects factors other than the detection of a movement violation. A particular worry in this regard is the possibility of a confounding between the effects of syntactic *anomaly* and syntactic *complexity* in Experiment 1. Specifically, sentences containing a movement violation (e.g. "I wonder which of his staff members the candidate was annoyed *that* ...") were syntactically more complex than the non-movement control sentences (e.g. "I wonder whether the candidate was annoyed *that* ..."), and this inequality in complexity preceded the critical word. It is conceivable, therefore, that differences in the ERPs to movement-violating and control words reflected differences in syntactic complexity rather than the presence or absence of a movement violation and, furthermore, that the apparent early response to the movement violation was in fact a differential response to the relative complexity of the sentence segment preceding the critical word. However, such an interpretation seems unlikely. The existing evidence suggests that the P600 is elicited by syntactic anomaly (resulting either from an outright rule violation or from an apparent rule violation engendered by processing strategies adopted by the processing system) rather than syntactic complexity (cf. Osterhout, 1994; Osterhout & Holcomb, 1995). The results of Experiment 1 are fully consistent with this claim. Although the movement and non-movement sentences presented in Experiment 1 clearly differed in complexity prior to the critical word, differences in the ERPs elicited by the two sentence types emerged only after the onset of the movement-violating word; ERPs to the four words immediately preceding the critical word were nearly identical across movement and non-movement sentences (see Fig. 3). Furthermore, no reliable differences at all were observed in the ERPs elicited by grammatical movement sentences and non-movement controls, even though these sentences clearly differed in complexity.⁷

This is not to say that syntactic complexity had no effects on sentence processing. Indeed, the subjects' sentence-acceptability judgements indicated that they had more difficulty processing the well-formed *and* ill-formed movement sentences than they did the non-movement control

⁷Although ERPs to words that preceded the critical word could be compared across sentence types, such comparisons for words that followed the critical word were precluded by excessive rates of artifact rejection. Rejection rates were a function of word position within the sentence such that ERPs to late-occurring words were contaminated far more frequently than were ERPs to early-occurring words. The subjects appeared to have difficulty preventing eye blinks when reading the latter portion of these very lengthy sentences.

sentences. The apparent absence of an ERP “complexity effect” probably reflects the manner in which ERPs are obtained. ERPs represent brain activity averaged over both items and subjects, time-locked to the onset of some stimulus (e.g. word onset). The most observable and robust ERP effects are those that have an invariant temporal relationship to the onset of the stimulus. Apparently, the brain response to a syntactic anomaly is sufficiently invariant and time-locked to the onset of the anomalous word that it remains readily observable in the grand-average ERP. By contrast, the electrophysiological concomitants of sentence complexity (assuming such concomitants exist) might not be as temporally linked across trials and subjects to the onset of a particular word, and hence might be less visible in the grand-average ERP.

Finally, if movement constraints are in fact interleaved with phrase structure constraints, then one might anticipate that violations of movement constraints would elicit a response quite similar to that elicited by word-order violations. The ERP response to these two anomaly types, as evidenced by diverging waveforms between well-formed and ill-formed conditions, did emerge with a similar time-course (i.e. within 300 msec after onset of the anomalous word). However, phrase structure violations have been reported to elicit at least two distinct effects: an early left hemisphere negativity between roughly 200 and 500 msec, and a late positive wave (P600) with an onset around 500 msec and a centro-posterior distribution (Neville et al., 1991; Osterhout & Holcomb, 1992, 1993; but see Hagoort et al., 1993, who observed only the positive shift). In contrast, in the current study, movement violations elicited only the positive wave (see also Neville et al., 1991), with a much earlier onset than that observed in the response to phrase structure violations. Again assuming that distinct brain responses reflect distinct cognitive and neural processes, one might interpret this ERP difference as indicating that the processes that detect and/or respond to violations of movement constraints are not identical to those that detect and/or respond to phrase structure violations, although both types of anomaly appear to engender a similar end state—that is, the elicitation of a P600-like positivity (cf. Neville et al., 1991). Clearly, more work is needed to uncover the significance of these between-anomaly differences.

The present results, taken together with the results of previous studies, indicate that a disparate collection of linguistic anomalies elicits a late positive shift in the ERP, at least under certain experimental conditions. To date, positive shifts have been elicited by anomalies involving phrase structure (Hagoort et al., 1993; Neville et al., 1991; Osterhout & Holcomb, 1992), verb subcategorisation (Osterhout & Holcomb, 1992; Osterhout et al., 1994) and agreement (Hagoort et al., 1993; Osterhout & Mobley, 1995; Osterhout et al., in press). Seemingly, the primary property shared by these anomalies is that they involve either an actual violation or an apparent

violation (due, for example, to the misanalysis of a syntactically ambiguous string) of a posited formal grammatical rule. Importantly, the P600 effect is qualitatively distinct from the N400 effect elicited by semantically/pragmatically inappropriate words. A reasonable hypothesis supported by these results is that the P600 acts as a general electrophysiological marker of syntactic anomaly (at least under certain experimental conditions), one that is qualitatively distinct from the response to certain types of semantic and pragmatic anomalies.

One objection to the claim that the P600 is a general marker of syntactic anomaly is that there is notable variation in the component characteristics of the effect across experiments. Hence, this effect might not represent a unitary ERP effect. Most notable in this regard are the apparent inconsistencies in the scalp distribution of these positivities across experiments and anomaly types. The positive shift has been variously reported to be largest posteriorly (Osterhout & Mobley, 1995; Osterhout & Holcomb, 1993; Osterhout et al., 1994), evenly distributed across the scalp (present study), or largest over anterior regions (Osterhout & Holcomb, 1992). Differences in distribution of the positivities elicited by subjacency-violating and ECP-violating words were also observed in the present study. Given that scalp distribution is a defining characteristic of a component, it is unclear whether or not these positivities reflect an identical set of cognitive and neural events. The positive shift observed here also had an earlier onset than that observed in previous reports; the typical onset of P600-like positivities is at ~ 500 msec after the presentation of the anomalous word. A likely explanation for the difference between our findings and previous research is the difference in the rate of stimulus presentation in the various studies. In the two experiments reported here, the stimulus onset asynchrony (SOA) between word onsets was 400 msec, whereas in prior research the SOA has ranged from 500 to 650 msec.

A second challenge to the claim that syntactic anomalies, in general, elicit a P600-like positive shift, comes from recent studies reporting a negative-going wave in the response to syntactic anomalies, rather than a P600. For example, Münte et al. (1993) recorded ERPs to the second word of word pairs. In the semantic conditions, these pairs were either semantically related ("gangster-robber") or unrelated. In the syntactic condition, the pairs were either "grammatical" (e.g. "you-spend") or "ungrammatical" (e.g. "your-write"). Target words in the semantically unrelated condition elicited a negative-going wave with a centro-parietal distribution (the N400 effect), whereas targets in the ungrammatical condition elicited a negative-going wave largest over anterior regions. A similar result was obtained by Rösler et al. (1993), who asked subjects to read sentences, some of which ended in either a semantically or syntactically anomalous word. Their subjects made lexical decisions to the critical words. The proper

interpretation of these effects, and the caveats they might necessitate with respect to the claim that syntactic anomalies elicit a positive shift, are not yet clear. For example, word pairs might not fully engage the sentence processing system, and we do not know for certain what effects the lexical decision task might have on the response to syntactically anomalous words. Another possible factor involves the effects of the critical word's position within the sentence. In general, studies reporting a predominantly positive-going response have presented the critical word embedded within the sentence, whereas those reporting a predominantly anterior negative-going effect have placed the word in sentence-final position. The possibility exists that by placing the anomaly in sentence-final position, one risks confounding the response to the anomaly with the ERP manifestations of sentence wrap-up, evaluative/decision, and response processes (cf. Osterhout, 1994).

Another important result of the current study was the failure to replicate a previously reported effect associated with holding a moved constituent in memory until the appropriate gap is located. Kluender and Kutas (1993a, b; see also King & Kutas, 1995) reported that words following a moved constituent elicited an increased negativity between about 300 and 500 msec that was largest over anterior regions of the left hemisphere (the LAN effect). For example, they compared ERPs to function words after the complementiser position in sentences containing a moved constituent, compared to sentences without a moved constituent (e.g. "Do you wonder who, *they* caught —, at it by accident?" vs "Do you think that *they* caught him at it by accident?"). Pronouns that followed the moved constituent *who* elicited a LAN effect, relative to the pronouns that followed the complementiser *that*. Kluender and Kutas interpreted the LAN as reflecting storage of the moved constituent in working memory. In the present study (Experiment 1), the displaced constituents were much longer than those used by Kluender and Kutas. Assuming that holding a complex noun phrase in working memory places more demands on cognitive resources than does holding a simple wh-word, one might predict a larger LAN in our study than that reported by these other studies. However, we failed to find any evidence of a LAN effect in the current study. Although this finding is seemingly inconsistent with the hypothesis that the LAN effect reflects working memory load engendered by moved constituents, the proper reconciliation of the present results with those reported by Kluender and Kutas is a matter that requires further research.

In summary, violations of two types of constraint on constituent movement elicited a positive-going brain response. This positive shift was similar to the "P600" previously reported to be elicited by a variety of syntactic violations, including violations of constraints on verb subcategorisation, phrase structure and agreement. Conversely, this effect was qualitatively distinct from the N400 response to semantically/

pragmatically anomalous words. Further, the response to the movement violations was remarkably rapid, beginning within 300 msec subsequent to presentation of the constraint-violating word, suggesting that movement constraints are applied during an early stage of sentence processing.

Manuscript received December 1994
Revised manuscript received April 1996

REFERENCES

- Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (Eds), *Festschrift for Morris Halle*, pp. 232–286. New York: Holt, Rinehart and Winston.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.
- Crain, S., & Fodor, J.D. (1987). Sentence matching and overgeneration. *Cognition*, 26, 123–169.
- Donchin, E. (1981). Surprise! ... Surprise? *Psychophysiology*, 18, 493–513.
- Duncan-Johnson, C.C., & Donchin, E. (1977). On quantifying surprise: The variation in event-related potentials with subjective probability. *Psychophysiology*, 14, 456–567.
- Epstein, S.D. (1991). *Traces and their antecedents*. New York: Oxford University Press.
- Fong, S. (1990). *Computational properties of principle-based grammatical theories*. Unpublished doctoral dissertation, MIT, Cambridge, MA.
- Forster, K.I. (1987). Binding, plausibility, and modularity. In J.L. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding*, pp. 63–82. Cambridge, MA: MIT Press.
- Frazier, L. (1987). Sentence processing: A tutorial. In M. Coltheart (Ed.), *Attention and performance XII*, pp. 559–585. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210.
- Freedman, S.A., & Forster, K.I. (1985). The psychological status of overgenerated sentences. *Cognition*, 19, 101–131.
- Friedman, D., Simson, R., Ritter, W., & Rapin, I. (1975). The late positive component (P300) and information processing in sentences. *Electroencephalography and Clinical Neurophysiology*, 38, 255–262.
- Garney, S. (1993). Event-related brain potentials in the study of language: An introduction. *Language and Cognitive Processes*, 8, 337–356.
- Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift as an ERP measure of syntactic processing. *Language and Cognitive Processing*, 8, 439–483.
- Hillyard, S.A., & Picton, T.W. (1987). Electrophysiology of cognition. In F. Plum (Ed.), *Handbook of physiology, Section 1: Neurophysiology*, pp. 519–584. New York: American Physiological Society.
- Huang, J. (1982). *Logical relations in Chinese and the theory of grammar*. Unpublished doctoral dissertation, MIT, Cambridge, MA.

- Jasper, H.H. (1958). Report to the Committee on Methods of Clinical Examination in Electroencephalography. Appendix: The ten–twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.
- King, J., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*, 376–395.
- Kluender, R., & Kutas, M. (1993a). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*, 573–633.
- Kluender, R., & Kutas, M. (1993b). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196–214.
- Kutas, M., & Hillyard, S.A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*, 99–116.
- Kutas, M., & Hillyard, S.A. (1980b). Reading between the lines: Event-related brain potentials during natural sentence processing. *Brain and Language*, *11*, 354–373.
- Kutas, M., & Hillyard, S.A. (1980c). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Kutas, M., & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.
- Lasnik, H., & Saito, M. (1984). On the notion of proper government. *Linguistic Inquiry*, *15*, 235–289.
- Lasnik, H., & Saito, M. (1992). *Move-alpha*. Cambridge, MA: MIT Press.
- Marslen-Wilson, W.D., & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*, 1–71.
- Mitchell, D. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, *4*, 123–154.
- Münte, T.F., Heinze, H., & Magnum, G.R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, *5*, 335–344.
- Neville, H.J., Kutas, M., Chesney, G., & Schmidt, A.L. (1985). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language*, *25*, 75–92.
- Neville, H.J., Nicol, J., Barss, A., Forster, K.I., & Garrett, M.F. (1991). Syntactically based processing classes: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *3*, 151–165.
- Neville, H.J., Mills, D., & Lawson, D. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, *2*, 244–258.
- Nicol, J. (1988). *Coreference processing during sentence comprehension*. Unpublished doctoral dissertation, MIT, Cambridge, MA.
- Osterhout, L. (1994). Event-related brain potentials as tools for comprehending language comprehension. In K. Rayner, L. Frazier, & C. Clifton, Jr (Eds), *Perspectives on sentence processing*, pp. 15–44. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Osterhout, L., & Holcomb, P.J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.
- Osterhout, L., & Holcomb, P.J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*, 413–438.
- Osterhout, L., & Holcomb, P.J. (1995). Event-related brain potentials and language comprehension. In M.D. Rugg & M.G.H. Coles (Eds), *Electrophysiology of mind: Event-related potentials and cognition*, pp. 171–215. Oxford: Oxford University Press.
- Osterhout, L., & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*, 739–773.
- Osterhout, L., Holcomb, P.J., & Swinney, D.A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 786–803.

- Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (in press). On the language-specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*.
- Pickering, M., Barton, S., & Shillcock, R. (1994). Unbounded dependencies, island constraints, and processing complexity. In C. Clifton, Jr, L. Frazier, & K. Rayner (Eds), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Rösler, F., Friederici, A., Putz, P., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience*, 5, 345–362.
- Sabol, M.A., & DeRosa, D.V. (1976). Semantic encoding of isolated words. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 58–68.
- Stowe, L. (1986). Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, 1, 227–245.

APPENDIX 1

Sample materials for Experiment 1

- 1a. I wonder which of his staff members the candidate was annoyed when/that his son was questioned by.
- b. I wonder whether the candidate was annoyed when/that his son was questioned by one of his staff members.
- 2a. I wonder which bread the baker was surprised when/that the customers didn't like.
- b. I wonder whether the baker was surprised when/that the customers didn't like the rye bread.
- 3a. I wonder which bills the bus driver was relieved when/that his company paid.
- b. I wonder whether the bus driver was relieved when/that his company paid the insurance bills.
- 4a. I wonder which theater your neighbor was upset when/that the city announced plans to demolish.
- b. I wonder whether your neighbor was upset when/that the city announced plans to demolish his favorite theater.
- 5a. I wonder which shares the broker was afraid when/that his partner had gotten rid of.
- b. I wonder whether the broker was afraid when/that his partner has gotten rid of the oil shares.
- 6a. I wonder which mountain your friend will be astonished when/that his father has succeeded in climbing.
- b. I wonder whether your friend will be astonished when/that his father has succeeded in climbing Mt. Everest.
- 7a. I wonder which dress the guests at the wedding were shocked when/that the bride wore.
- b. I wonder whether the guests at the wedding were shocked when/that the bride wore a red dress.
- 8a. I wonder which movie actor the lawyer was pleased when/that his daughter went out with.
- b. I wonder whether the lawyer was pleased when/that his daughter went out with the movie actor.
- 9a. I wonder which procedure the surgeon was indignant when/that the hospital condemned.
- b. I wonder whether the surgeon was indignant when/that the hospital condemned the procedure.

- 10a. I wonder which scene the movie star was overjoyed when/that the director cut from the film.
- b. I wonder whether the movie star was overjoyed when/that the scene was cut from the film.
- 11a. I wonder which fossil the anthropologist was glad when/that the students found.
- b. I wonder whether the anthropologist was glad when/that the students found the bird fossil.
- 12a. I wonder which routine the acrobats were frustrated when/that the horse trainer spoiled.
- b. I wonder whether the acrobats were frustrated when/that the horse trainer spoiled their routine.
- 13a. I wonder which problem the car dealer was dumbfounded when/that the mechanic fixed.
- b. I wonder whether the car dealer was dumbfounded when/that the mechanic fixed the problem.
- 14a. I wonder which treatment the doctor was overwhelmed when/that the patient recovered under.
- b. I wonder whether the doctor was overwhelmed when/that the patient recovered under the risky treatment.
- 15a. I wonder which film the photographer was intrigued when/that the chemical ruined.
- b. I wonder whether the photographer was intrigued when/that the chemical ruined the film.

APPENDIX 2

Sample materials for Experiment 2

- 1a. The teacher seems that it is likely to give a test.
- b. It seems that it is likely the teacher will give a test.
- 2a. The governor seems that it is unlikely to attend the meeting.
- b. It seems that it is unlikely the governor will attend the meeting.
- 3a. The biologist seems that it is believed to save many dolphins.
- b. It seems that it is believed the biologist can save many dolphins.
- 4a. The physician seems that it is understood to donate the new medicine.
- b. It seems that it is understood the physician donates the new medicine.
- 5a. The musician seems that it is sure to perform live tonight.
- b. It seems that it is sure the musician will perform live tonight.
- 6a. The farmer seems that it is likely to grow the pumpkins.
- b. It seems that it is likely the farmer will grow the pumpkins.
- 7a. The senator seems that it is unlikely to win the primary.
- b. It seems that it is unlikely the senator will win the primary.
- 8a. The dancer seems that it is thought to practice hours daily.
- b. It seems that it is thought the dancer should practice hours daily.
- 9a. The plumber seems that it is known to fix the pipe.
- b. It seems that it is known the plumber will fix the pipe.
- 10a. The carpenter seems that it is certain to build the house.
- b. It seems that it is certain the carpenter will build the house.
- 11a. The dealer seems that it is argued to lower his prices.
- b. It seems that it is argued the dealer was lowering his prices.
- 12a. The photographer seems that it is bound to shoot the scene.
- b. It seems that it is bound the photographer will shoot the scene.
- 13a. The policeman seems that it is apt to give the address.
- b. It seems that it is apt the policeman will give the address.

- 14a. The librarian seems that it is known to quit her job.
 - b. It seems that it is known the librarian might quit her job.
- 15a. The ranger seems that it is feared to change the policy.
 - b. It seems that it is feared the ranger will change the policy.

