2 Car talk, car listen

Kathryn Bock, Gary S. Dell, Susan M. Garnsey, Arthur F. Kramer, and Tate T. Kubose
University of Illinois at Urbana-Champaign, USA

Among age-old speculations about the origins of human language is a suggestion that use of the oral channel evolved to leave the hands and the rest of the body free for other activities. Whether this is true or not, there are few human activities that proceed unaccompanied by talking or listening. With the scope of everyday activities expanding to include some that are life-threatening unless performed with sufficient care, questions about the demands of language use on attention to other things have assumed new priority. This new priority runs up against an old, unresolved, but theoretically central psycholinguistic debate over how language production and language comprehension are related to each other. The terms of this relationship involve shared or divided components of linguistic knowledge and shared or divided resources of perceptual, motor, and cognitive skill. Our question in this chapter is how production and comprehension differ in their demands on attention or working memory (Baddeley, 2003), as reflected in how much production and comprehension interfere with other tasks. At bottom, we are interested in whether talking is harder than listening.

The automatic answer to this question seems to be "of course". Talking involves retrieving linguistic information from memory and assembling utterances in working memory, using only information from the intended meaning; listening benefits from specific form cues and accompanying recognition processes. The retrieval challenges for speakers are imposing: The estimated vocabulary of an educated English speaker is over 45,000 words. Failures of retrieval are common enough to have inspired a famous observation of William James (1890) about tip-of-the-tongue states and a large research literature on the phenomenon, beginning with Brown and McNeill (1966). Assembling utterances entails structuring and ordering the words into one of the infinite number of possible phrases or sentences of English, drawing on the 12,000-or-more syllables of which English words are composed, and articulating the syllables at a rate of three or four per second, using more muscle fibres than any other mechanical performance of the human body (Fink, 1986). It is ability to speak a foreign language, not ability to understand one, that constitutes the commonplace standard of bilingual competence: Every word uttered, indeed every sound, is on display for evaluation. Talking to an audience, public speaking, is many people's greatest fear. Public listening inspires no such terror.

This unreflective view of the challenges of language production overlooks the more subtle but no less vexing challenges of normal language understanding. Normal understanding involves the auditory modality, and audition has its own impressive accounting of peripheral mechanics, in the form of the 1,600,000 or so moving parts of the hair cells in the cochlea. Cognitively, listening to language means segmenting a continuous stream of sound into identifiable components and disambiguating the inevitably ambiguous products at whatever rates they are produced. As adventurous tourists testify, it means dealing with the uninterpretable cascades of sound that greet one's tentative efforts to converse in a foreign language. It is generally easier to produce "ou est la toilette?" than it is to comprehend a fluently spoken response. One hopes instead for a pointed gesture.

Conceivably, intuitions about the greater difficulty of language production are due merely to differences in how readily we can evaluate our own performance. Talking provides multiple opportunities to compare our speech against standards ranging from how we sound and how fluent we are to how adequately we are conveying what we mean to say. After all, speakers know what meaning they intended. All of this is possible because we presumably listen to and comprehend what we say, fairly automatically and unstoppably. In comparison, diagnosing a problem in our comprehension of others' speech requires information to which we have no direct access. We become aware of ambiguity, incomplete understanding, and outright mistakes only when confronted with other evidence that conflicts with an initial interpretation, sometimes in the form of trying to say the same thing ourselves. This raises the suspicion that the major thing that makes production hard is the standard we hold ourselves to.

Still, the automatic reaction that speaking is harder has something going for it empirically. Beyond the patent information-processing challenges of language production, there is evidence that children's early comprehension vocabularies are 3 months ahead of their production vocabularies (Benedict, 1979). For adults, the active or productive vocabulary is vastly smaller than the passive or comprehension vocabulary, with estimates of 30,000 words for the former compared to 75,000 for the latter (Levelt, 1989). Language disorders of all kinds, from stuttering in children to aphasia in adults, are more strongly associated with problems of production than with deficits in comprehension. All of these things are consistent with the intuition that production is the harder and more fragile aspect of language performance.

Of course, these types of evidence are vulnerable to the differential-standards objection. Perhaps children are in principle capable of producing what they comprehend, but their articulatory apparatus stands in the way of making their speech comprehensible. What adults require as evidence of comprehension is simple and mastered much earlier: When infants reliably turn toward the referent of a word, they are likely to be credited with understanding it (Huttenlocher, 1974). Yet, there is next to nothing in the psychological literature, either about human perception and performance or about psycholinguistics, that allows the intuition of greater production difficulty to be quantified with respect to normal adult language use.

Reasons for the absence of rigorous comparisons are easy to come by. Auditory and articulatory events differ enormously at the periphery, with auditory comprehension involving auditory sensory systems and production involving oral motor systems. The events are subserved by different cortical regions in different parts of the human brain. There are differences in what the processing systems have to accomplish that require differences in how they work: Comprehension creates interpretations of utterances, constructing meaning from ambiguous input; production creates utterances, constructing forms from *unambiguous* input (Bock, 1995). Comprehension requires perceptual analysis to segment and recognize words and parse structures; production requires conceptual synthesis to recall and assemble words and structures. Finally, the demanding events seem to occur at different points in time, with the demands of comprehension peaking near the ends of utterances and the demands of production peaking near the beginnings, even prior to speech onset. These differences are a challenge to standards of experimental control on a par with the proverbial comparison of apples and oranges.

Dual-task comprehension and production

Because of the many intrinsic differences between comprehension and production, in the present work we applied dual-task logic to the question of their relative difficulty. Despite their pitfalls (Jonides & Mack, 1984), dual-task methodologies create a yardstick on which otherwise incommensurate types of cognitive performances can be roughly compared.

There have been a few controlled comparisons of the dual-task effects of speaking and hearing single words (Shallice, McLeod, & Lewis, 1985). The results of this work suggest that the systems for articulation and audition are separable and impose different demands (see also Martin, Lesch, & Bartha, 1999). Missing, however, are systematic comparisons of comprehension and production of connected speech, which is what normal adult language use consists of. Even when adults speak to babies, fewer than 10% of their utterances consist of isolated words (van de Weijer, 1999).

We know that hearing or reading a sentence interferes with performance on secondary tasks such as detecting a nonlinguistic auditory stimulus (Garrett, Bever, & Fodor, 1966; Holmes & Forster, 1970) or making a lexical decision to an unexpected letter string (Shapiro, Zurif, & Grimshaw, 1987). Importantly, this interference varies as a function of the sentence's complexity (Gordon, Hendrick, & Levine, 2002) and where in a sentence the secondary task is performed (Ford, 1983). Similar evidence of difficulty arises during speaking, but the timing differs. Ford and Holmes (1978) required speakers to respond with a button press to tones that occurred at unexpected times during extemporaneous monologues about general topics such as "family life". Response times were slower at ends of clauses only when the current clause was followed by at least one additional clause, suggesting that the planning or preparation of upcoming speech is more distracting than current execution. In spontaneous speech, pauses and hesitations tend to occur earlier rather than later in utterances (Clark & Wasow, 1998; Maclay & Osgood, 1959), suggesting that problems arise nearer the beginning than the end of speech, consistent with the preparation hypothesis.

A more rigorous illustration of the timing of production demands can be found in work by Bock, Irwin, Davidson, and Levelt (2003). They used a controlled task that elicited utterances of different complexity as descriptions of identical scenes, and speakers' eye movements were monitored as they described the scenes. The results showed increases in several measures of preparation at earlier utterance positions. More complex utterances were accompanied by longer speech onset latencies as well as increased eye-voice spans and longer gaze durations to parts of scenes mentioned earlier in the utterances.

The finding that production takes its toll early in utterances stands in contrast with long-standing evidence that the ends of sentences and clauses are sites of high comprehension demand (Mitchell & Green, 1978). Bock et al. (2003; see also Bock, Irwin, & Davidson, 2004) proposed a process of utterance *ground breaking* in production, analogous to the phenomenon of clause wrap-up in comprehension. Ground breaking requires that speakers (1) dis-integrate the elements of a scene or mental model from their perceptual or mental contexts, in order to refer to them in an upcoming utterance; and (2) map the elements to the words and structures of language. The harder the dis-integration and mapping, the more difficult the utterance preparation.

Clearly, there are demands of both comprehension and production, but existing research provides little basis for systematic comparison. As a result, it remains possible and perhaps even likely that, held to comparable, objective performance standards, comprehension and production differ relatively little in how much they disrupt concurrent nonlanguage activities. Successful listening and successful speaking may impose similar penalties on performance, with major implications for psycholinguistic theories of language processes. In addition, the issue of the normal demands of language use has increasingly important practical and societal implications for whether and how we prescribe (or proscribe) the ability to engage in conversation.

To address both the theoretical and the practical questions about the demands of different facets of language performance, we have begun to examine how hard speaking and listening are in terms of their comparative consequences for the performance of another common, fairly automatic, and culturally central adult activity, driving an automobile.

A growing body of research has considered the impact of talking on driving in a very specific context. With increasing use of cellular telephones by drivers, there is an urgent need to understand their effects on driving performance (see Alm & Nilsson, 2001, for a review). For the most part, these effects appear to be negative. For instance, epidemiological research has shown that the chances of being involved in an automobile collision are greater for drivers using wireless telephones (Redelmeier & Tibshirani, 1997; Violanti, 1998; Violanti & Marshall, 1996). The initial assumption was that physical manipulation of the telephone, and not actual conversation, presents the greatest danger to driving safety. Recent research has shown that this is not necessarily the case. McKnight and McKnight (1993) showed that both placing calls and holding a conversation resulted in poorer simulated driving performance than no distraction at all. Other studies have shown that hands-free telephones provide little added safety over handheld versions with respect to driving performance (Lamble, Kauranen, Lassko, & Summala, 1999; Strayer & Johnston, 2001). The conclusion from this work is that the major distraction in using cellular phones is not the manipulation of the telephone, but the perceptual or cognitive distractions of conversation.

Taking this research as a springboard, we set out to separate the relative demands of comprehending and producing connected speech on driving performance in a high-fidelity simulator. As in the existing psycholinguistic literature, applied research on driving has not separated comprehension and production difficulty under conditions that match them for the complexity of the linguistic materials. Recarte and Nunes (2003) conducted a remarkable study that ranged over many potential sources of driver distraction, including language comprehension and production, finding that comprehension was often no more disruptive than uninterrupted driving. However, Recarte and Nunes presented drivers with 2-minute passages of speech which the drivers had to recall aloud, with measures of driving performance taken during the comprehension and production (recall) episodes. There were no apparent requirements for recall accuracy, though, which would allow perfunctory comprehension to trade off with laboured, retrieval-intensive production. Without criteria for comprehension accuracy and matching of conceptual and linguistic complexity, it is impossible to tell whether the source of difficulty is the language task itself or the general memory and conceptual processes that accompany the encoding and retrieval of the material.

In our examination of the effects of speech production and speech comprehension on driving performance, we used a dual-task paradigm with simulated driving as the primary task and a language-production or language-comprehension task as the secondary task. That is, we compared driving performance under single-task (i.e., driving only) and dual-task (i.e., driving with a concurrent speech task) conditions, and we also compared driving performance when the secondary task, the concurrent speech task, involved chiefly language production or chiefly language comprehension. The advantage of driving as a primary task, apart from its social and applied significance, is that it offers multiple, continuous performance measures, extending over the same time course as speech itself. The measures include proximal aspects of driving skill, such as steering, braking, and moving the accelerator, as well as the distal consequences of these activities, including maintenance of lane position, following distance, and acceleration.

The secondary language tasks involved producing or understanding statements about the spatial relationships between pairs of buildings on the campus of the University of Illinois at Urbana-Champaign. Normative knowledge of campus buildings and locations was assessed in preliminary studies that determined which buildings could be queried in the task. The participants in the actual talk-and-drive tasks were also tested on their campus knowledge prior to performance of the experimental tasks, and their memory of building locations was refreshed with the map in Figure 2.1.

In the production task, participants were cued to formulate and produce two-clause statements about the buildings. They were given two building names (e.g., Foellinger Auditorium, English Building) and then a map direction (e.g., south). They then had 20 s to say what other campus buildings were closest to the named buildings in the given direction, using a structurally simple, two-clause sentence. For instance, a correct response to "Foellinger

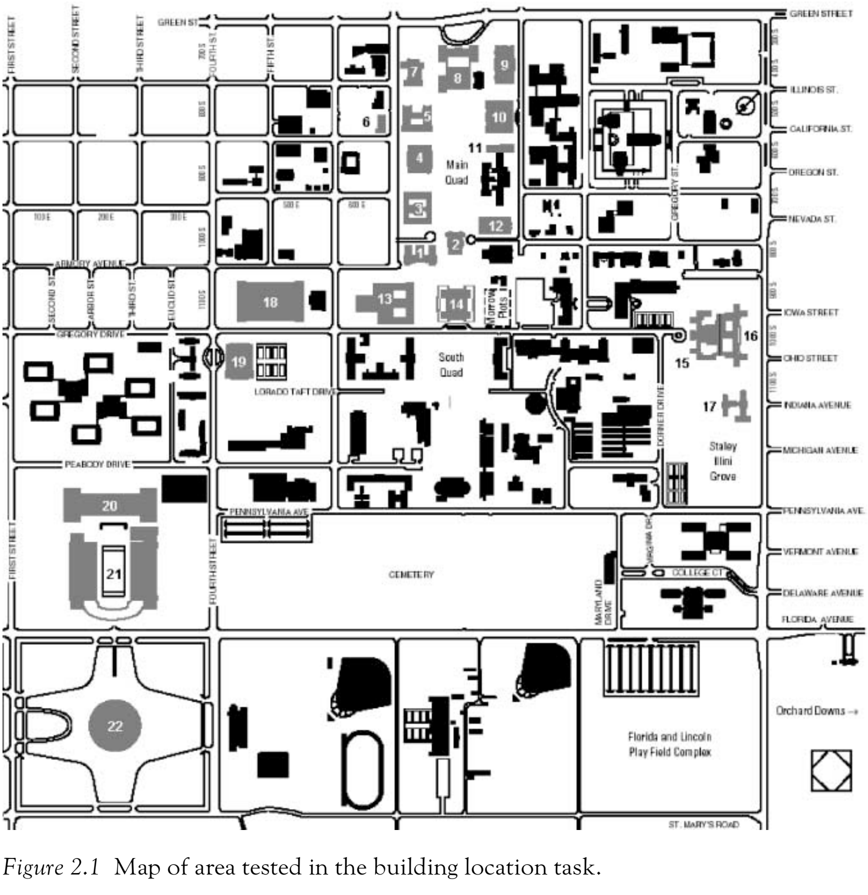

Figure 2.1 Map of area tested in the building location task.

Auditorium, English Building, South" would be "The Undergraduate Library is south of Foellinger Auditorium and Lincoln Hall is south of the English Building." Comprehension used recordings of the actual utterances generated during the production experiment, presented to different participants. The participants listened to these utterances and had to verify whether each clause of the utterance was true. Figure 2.2 displays the sequence of events in the corresponding production and comprehension tasks.

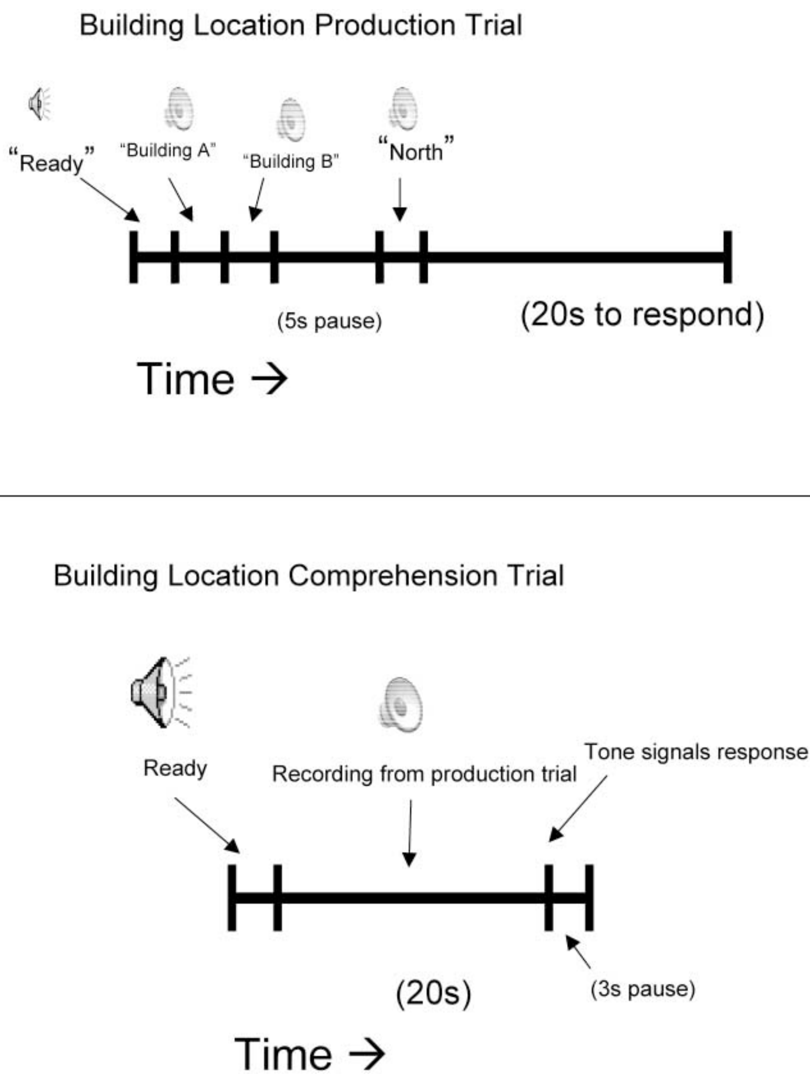

Figure 2.2 Events on production and comprehension trials.

In this way, production and comprehension were closely matched. The produced and comprehended sentences were the same. The same knowledge sources were called upon, and both tasks required participants to deal with exactly two building pairs. Both speech tasks required spatial reasoning of the kind that occurs when giving or receiving directions in the car, in person, or by telephone. This permitted a comparison of speaking and understanding under conditions that ensured similar cognitive processes, processes which are shared by production and comprehension.

The language and driving tasks were performed in the Beckman Institute Driving Simulator. The simulator is a real automobile, an automatic-transmission 1998 Saturn SL, positioned between eight projection screens displaying an integrated 260 view to the front and rear of the vehicle (Figure 2.3). The displays mimic movement through natural environments. The driving scenarios, the movements of the vehicle, and the behaviours of interacting vehicles were simulated with commercial simulation software adapted for the purposes of the experiment. Figure 2.4 shows a simulated environment photographed from outside the vehicle. Seen from inside, the display appears as a continuous scene to the front and rear.

In two experiments, Kubose, Bock, Dell, Garnsey, Kramer, and Mayhugh (2006) examined the effects of speech production and speech comprehension on continuous measures of velocity, lane maintenance, and headway maintenance (maintenance of a safe distance behind a lead vehicle). Simulated wind gusts, variable in strength and direction, made vehicle control more difficult, similar to a technique in a driving task designed by Graham and Carter (2001). In one experiment, driving was easy: Participants simply had to maintain their speed (55 miles per hour) and position in the right-hand lane of a straight, two-lane rural highway with intermittent oncoming traffic, compensating for wind gusts. In the second experiment, the driving task was similar but harder: On the same two-lane road, with the same intermittent oncoming traffic and wind gusts, drivers had to maintain a safe following distance (headway) from a lead vehicle. Headway maintenance was complicated by simulating the erratic slowing and acceleration of a box truck ahead of the driver's vehicle.

Kubose et al. (2006) found that driving was indeed worse under dual-task conditions than under single-task conditions, consistent with other dual-task driving studies (Rakauskas, Gugerty, & Ward, 2004). As Figure 2.5 (top panel) illustrates, in easy driving there was increased variability in velocity when driving was accompanied by a speech task than when driving was the only task, and in the harder driving task (bottom panel), there was increased variability in headway time when speech accompanied driving. However, ~~despite~~ the expectation of greater difficulty due to production than to comprehension, the differences in the effects of the two secondary tasks were negligible. Figure 2.5 shows that the impact of production on driving was no worse than the impact of comprehension on driving. Paradoxically, on one measure, talking actually seemed to improve driving performance

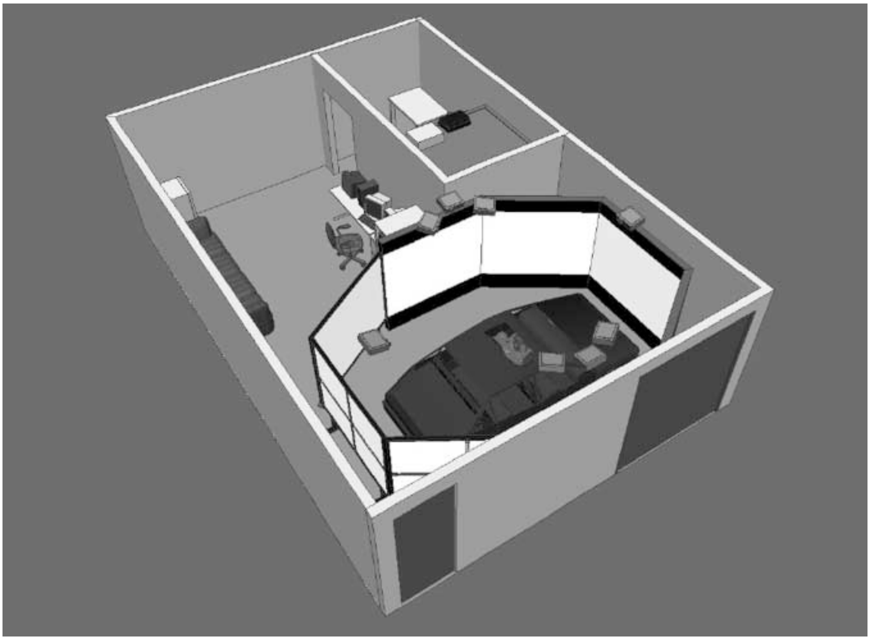

Figure 2.3 Beckman Institute Driving Simulator.

Figure 2.4 A simulated environment photographed from outside vehicle.

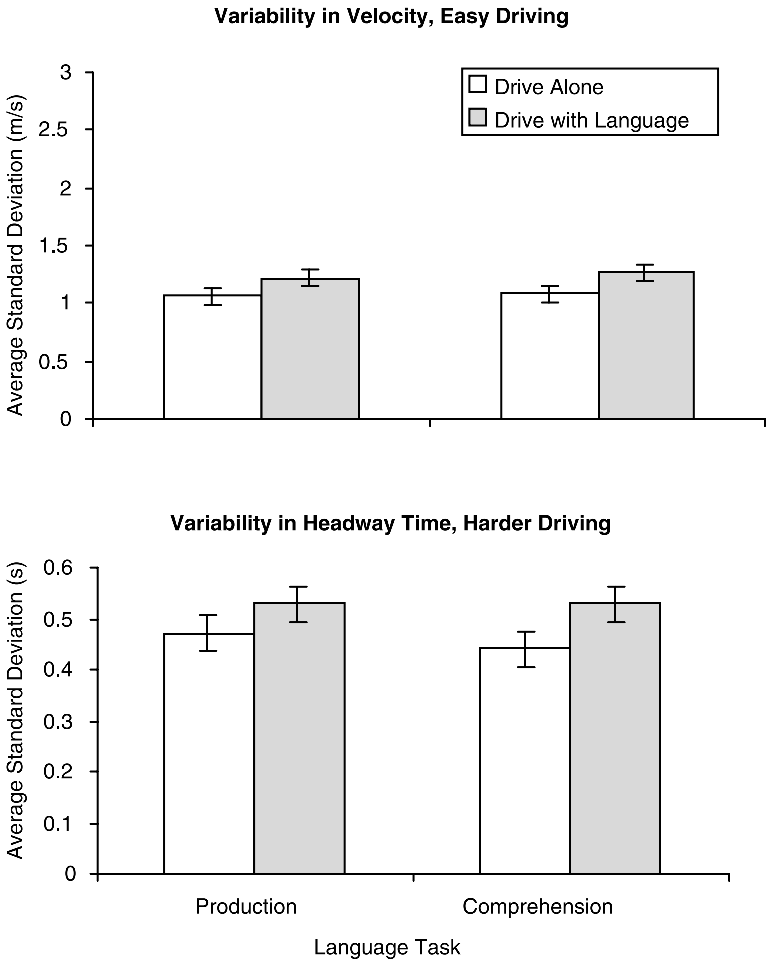

Figure 2.5 Change in variability of driving performance when driving alone or driving while talking (Production) or understanding (Comprehension) (from Kubose et al., 2006).

relative to driving in silence. With an accompanying production task, there was less variability in the maintenance of lane position for easier as well as harder driving. Comprehension, in contrast, had no effect relative to silent driving on either of these measures.

There are a number of ways to interpret these results. Routine driving is a highly automated task for most young adults, and its performance may tend to improve when attention is directed elsewhere. Experts can get worse when they try to focus on the components of their skills (Beilock & Carr, 2001; Beilock, Carr, MacMahon, & Starkes, 2002). This would account for the unexpected reduction in the variability of lane position that accompanied production, and would be consistent with the hypothesis that production demands more attention than comprehension. In light of the generally similar effects of production and comprehension on dual-task performance, however, another interpretation is that when the contents of the language (the messages, the syntax, the words, the phonology, and so on) for comprehension and production are well equated, the attentional demands are very much the same. Obviously, given the inherent disparities between understanding and generating language, this is a claim that deserves a sceptical reception and requires considerable buttressing. In the next section, we inspect it more closely, calling on finer-grained measures of driving performance time-locked to specific aspects of speech events.

Speaking and listening: Breaking ground and wrapping up

In terms of correct performance on the secondary language tasks, Kubose et al. (2006) found few differences between comprehension and production. Overall accuracy of producing and verifying building locations was better than 85%, with only a 1% decrease in accuracy under dual-task conditions. Comprehenders and producers showed similar levels of performance and similarly small reductions when the language tasks accompanied driving, consistent with their generally similar performance profiles.

These global similarities were found on measures recorded over fairly long intervals. The comprehension and production trials on which dual-task performance was assessed could take as long as 20 s; the utterances themselves consumed 8.6 s, less than a half of the whole period, on average. To determine how the peak local demands associated with comprehension and production compared under dual-task conditions, we determined the onsets and offsets of speech on every driving trial, time-locked the driving measures to the speech events, and examined the driving performance changes that accompanied the production and comprehension of the building location statements.

Figures 2.6 and 2.7 depict speech-localized variations in performance during production (top panel) and comprehension (bottom panel). Time 0 on each graph represents the onset of the 20-s data-recording periods. During the ensuing 20 s, the timing of the events during the production and comprehension tasks was as follows. On production trials, the offset of the recorded direction prompts (north, east, etc.) occurred at 7.25 s. The average onset of speaking came 2.74 s later, with the offset of the first clause occurring approximately 3.75 s after that. During these 3.75 s, speakers said things along the lines of "The Undergraduate Library is south of Foellinger Auditorium . . .". The utterance ended after another clause, on average, 8.5 s after speech began. Events on the comprehension trials occurred along the same timescale, with the important exception that the comprehenders did not hear the recorded direction prompts, but only what the speakers said. The onset of the audio for comprehenders occurred 6.25 s after trial onset.

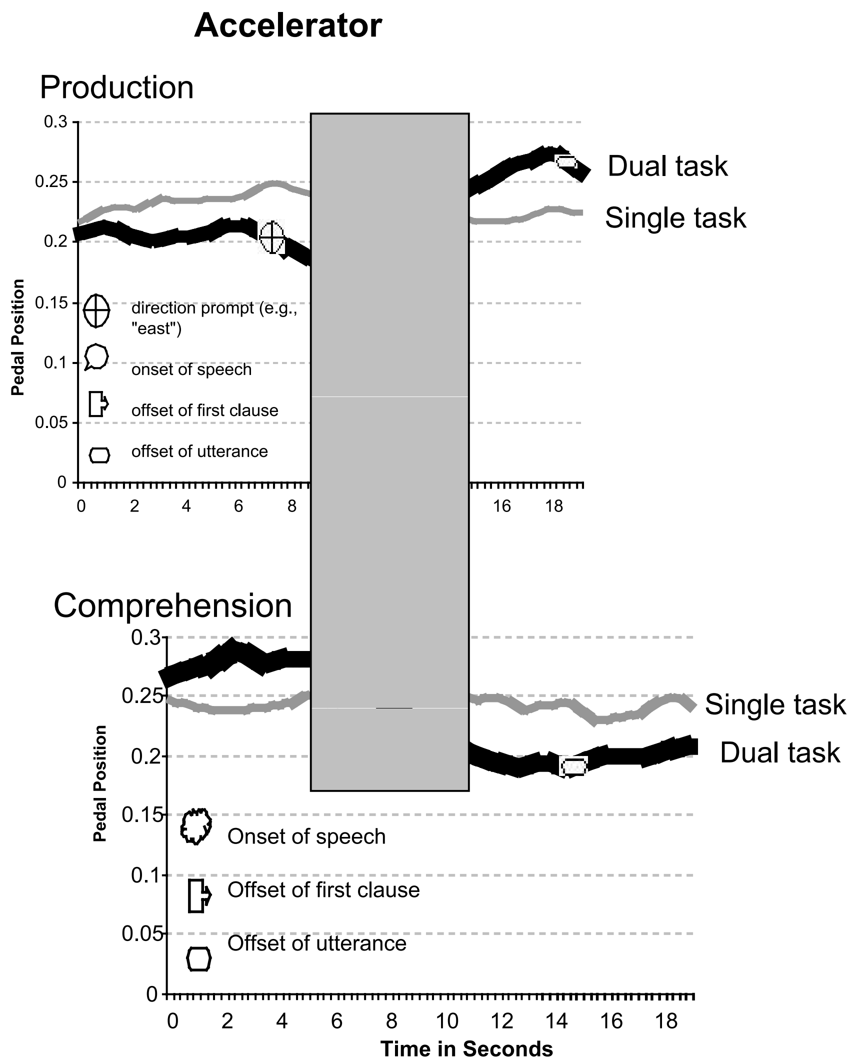

Figure 2.6 Localized changes in driving performance accompanying speech events.

The figures align speech onset across the two tasks, the onset of speaking in the production task and the onset of listening in the comprehension task. The graphed values are aggregates over performance during each quarter-second within the recording period. The heavy black line represents the dual-task condition, and the grey line represents the single-task condition. The boxed areas in each graph highlight the changes in the measures during production and comprehension of the first clauses of the building location descriptions.

The measures shown are direct reflections of the drivers' operation of the

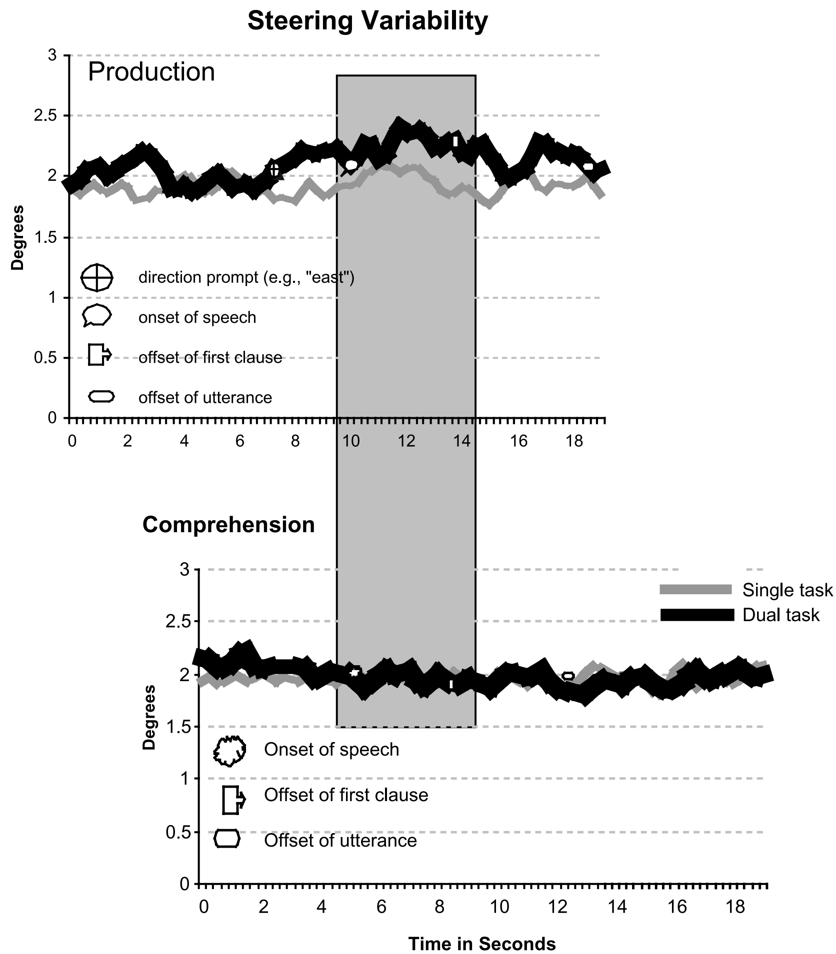

Figure 2.7 Localized changes in steering accompanying speech events.

vehicle's accelerator and steering wheel under the more challenging driving conditions that required headway maintenance. Steering variability is shown in degrees, and accelerator position in a range between 0 and 1, where 0 is no pressure and 1 is maximum pressure ("floored"). Because steering and acceleration both demanded continuous adaptations in order to deal with wind gusts and changes in the speed of the lead vehicle, they give a moment-by-moment picture of how drivers controlled the vehicle's perceived position and speed.

For accelerator position, shown in Figure 2.6, there were localized changes in driving performance accompanying speaking and listening at the expected points of greatest production and comprehension difficulty. Recall that, in production, we anticipated changes in driving occurring close to the onsets of clauses, whereas, in comprehension, we anticipated changes closer to the ends of clauses. The top panel of Figure 2.6 shows that when speakers received the direction prompt that allowed them to begin formulating their utterances, they reduced pressure on the accelerator and continued that reduction until they began speaking, at which point they began to increase pressure and continued increasing it throughout the utterance. For comprehenders, shown in the bottom panel, reduction in accelerator pressure began with the onset of the recorded utterance and continued through the end of the first clause. So, in production, the greatest reduction in accelerator pressure occurs during formulation, just prior to utterance onset; in comprehension, the greatest reduction begins close to the end of the clause and continues through the second clause. Both of these effects were statistically significant ($t(21) = -2.96$, $p < .01$ for production and $t(23) = 5.82$, $p < .001$ for comprehension) and, despite the differences in their timing, the magnitudes of the changes in accelerator position were similar relative to the negligible changes for the single-task controls.

The results for steering in Figure 2.7 revealed a different pattern. In comprehension, shown in the bottom panel, the absence of changes in the comprehenders' performance is consistent with the absence of dual-task effects in the maintenance of lane effects for comprehension. Producers, however, showed increased variability in steering ($t(21) = -2.728$, $p = .013$), beginning with the presentation of the direction prompt—again, the trigger for the onset of utterance formulation. However, in light of the better maintenance of lane position during dual-task than during single-task performance in production, the most viable interpretation of this increased variability is in terms of how successfully speakers dealt with wind gusts. To maintain lane position against unpredictable gusts of wind requires continuous adjustments to steering, and it appears that while they were speaking the drivers in the production condition did this better than when they were only driving.

The local variations in driving performance during speaking and listening disclose that the changes are concurrent with the accompanying speech events, and the nature of the changes is consistent with the broad differences measured over longer intervals. The process of production was accompanied by *improved* lane maintenance, and the changes in steering needed for better lane maintenance were temporally linked to the features of production that make the greatest demands. At analogous times, comprehension had no differential effect on steering. When talking and listening did interfere with driving, the interference came at different times, but the kinds and amount of disruption were similar. The times were linked to the beginnings and ends of clauses, the structures that correspond most closely to "idea units" in language. Apparently, in production, it was the mental ground breaking needed to assemble these units that was disruptive; in listening, it was the mental wrap-up that caused problems.

Implications

The time-course data showed the expected differences in the profiles of difficulty for comprehension and production. They also showed similarities in the amounts of disruption, when disruptions arose. An open, pressing question is how the contents of the language employed contributed to these similarities. The spatial thinking that was required in order to formulate and understand utterances in the building-location tasks may be more taxing, or simply different, than other messages conveyed in language, due the special nature of the disruptions associated with spatial cognition (cf. Irwin & Brockmole, 2004). If so, messages with different contents could have different effects. Although Recarte and Nunes (2003) found very few content-associated variations in driving disruption in a comparison of concrete and abstract language, it is unclear whether their concrete passages called on spatial imagery. But even if other kinds of language impose fewer demands, the importance of spatial cognition to driving and to ordinary communication during driving makes any negative impact of spatial language on comprehension a finding to be reckoned with. It means that following verbal directions, whether from a passenger or from the audio of a telematic device, can be very hard.

As it stands, the tentative conclusion from the present findings is that when the linguistic and conceptual complexity of language is controlled, and accuracy of understanding is required, language comprehension can interfere just as much with driving as talking does. A second conclusion is that when production and comprehension differed in the magnitudes of their effects, production actually served to improve driving performance. There is nothing in these results to suggest substantial disparities in the degree to which speaking and listening interfere with other activities, except in their timing.

Whether we see speaking as being objectively more demanding than understanding depends on the interpretation of the paradox of improved lane position that accompanied speaking in two experiments (Kubose et al., 2006) and the corresponding changes in steering that were time-linked to the events of production. These changes could reflect an improved ability to perform expert, highly practised, routinized skills when attention is drawn away from them. This is compatible with Beilock's (Beilock & Carr, 2001; Beilock et al., 2002) interpretation of how and why conscious attention interferes with skilled performance. If it is right, we would expect to see production interfere more than comprehension with other, less routine, less automatic skills. Likewise, we might expect to see improvements in the performance of some other highly practised skills when they are accompanied by talking. These speculations remain to be tested.

An alternative interpretation of these effects is specific to driving and, perhaps, similarly dangerous activities. Alm and Nilsson (1994) proposed that drivers are sensitive to the problems that talking can cause, and therefore selectively increase attention to the elements of a task that are dangerous. Kubose et al. (2006) tried to test this by changing the most threatening elements of the driving task, but they found the same patterns of results over two experiments. When lane position was the most important factor in safety, drivers maintained lane position better when talking than when not. When headway maintenance was also essential to safety, drivers dropped further behind the lead vehicle while performing the language tasks, regardless of whether the task involved comprehension or production. Only the speakers showed improved lane maintenance at the same time.

So, drivers who were talking stayed more reliably in their lanes, coping better with wind gusts, than they did when they were not talking. Assuming that their increased variability in steering (Figure 2.7) reflected coping with the wind, the talkers did not exhibit improved lane maintenance throughout the driving trials. Instead, the improvements actually accompanied speaking. There was no similar reaction by the comprehenders, despite their similar tendency to drop behind the lead vehicle. Therefore, if compensation for increased danger is indeed a strategic reaction to the secondary language tasks, we could conclude that drivers regard speaking as more dangerous than listening and as requiring additional forms of compensation, a conclusion that would support an interpretation of speaking as the harder task. Alternatively, it could be that listening elicits fewer efforts to compensate for danger than speaking does. If so, one might reason that listeners must be *more* distracted than speakers, implying that understanding is harder than talking. Although this stretches credulity and strains the logic of dual-task performance, it is nonetheless fully consistent with a compensatory hypothesis.

A final interpretation is that something about the act of speaking supports more active steering, regardless of attentional demands. Without knowing precisely what that something could be or how it works, we point to a phenomenon called "motor overflow" in the human-performance literature. Motor overflow is a spread of motor system output to muscles that would otherwise be at rest (Bodwell, Mahurin, Waddle, Price, & Cramer, 2003). Failure to inhibit motor overflow could reflect demands on attention, too, but since comprehenders presumably have no motor overflow to inhibit, its workings are relevant only to the driving performance of the talkers. Motor overflow channelled through the natural skill of wayfinding and the acquired, well-practised skill of steering may serve to enhance ~~transiently~~ the ability to stay on course.

In short, there is no selective support in our findings for the hypothesis that producing language is harder than understanding it. We do have evidence from variations in velocity and headway maintenance that production and comprehension alike change driving safety for the worse, and that the changes are temporally linked to the acts of talking and understanding.

The mechanisms of production and comprehension

The simplest conclusion from our results is that talking and understanding are both hard, because on the simplest interpretation of our dual-task measures, they seem to be about equally disruptive. Because these findings come from a dual task that was designed to equate the linguistic complexity and cognitive contents of the language used, this conclusion implies that the peripheral channels, together with the underlying mechanisms, do not differ substantially in their demands on or interference with other processes.

The peripheral channels for speaking and hearing language are clearly different in many respects, and seem to take independent tolls on attention (Martin et al., 1999; Shallice et al., 1985). What is largely unknown and a subject of long-standing debate is whether and how the central mechanisms for comprehending and producing speech divide. Classical models of aphasia posit different neural underpinnings for comprehension and production (Martin, 2003), but modern research has blurred these boundaries with evidence that the differing problems of aphasics may be better explained in terms of the general severity of the impairment or specific linguistic deficits (i.e., deficits to components of language such as the lexicon or syntax) than in terms of modality.

If comprehension and production are subserved by identical cognitive mechanisms, it would make sense for them to be nearly equivalent in their ability to interfere with other activities. However, for many of the same reasons that we lack evidence about their relative difficulty, we do not know whether they work in the same ways. Theories about the relationship between production and comprehension run the gamut from near equation of the two systems (MacKay, 1987; Kempen, 1999) to endorsement of the substantial functional and neuroanatomical separation of modality-specific information (Caramazza, 1997). In between are claims that the systems are separate but intertwined in various ways. Bever (1970; Townsend & Bever, 2001) proposed that something similar to the production system serves as a backstop for language comprehension, checking its operation with analysis by synthesis. On this kind of view, the comprehension system normally relies on sophisticated lexical guesswork without much early reliance on syntax, bringing the production system's syntax into play during comprehension to verify candidate interpretations. Garrett (2000) argued not only that the production system may serve as an analysis-by-synthesis backstop for the comprehension system, but also that the comprehension system may serve as a synthesis-by-analysis backstop for production. In some theories of language production, comprehension operates to monitor the accuracy of the process (Levelt, 1989).

While falling short of establishing how production and comprehension differ or converge, there is growing evidence for similarities in their workings. Parallel findings for number agreement in English (Bock & Cutting, 1992; Pearlmutter, Garnsey, & Bock, 1999), for structural adaptations from language comprehension to language production that parallel adaptations within production (Bock, Chang, Dell, & Onishi, in press), and for similarities in lexical-structural preferences (Garnsey, Pearlmutter, Myers, & Lotocky, 1997), all indicate that whatever the differences between the systems, they must have a lot in common. To support the prosaic successes of human communication, they have to. Our preliminary results seem to extend the similarities of normal language comprehension and production to their capacity to interfere with other activities.

Some practical implications and some practical limitations

To return briefly to the practical goal of our work, we make one addition to the growing body of research showing that the use of automotive telematic devices can be hazardous to one's health. It is now fairly well established that the major disruptions to driving from the use of cellular telephones come not from the manipulation of the devices, but from the cognitive demands of ordinary conversation (Nunes & Recarte, 2002; Strayer & Johnston, 2001). This makes moot the current regulations in some states and cities that require the use of hands-free telephones in automobiles. Until the demands of conversation are better understood, we are in a poor position to legislate the details of when and how people should use the mobile communication apparatus that is available to them, down to and including the ubiquitous radio. We need safer, better-informed designs for the increasingly sophisticated and increasingly intrusive devices of the present and near future, including navigation aids and electronic mail. Current research hints that the use of cellular telephones is already more distracting than anything else we do while driving cars and can increase the probability of accidents by a factor of four (Redelmeier & Tibshirani, 1997). To this, our results provisionally add the bad news that the problems can be exacerbated as much by drivers' careful listening as by their talking.

The provision is that there is obviously much more to be done in order to establish how general (which is to say, how limited) this conclusion may be. The narrow range of driving situations we sampled, the artificiality of even a high-fidelity driving simulator, and the restricted nature of the messages that drivers produced and understood, all argue that strong conclusions would be premature. Our findings nonetheless offer good reasons to revisit the seductive idea that listening to language is little more distracting for drivers than doing nothing else at all (Recarte & Nunes, 2003).

Summary

A widely credited assumption about language is that producing it is much harder than comprehending it. The assumption is often treated as an explanation for why infants start to understand before they start to talk, for why second-language learners can understand languages that they cannot speak, and, more generally, for why the vocabulary and structural complexity accessible to comprehenders seem to exceed those accessible to speakers. But in spite of its intuitive appeal, the assumption of significant disparities in the difficulty of comprehension and production has scant empirical support. The record is thin for understandable reasons. There are problems in disentangling the component processes of comprehension and production, problems in equating the information called upon during task performance, and problems in pinpointing where variations in performance arise when someone hears or says a normal, structured sequence of words and phrases. We have begun to address these problems in experiments that equate the speech that is produced and understood within a task environment that provides continuous performance measures. Specifically, speakers and listeners respectively produce and understand the same utterances while driving in a simulator that provides continuous assessment of acceleration, velocity, lane position, steering, following distance, braking latency, and so on. Utterances are elicited with a task that controls utterance content while ensuring utterance novelty for the speakers who spontaneously produce them as well as the listeners who later understand them. Relative to single-task control conditions (driving only and speaking or listening only), we find differences between speakers and listeners in the management of component tasks, but no generalized, differential degradation in the driving performance of speakers compared to listeners. Performance degradation is time-locked to specific events of comprehension and production. The findings have implications for claims about the nature of the challenges that automotive telematic systems present to safe driving as well as for theories of executive processes in language use.

Acknowledgements

Authors are listed alphabetically. We are grateful to Pierre Wiltzius and the Beckman Institute, to Ko Kurokawa and General Motors, to the National Institutes of Health (grant numbers T32-MH18990 and R01-MH66089), and to the National Science Foundation (grant number BCS-0214270) for their support. We also thank Hank Kaczmarski, Braden Kowitz, Camille Goudeseune, Jim Farrar, and Debbie Carrier for helping to keep the simulator laboratory running smoothly; Jill Barr, Valerie Brisky, Natalie Hall, Christopher Johnson, Etienne Korvorst, Laura Matzen, Jeff Nelson, Ankur Patel, Erik Pedersen, Seth Redenbo, Brian Roland, and Aaron Yao for their help with data collection and analysis; Simon Garrod, Alan Baddeley, and other conference participants for helpful comments on the work; and Antje Meyer, Andrea Krott, and Linda Wheeldon for generously providing the venue, the funding, and the patience that made this chapter possible.

References

Alm, H., & Nilsson, L. (1994). Changes in driver behaviour as a function of hands-free mobile phones—a simulator study. *Accident Analysis and Prevention*, *26*, 441–451.

Alm, H., & Nilsson, L. (2001). The use of car phones and changes in driver behaviour. *International Journal of Vehicle Design*, *26*, 4–11.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, *130*, 701–725.

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When attention becomes counterproductive: Divided versus skill-focused attention in performance of sensorimotor skills by novices and experts. *Journal of Experimental Psychology: Applied*, *8*, 6–16.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*, 183–200.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.

Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Ed.), *Handbook of perception and cognition. Vol. 11: Speech, language, and communication* (pp. 181–216). Orlando, FL: Academic Press.

Bock, J. K., Chang, F., Dell, G. S., & Onishi, K. H. (in press). Persistent structural priming from language comprehension to language production. *Cognition*.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Bock, J. K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In J. M. Henderson & F. Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 249–278). New York: Psychology Press.

Bock, J. K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*, 653–685.

Bodwell, J. A., Mahurin, R. K., Waddle, S., Price, R., & Cramer, S. C. (2003). Age and features of movement influence motor overflow. *Journal of the American Geriatrics Society*, *51*, 1735–1739.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177–208.

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, *37*, 201–242.

Fink, B. R. (1986). Complexity. *Science*, *231*, 319.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 203–218.

Ford, M., & Holmes, V. (1978). Planning units and syntax in sentence production. *Cognition*, 6, 35–53.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Garrett, M. (2000). Remarks on the architecture of language processing systems. In Y. Grodzinsky, L. P. Shapiro, & D. Swinney (Eds.), *Language and the brain: Representation and processing* (pp. 31–69). San Diego, CA: Academic Press.

Garrett, M., Bever, T., & Fodor, J. (1966). The active use of grammar in speech perception. *Perception and Psychophysics*, *1*, 30–32.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, *13*, 425–430.

Graham, R., & Carter, C. (2001). Voice dialling can reduce the interference between concurrent tasks of driving and phoning. *International Journal of Vehicle Design*, *26*, 30–47.

Holmes, V. M., & Forster, K. I. (1970). Detection of extraneous signals during sentence recognition. *Perception and Psychophysics*, *7*, 297–301.

Huttenlocher, J. (1974). The origins of language comprehension. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 331–368). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Irwin, D. E., & Brockmole, J. (2004). Suppressing *where* but not *what*: The effect of saccades on dorsal- and ventral-stream visual processing. *Psychological Science*, *15*, 467–473.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Dover.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, *96*, 29–44.

Kempen, G. (1999). *Human grammatical coding*. Unpublished manuscript, Leiden University, Leiden, the Netherlands.

Kubose, T., Bock, J. K., Dell, G. S., Garnsey, S. M., Kramer, A. F., & Mayhugh, J. (2006). The effects of speech production and speech comprehension on simulated driving performance. *Journal of Applied Cognitive Psychology*, *20*, 43–63.

Lamble, D., Kauranen, T., Laasko, M., & Summala, H. (1999). Cognitive load and detection thresholds in car following situations: Safety implications for using mobile (cellular) telephones while driving. *Accident Analysis and Prevention*, *31*, 617–623.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. Berlin: Springer-Verlag.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, *15*, 19–44.

Martin, R. C. (2003). Language processing: Functional organization and neuroanatomical basis. *Annual Review of Psychology*, *54*, 55–89.

Martin, R., Lesch, M., & Bartha, M. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, *41*, 3–29.

McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention*, *25*, 259–265.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*, 609–636.

Nunes, L. M., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F*, *5*, 133–144.

Pearlmutter, N. J., Garnsey, S. M., & Bock, J. K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456.

Rakauskas, M., Gugerty, L., & Ward, N. J. (2004). Effects of cell phone conversations on driving performance with naturalistic conversation. *Journal of Safety Research*, *35*, 453–464.

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied*, *9*, 119–137.

Redelmeier, D. A., & Tibshirani, R. J. (1997). Associations between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, *336*, 453–458.

Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *Quarterly Journal of Experimental Psychology*, *37A*, 507–532.

Shapiro, L. P., Zurif, E., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition*, *27*, 219–246.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*, 462–466.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cellphone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, *9*, 23–32.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

van de Weijer, J. (1999). *Language input for word discovery*. Nijmegen, the Netherlands: MPI Series in Psycholinguistics.

Violanti, J. M. (1998). Cellular phones and fatal traffic collisions. *Accident Analysis and Prevention*, *30*, 519–524.

Violanti, J. M., & Marshall, J. R. (1996). Cellular phones and traffic accidents: An epidemiological approach. *Accident Analysis and Prevention*, *28*, 265–270.