

## Sentence Comprehension Deficits in Alzheimer's Disease: A Comparison of Off-Line vs. On-Line Sentence Processing

Daniel Kempler,\*† Amit Almor,\*‡ Lorraine K. Tyler,§  
Elaine S. Andersen,\*|| and Maryellen C. MacDonald\*‡

*\*Program in Neural, Informational and Behavioral Sciences; †Department of Otolaryngology and School of Gerontology, ‡Department of Psychology, and ||Department of Linguistics, University of Southern California; and §Department of Psychology, Birkbeck College, University of London, United Kingdom*

Two studies explored whether sentence comprehension impairments in Alzheimer's disease (AD) are due to deficits in syntactic processing or memory. Study 1 used a picture-pointing sentence comprehension task to measure the final outcome of comprehension in an off-line fashion. It showed the comprehension of 30 patients with AD to be impaired, but suggested that the deficits could not be attributed solely to syntactic impairments. Study 2 investigated the effects of memory on sentence comprehension by comparing off-line (grammaticality judgment) with on-line (cross-modal naming) language processing in 11 AD and 9 control subjects. The results revealed impaired performance in the off-line task but normal performance in the on-line task using the same sentences. Performance on the off-line task correlated with independent measures of verbal working memory. These data are used to argue that sentence comprehension impairments are related to verbal working memory deficits in AD. © 1998 Academic Press

### Sentence comprehension requires the processing of both word meaning

This work was supported by a grant from the National Institutes of Health to Maryellen MacDonald (NIH RO1 AG11773). We acknowledge the following for referring Alzheimer patients for this research: Dr. Victor Henderson, Professor of Neurology and Director of the Clinical Core of the Alzheimer's Disease Research Center at the University of Southern California (NIH AG05142); Dr. Helena Chui, Co-Director, Alzheimer's Disease Diagnostic and Treatment Center, Rancho Los Amigos Medical Center (DHS 94-20356); and Dr. Lee Willis, USC Department of Neurology and Geriatric Neurobehavior Center, Rancho Los Amigos Medical Center. We gratefully acknowledge the persistence and dedication of Lori Altmann, Mariela Gil, Laila Lalami, Karen Marblestone, Sarah Schuster and Karen Stevens who collected these data, and we thank the subjects and their families for their participation in this project.

Address reprint requests to Daniel Kempler, Speech Pathology OPD 2P52, LAC + USC Medical Center, 1200 North State Street, Los Angeles, CA 90033. Fax: (323) 226-2780. E-mail: kempler@hsc.usc.edu.

and syntactic structure. For example, in interpreting the sentence "The man was followed by the woman" a listener must understand the difference between *man* and *woman* and know that a passive sentence structure means that the second argument (woman) is the agent of the action. Since speech takes place over time, and a listener must remember details from early parts of a sentence in order to integrate new words into the meaning of that sentence, memory also plays an important role in sentence processing. This is exemplified by grammatical dependencies that span many words, as when a subject and verb are separated by a clause (e.g., "The man, who wandered around the old and dangerous parts of the city night after night, was followed by the woman.") Comprehension deficits can arise because of impairments in processing semantic or syntactic information or because of problems remembering elements of a sentence during the comprehension process. The primary aim of the research reported here is to locate the source of language comprehension impairments in patients diagnosed with Alzheimer's disease (AD). In particular, we wish to determine the extent to which sentence comprehension problems can be attributed to primary deficits in syntactic ability vs. memory.

Locating the source of language comprehension impairments in AD has both clinical and theoretical importance. From a clinical point of view, it is useful to understand the cause of comprehension difficulty in order to plan appropriate intervention. For example, if the problem is due to attention and memory deficits, intervention strategies might include repetition of the input; if syntactic deficits are involved, intervention might involve linguistic simplification of the input and/or alternative communication modalities. In addition to these practical issues, the patterns of relative impairment vs. sparing of language and memory functions in impaired populations furnish important data for the ongoing debate surrounding the relationship between language and nonlinguistic cognitive functions in healthy subjects.

## SENTENCE COMPREHENSION IN AD

AD patients have difficulty understanding language. This has been demonstrated in a number of procedures, including enactment (e.g., Emery, 1983; 1988), sentence-picture matching (e.g., Rochon, Waters & Caplan, 1994) and the Token Test (e.g., Tomoeda, Bayles, Boone, Kaszniak & Slauson, 1990). Because these comprehension deficits occur in the context of multiple cognitive impairments (i.e., dementia), it is not clear what is causing the language comprehension problem. Deficits in lexical semantics, syntactic ability and memory may all play a role. Since AD patients are notoriously anomie and are impaired on a range of semantic tasks (Bayles & Kaszniak, 1987, Huff, 1988; Huff, Mack, Mahlmann, & Greenberg, 1988; Kempler, 1995), it would be logical to postulate that semantic memory impairment underlies their sentence comprehension deficit. However, there are several

reasons to reject this hypothesis. First, in contrast to naming, word comprehension appears to be quite well preserved in AD (Kempler, 1988). Second, AD patients appear to have difficulty understanding sentences, even when they demonstrate normal word comprehension (Small, Andersen & Kempler, 1997). Therefore, it does not appear that semantic memory deficits underlie the sentence comprehension impairment in AD, at least in any straightforward way.

On the other hand, deficits in syntactic processing and memory are both likely causes of sentence comprehension impairment in AD. Emery (1985) and Grober and Bang (1995) have argued that AD patients' comprehension impairment is due to a "genuine syntactic deficit" (Grober & Bang, p 95). A syntactic processing deficit is supported by the pattern of comprehension impairment these patients exhibit: i.e., they have little or no difficulty understanding non-reversible passives which can be understood on the basis of word meaning alone (e.g., "The package is carried by the boy."), but make errors on reversible passives which require syntactic processing for accurate comprehension ("The boy is kissed by the girl."). Grober and Bang (1995) further argue that the comprehension deficits in AD are not due to memory impairment because AD patients show this same comprehension pattern even when "storage demands are minimized" by allowing patients to view the stimulus sentences while selecting an answer (p 104).

Other researchers have taken an alternative position, namely that memory and other nonlinguistic deficits underlie sentence comprehension impairments in AD (e.g., Rochon, Waters & Caplan, 1994; Waters, Caplan & Hildebrandt, 1991; Waters, Caplan & Rochon, 1995). These authors have found that AD patients' errors do not necessarily increase as syntactic complexity increases, as would be predicted by a syntactic deficit hypothesis. In their research, AD patients performed no differently when comprehending simple active sentences (e.g., "The lion killed the elephant.") vs. syntactically more complex structures of equal length, such as truncated passives (e.g., "The pig was touched.") Rather, they found that semantic or conceptual complexity affected comprehension: AD patients had more difficulty interpreting sentences that contained more than one proposition. Waters et al. attribute this pattern of comprehension deficits to an impairment of "postinterpretive" processing. They argue that AD patients are able to interpret the sentences normally, but have difficulty when they need to use those interpretations for some additional purpose, such as matching the sentence meaning to a picture. Postinterpretive processes are considered nonlinguistic, and include, for example, visual processing of stimulus pictures in a sentence-picture matching task and the working memory (WM) necessary to maintain an active representation of the sentence meaning while selecting a match from a picture array.

The Waters et al. account makes several predictions. First, in off-line picture pointing tasks, AD patients' performance will be affected by semantic

and propositional complexity. Second, AD patients should perform normally in those tasks that minimize the contribution of postinterpretive processes. Such tasks are typically referred to as “on-line” in that they are claimed to tap into the immediate, automatic, first-pass analysis of the language input (e.g., Marslen-Wilson & Tyler, 1975; Tyler, 1992b), and minimize the use of memory and metalinguistic processes. Tyler (1992 a & b) has reported data from a number of aphasic patients whose performance on on-line tasks is substantially better than their performance on off-line tasks, and she has argued that exclusive reliance on off-line tasks can lead to an underestimation of some aphasic patients’ language comprehension abilities.

We investigate a similar possibility here: AD patients’ poor performance on sentence comprehension tasks may be due not to deficits in syntactic knowledge but rather to other factors that interfere with task performance, such as working memory (WM) impairments, which are endemic in this population (e.g., Morris & Baddeley, 1988). We test this hypothesis by using two types of language comprehension task to evaluate the comprehension abilities of AD patients: off-line tasks, including sentence-picture matching and grammaticality judgment, which require the subject to hold the sentence in WM and then make a metalinguistic judgment about it; vs. an on-line task, i.e., cross-modal naming, in which responses are collected in the course of on-line language comprehension and no judgments are required. If a syntactic deficit underlies AD patients’ problems with language comprehension, then these patients should perform more poorly than control subjects on both on- and off-line tasks. On the other hand, if these patients have intact syntactic processing abilities, but deficits in working memory, then performance should be impaired only on the off-line tasks.

We present data from two studies. Study 1 first establishes the extent of AD patients’ comprehension deficits on a variety of sentence constructions in an off-line task. Study 2 then compares performance on an off-line task vs. a parallel on-line task to explore the effects of post-interpretive demands on comprehension. In addition, as part of Study 2, we collected data from several working memory measures, including some new measures that were designed to tap the working memory processes involved in language comprehension. These measures allow us to investigate the extent to which deficits in sentence processing can be attributed to working memory impairments.

### STUDY 1: SENTENCE-PICTURE MATCHING

There were two aims of this study. First, we wanted to replicate the typical finding of sentence comprehension deficits in AD through a standard sentence-picture matching task. In order to compare off-line performance with on-line performance (Study 2), we needed to be sure that our AD subjects are comparable to others in the literature in that they make errors on standard picture pointing tests. Second, we wanted to investigate the role of sentence structure in the comprehension performance of AD subjects.

TABLE 1  
Subject Information, Study 1

	MMSE	Age	Education
	Mean ( <i>SD</i> , range)	Mean ( <i>SD</i> , range)	Mean ( <i>SD</i> , range)
AD ( <i>n</i> = 30)	17.9 (5.0, 7–26)	76 (7.7, 61–89)	13 (3.1, 8–20)
NC ( <i>n</i> = 23)	28.4 (1.5, 25–30)	75 (5.7, 63–88)	16.6 (3.0, 8–22)

*Note.* Demographic information for patients with Alzheimer's disease (AD) and normal controls (NC), including severity of dementia as gauged by performance on the Mini-Mental State Examination (MMSE, total possible = 30), age at the time of testing, and years of education.

### Method

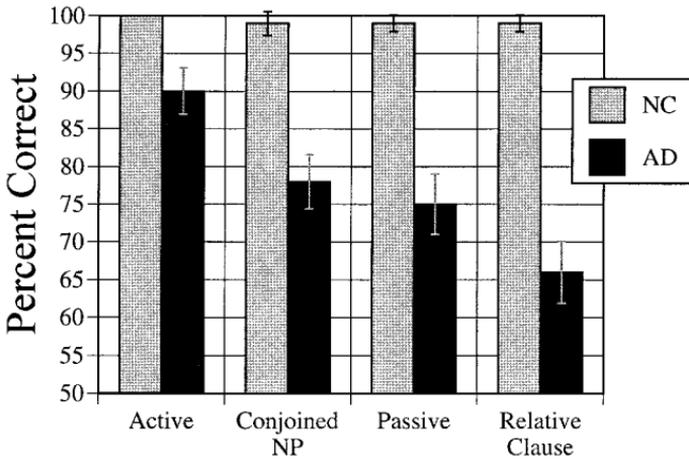
*Subjects.* Thirty individuals diagnosed with Probable Alzheimer's disease and 23 age-matched normal control (NC) subjects participated in the first study (see Table 1). The Alzheimer subjects were referred by the University of Southern California Alzheimer's Disease Research Center and the Alzheimer's Disease Diagnostic and Treatment Center at Rancho Los Amigos Medical Center. All AD subjects met the NINCDS-ADRDA criteria for probable Alzheimer's disease (McKhann et al., 1984). Results of neurological, laboratory (including computed tomography or magnetic resonance scan), and neuropsychological assessment failed to suggest other causes of dementia. The AD subjects were mildly to moderately demented as gauged by the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). All subjects were native speakers of Standard American English. Twenty-seven men and 26 women participated. There were no significant differences between the AD and control subjects in age ( $t(51) = .49, p = .63$ ), but the difference in education was significant ( $t(51) = 3.9, p < .001$ ). The difference in education between the AD and control groups is largely unimportant here because the task is so simple that control subjects perform near ceiling and the relevant comparisons (reported below) are performance differences between the sentence structures within the AD group.

*Materials and procedures.* Sentence comprehension was measured using a sentence-picture matching protocol. The stimuli were 24 sentences, including exemplars of four grammatical constructions: simple active, active with a conjoined noun phrase (NP), full passive, and relative clause (subject-subject and object-subject relatives). All sentences used common nouns and transitive verbs, but differed in number of words (range 5–9), the number of participants (2 vs. 3) and whether the sentence was syntactically complex (i.e., contained noncanonical word order or embedding) (see Table 2). The number of participants is important as an indication of the semantic density of the sentence, and also reflects the complexity of the picture

TABLE 2  
Sentence Types and Examples Used in the Off-Line Sentence-Picture Matching Protocol

Sentence type	Example	Syntactically complex?	No. of participants
Active	The boy pushes the girl.	No	2
Conjoined NP	The boy scratches the dog and the cat.	No	3
Passive	The boy is kissed by the girl.	Yes	2
Relative clause	The dog chases the girl, that chases the boy	Yes	3

## Comprehension of Four Sentence Types



**FIG. 1.** Performance (mean percent correct and standard error) for Alzheimer (AD) and Normal Control (NC) subjects on a sentence-picture matching task, using four different sentence types.

array: the pictures that accompanied the 3-participant sentences contained more figures and were therefore visually more complex than the pictures that accompanied the 2-participant sentences. All of the sentences were reversible in that the participants could have reversed roles without affecting plausibility. The distracter pictures represented syntactic foils: they depicted the same action as the target (pushing, carrying, kissing, etc.) and the same participants, but in different roles. For example, the item “The boy is kissed by the girl” was accompanied by the target picture (a girl kissing a boy) and the syntactic foil (a boy kissing a girl). Each subject heard a spoken sentence and pointed to one of the two vertically arrayed line drawings that best depicted the meaning of the sentence.

These sentence structures allowed us to test several hypotheses about the source of the comprehension deficits. For example, if syntactic complexity creates comprehension problems, patients should perform more poorly on passive than active sentences; if number of participants is a crucial factor, performance should be compromised on the active sentences with a conjoined NP compared to the simple active sentences.

### Results

Performance of the AD and control subjects is presented in Fig. 1. The AD patients performed overall at 77% correct (range 46–100%). This is significantly worse than the control subjects, who performed near ceiling (mean = 99% correct; range 88–100%) ( $\chi^2(1) = 130, p < .001$ ). Separate chi square comparisons for each sentence type showed that AD subjects performed worse than NCs on all sentence types (ACTIVE:  $\chi^2(1) = 12.53, p = .001$ ; PASSIVE  $\chi^2(1) = 12.1, p < .001$ ; CONJOINED NP  $\chi^2(1) = 49.3, p < .001$ ; RELATIVE CLAUSE  $\chi^2(1) = 92.4, p < .001$ ).

Looking at the AD subjects only, chi-square comparisons of each sentence

type with each other sentence type showed: better performance on the simple actives than all other sentence types (ACTIVE vs. PASSIVE  $\chi^2(1) = 9.5$ ,  $p = .002$ ; ACTIVE vs. CONJOINED NP  $\chi^2(1) = 7.5$ ;  $p = .006$ ; ACTIVE vs. RELATIVE CLAUSE  $\chi^2(1) = 24.2$ ,  $p < .001$ ); no difference between the PASSIVE vs. CONJOINED NP ( $\chi^2(1) = .51$ ,  $p = .477$ ); an almost significant difference between the PASSIVE vs. RELATIVE CLAUSE ( $\chi^2(1) = 3.3$ ,  $p = .07$ ); and a significant difference between CONJOINED NP vs. RELATIVE CLAUSE ( $\chi^2(1) = 9.5$ ,  $p = .002$ ). There was a significant correlation between overall dementia severity as gauged by the MMSE score and sentence comprehension ( $r = .6$ ,  $p < .001$ ).

### *Discussion*

A group of 30 mild-moderately demented AD subjects were impaired on a sentence-picture matching test of sentence comprehension. This replicates previous research demonstrating impaired comprehension with comparable tasks (e.g., Grober, & Bang, 1995; Rochon et al., 1994; Small, Kemper & Lyons, 1997). The AD subjects in this study were significantly impaired on all sentence types, including the simplest structure—simple active. They performed best on active sentences, worst on sentences containing relative clauses, with intermediate performance on passives and actives containing a conjoined NP.

Can we tell from these data if the sentence comprehension impairment is due to syntactic vs. postinterpretive processing deficits? The data offer some support for each hypothesis. In support of a syntactic deficit, the subjects performed worse on the two syntactically complex sentence structures than the syntactically simple structures, matched for number of participants: worse on passives than simple actives (both contained 2 participants), and worse on relative clauses than actives with conjoined NPs (both contained 3 participants). Although this comparison is minimally confounded by sentence length (the syntactically complex sentences are slightly longer than the syntactically simple sentences), the difference in sentence length is too small to seriously support an explanation based exclusively on sentence length: passives are 2 words longer than the actives; relative clauses are 1 word longer than the conjoined NPs.

However, two comparisons suggest that the comprehension deficit is not caused solely by a syntactic processing impairment. First, looking at performance on just the simpler sentences, the patients performed worse on the actives with a conjoined NP than on the simple actives. Since both of these structures are syntactically simple (canonical word order, etc.), this contrast is likely due to the difference in sentence length (8 vs. 5 words) or the number of participants (3 vs. 2), or both. Worse performance on the longer sentences would be predicted by the postinterpretive deficit hypothesis: Because words (participants) must be kept active as part of the sentence representation while

subjects scan the response array, a WM impairment that limits the number of elements actively maintained in the semantic representation would affect those sentences with more words and more participants, regardless of syntactic complexity.

A second comparison also suggests that a syntactic deficit does not account for the results. If AD subjects suffer from a just syntactic deficit, we would predict, given sentences of comparable length, their performance would be worse on the syntactically more complex (i.e., passives) than on the simpler (i.e., conjoined NP) sentences. This prediction was not borne out: The AD subjects performed similarly on the passive and conjoined NP sentences. Although a syntactic processing impairment might explain the errors on the passive sentences, it does not explain the equal number of errors on the conjoined NP sentences, which are syntactically simpler (since they contain canonical subject-verb-object word order). These data are compatible with the notion of a postinterpretive deficit, which affects performance on the task due to demands on, e.g., visual processing and working memory, which are both required by the task and impaired in AD. In this later explanation, it is the processes involved in making the metalinguistic judgment required by this task (comparing and selecting a picture which matches the sentence meaning), not linguistic processes per se, that are impaired. The important role of nonlinguistic (e.g., memory) impairments in sentence comprehension would be more clearly demonstrated if we could show that comprehension improves when the postinterpretive load is lessened. The next study compares sentence comprehension in two tasks that vary in the amount of postinterpretive processing they impose.

## STUDY 2: ON-LINE VS. OFF-LINE PERFORMANCE

The first study confirmed that AD patients have difficulty interpreting sentences when comprehension is measured by an explicit off-line task. Our second study asks whether AD patients will exhibit a sentence processing deficit when comprehension is measured by an on-line task that reflects the automatic interpretation of the language input and minimizes metalinguistic and memory demands. To this end, we compared the comprehension ability of AD patients in an on-line task (cross-modal naming) with their comprehension in an off-line task (grammaticality judgment), using the same materials in both tasks. The cross-modal naming paradigm capitalizes on AD patients' relatively preserved ability to read aloud throughout the moderate stages of the disease (Bayles, Tomoeda & Trosset, 1992), and was selected after pilot testing revealed that the AD patients were able to perform this type of task with relatively little difficulty. The grammaticality judgment task was selected because it allowed the use of the same stimuli as used in the cross-modal naming paradigm (pairs of grammatical and ungrammatical

TABLE 3  
Subject Information, Study 2

	MMSE	Age	Education
	Mean ( <i>SD</i> , range)	Mean ( <i>SD</i> , range)	Mean ( <i>SD</i> range)
AD ( <i>n</i> = 11)	21 (1.6, 16–24)	81 (6.6, 64–89)	15 (2.3, 12–20)
NC ( <i>n</i> = 9)	29 (1.1, 27–30)	77 (4.4, 70–85)	17 (2.8, 13–22)

*Note.* Demographic information for patients with Alzheimer's disease (AD) and normal controls (NC), including severity of dementia as gauged by performance on the Mini-Mental State Examination (MMSE, total possible = 30), age at the time of testing, and years of education.

sentences) in an off-line task. Since a key hypothesis under investigation is that working memory impairment is largely responsible for off-line comprehension deficits, we also collected independent measures of verbal working memory for direct comparison with the language comprehension measures.

### Method

*Subjects.* Eleven subjects diagnosed with Probable Alzheimer's disease and nine healthy age-matched controls participated in this study. This sample is a subset of the subjects who participated in Study 1 (see above for diagnostic information, and Table 3 for demographic information). The healthy control and AD groups in Study 2 were not significantly different in either age ( $t(18) = 1.38, p = .19$ ) or years of education ( $t(18) = -1.6, p = .14$ ). There were 6 women in each group.

*Materials and procedure.* The stimulus sentences used in the cross-modal naming test were 80 sentences consisting of 40 grammatical/ungrammatical sentence pairs. Grammatical sentences were constructed and then altered to create ungrammatical counterparts by the substitution, addition or deletion of one word. The sentences were constructed so that they were either grammatical or ungrammatical at a particular target word (italicized in the following examples). Two types of grammatical violations were included: (1) subject verb disagreement ("After the safety course, the students *were/\*was* preventing fires") and (2) errors of transitivity, including both (i) transitive verbs not followed by a direct object ("Bill poured *cereal.*" vs. "\*Bill poured with *cereal.*"), and (ii) intransitive verbs followed directly by a noun ("Susan went to *meetings*" vs. "\*Susan went *meetings*").

The stimuli were digitally recorded using MacRecorder and a Macintosh computer in a sound attenuated booth with a constant mouth-to-microphone distance of 8 inches. Each stimulus sentence was recorded with a semantically neutral preceding context sentence which, in the experiment, served to alert and orient the subject to the stimuli. The target words (and any subsequent material) were digitally excised. The resulting sentence fragments were used in the experiment, which was controlled by the PsyScope software package (Cohen, MacWhinney, Flatt, & Provost, 1993).

Each subject was tested in a quiet room, sitting in front of a Macintosh Classic computer. Auditory stimuli were presented via a high quality loudspeaker, with volume adjusted to ensure that each subject could hear the stimuli adequately. For each stimulus, subjects heard a complete context sentence followed by a sentence fragment. At the offset of the sentence fragment, a single word, which was a continuation of the sentence, appeared on a computer monitor.

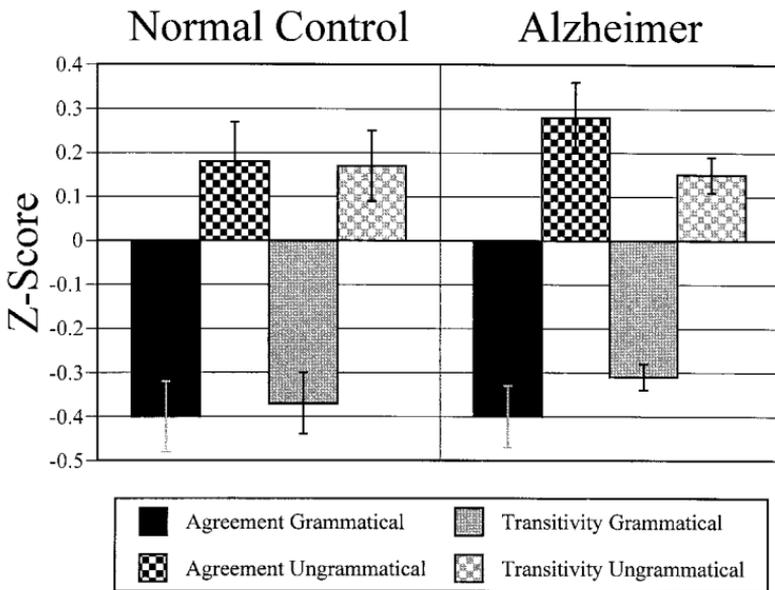
Subjects were asked to read the word aloud as quickly and accurately as possible (Tyler & Marslen-Wilson, 1977; Marslen-Wilson, Tyler & Koster, 1993). Each auditory sentence fragment in conjunction with the visually presented target created either a grammatical or ungrammatical construction. In some cases the visual target made the sentence fragment into a complete sentence; in other cases, the fragment did not complete (but merely continued) the sentence. In either case, the auditory presentation stopped at the onset of the visual target. In order to assure that the subjects were attempting to integrate the context fragment and the visual target (and not just reading the words aloud, see Tyler & Marslen-Wilson, 1977), we included an additional task. After naming each target word, control subjects were asked whether the word was a good or bad continuation of the sentence; Pilot testing with AD subjects indicated that this secondary task was so distracting, that they could not produce accurate and fast responses to both the primary (word reading) and secondary tasks. Therefore, an alternative directed attention task was developed for the AD subjects: they were asked yes/no comprehension questions for 25% of the context sentences. This secondary task successfully directed their attention to the stimuli and did not interfere with their ability to perform the primary task. The responses to the secondary task were not scored, but merely served to encourage the subjects to pay close attention to the stimuli.

We tested the subjects in four sessions, a minimum of 1 week apart for the AD patients and three weeks apart for the control subjects. The stimuli presented in each session were counterbalanced so that a subject never heard the same sentence fragment more than once in a single session, and therefore never encountered a grammatical and ungrammatical version of the same item in the same session. Stimuli were pseudo-randomly mixed with fillers of different sentence structures. Forty percent of the items were from this experiment and 60% were fillers of other sentence structures. Half of the items were grammatical and half were ungrammatical. Ten practice items began each session.

Latencies for reading the target words in the two experimental conditions (grammatical vs. ungrammatical) were used to indicate whether subjects were sensitive to the grammatical constraints of the preceding fragment. Slow naming of the ungrammatical compared to the grammatical continuations is expected in normal subjects, and indicates the processing of grammatical information. If AD subjects suffer from grammatical processing deficits, we would expect reduced differences between the grammatical and ungrammatical conditions. However, if the source of the comprehension impairment in AD is due to memory limitations, we would expect the AD subjects to perform similarly to the controls in this task, for which memory demands are minimal.

The grammaticality judgment task used 40 of the 80 stimuli from the cross-modal experiment. Half of the items were from the subject-verb agreement list and half were drawn from transitive/intransitive sentences. Half of the items in each set were grammatical and half were not. Subjects heard only one version (grammatical or ungrammatical) of any particular sentence. In the grammaticality judgment task, the subject was presented with a complete spoken sentence, and asked if the sentence was 'good English' or not. These items were administered in a different session than the cross-modal naming items.

Working memory (WM) was assessed by a combination of standard memory tasks and several new tasks designed to assess storage and manipulation of verbal information. The standard memory tasks were taken from the Wechsler Memory Scale (Wechsler, 1987), and included counting from one to 40 by threes and backward digit span (repeating an increasingly long series of numbers in reverse order from their presentation order). New WM tasks were designed to involve the mental procedures thought to underlie standard verbal working memory span tasks such as those used in Daneman and Carpenter's (1980) listening span, but to be simple enough for demented patients to perform. The specially devised WM measures were comprised of (i) month ordering: putting into calendar sequence an increasingly long set of months presented out of calendar order and (ii) digit ordering: putting into ascending order an increasingly long series of randomly ordered digits. These tasks create similar working



**FIG. 2.** Reading latencies measured in milliseconds and normalized by z transformation based on the individual subject's means and standard deviations. Graph shows mean latencies for reading grammatical vs. ungrammatical sentence continuations in subject-verb agreement and transitivity conditions, by Normal Control subjects and patients with Alzheimer's disease. Relatively faster reading latencies are negative (less than the mean) while slower reading times are positive (greater than the mean). Variability is shown by standard error bars.

memory and processing demands to those contained in auditory language comprehension, in that the subject must activate semantic information from the auditory stimuli in order to re-order the input items. There were four items presented at each span level, and testing was stopped after two items at a span level were sequenced incorrectly.

## Results

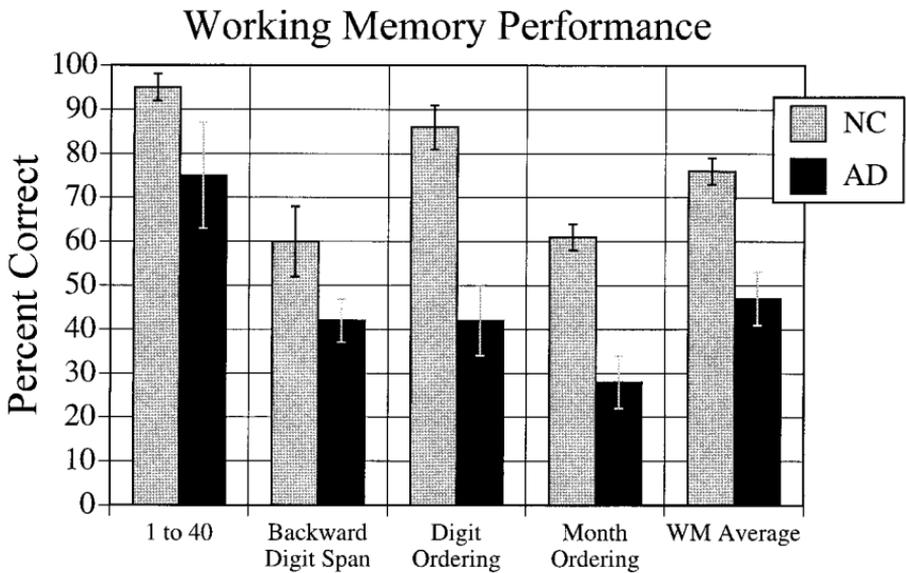
*On-line:* Cross-modal naming. Reaction times from the cross-modal naming experiment were analyzed after eliminating all machine errors (5% of the data), reading errors (3% of responses) and latencies well beyond the observed range: normal control responses above 2300 msec, and all AD responses longer than 4 sec (1% of the data). Next, the data from each sentence structure (subject-verb agreement and transitives/intransitives) were normalized by transforming them into z-scores based on each subject's mean and standard deviation for each sentence structure for each session. This was done to minimize irrelevant effects due to variations in the speed of individual subjects across sessions. The data were then pooled across sessions for each subject and each sentence type. A grammaticality effect for the two populations (for each sentence type) was derived by calculating the difference between the grammatical and ungrammatical response times. These results are shown in Fig. 2.

Because preliminary analysis of each sentence structure failed to show any effect due to construction type ( $F < 1$ ), the data from both constructions are combined in the following analyses. An ANOVA examining the grammaticality effect (within subjects factor) in the two groups (between subjects factor) showed a reliable effect of grammaticality ( $F(1, 18) = 80.19, p < .001$ ), no effect of group ( $F < 1$ ), and no significant interaction ( $F < 1$ ). The results therefore demonstrate that AD patients and healthy controls were equally sensitive to grammatical violations in this on-line experiment.

*Off-line: Grammaticality judgments.* Correct responses to the grammaticality judgment task were converted to percent correct scores. Healthy controls performed near ceiling (mean 95% correct overall, range: 88–100%), while the AD patients performed significantly worse (mean 80% correct, range: 58–95%) ( $t = -3.92, p < .001$ ). These data are similar in pattern and magnitude to the off-line comprehension data obtained in our first study with a picture-pointing task. Since the subjects that participated in Study 2 are a subset of those that participated in Study 1, we have both sentence-picture matching and grammaticality judgment performance from these 20 subjects. Comparing performance on these two off-line tasks revealed a significant correlation (AD & NC subjects combined;  $r = .86, p < .001$ ; AD patients only:  $r = .86, p < .001$ ). This finding suggests that despite the difference in sentence structures and task constraints in the two off-line comprehension tests, these two measures elicited a comparable degree of impairment in this group of subjects.

*Working memory.* Working memory measures were converted to percent correct scores and are displayed in Fig. 3. Control subjects performed better than the AD subjects on all measures, but the difference between the two groups was greatest for the two new WM measures: (1 to 40:  $t(18) = -1.5, p = .08$ ; backward digit span,  $t(18) = -1.9, p = .03$ ; digit ordering,  $t(18) = -4.3, p < .001$ ; month ordering,  $t(18) = -4.6, p < .001$ ). A composite WM score was derived by averaging performance on all four measures, and this score reliably distinguished between the AD and NC groups ( $t = -3.8, p = .0007$ ).

To compare performance on the comprehension and memory measures, pairwise correlations were calculated using each subject's grammaticality effect (the difference between grammatical and ungrammatical response times in the on-line task), percent correct on grammaticality judgment, and composite WM score. The only significant correlation was between the grammaticality judgment and the WM score ( $r = .69, p < .01$ ). On-line sensitivity to syntactic anomalies (as reflected by the grammaticality effect in cross-modal naming) appears to be independent of performance on the judgment task ( $r = .19; p = .42$ ), and independent of WM impairment ( $r = .07, p = .76$ ). Dementia severity (MMSE) was significantly correlated only with performance on grammaticality judgments ( $r = .83, p < .01$ ).



**FIG. 3.** Performance (mean percent correct and standard error) for Alzheimer (AD) and Normal Control (NC) subjects on four working memory tasks, and mean performance on all four memory tasks combined.

### *Discussion*

While Study 1 used off-line comprehension performance to investigate the role of syntactic and memory impairments in sentence comprehension in AD, the second study took a different approach. We reasoned that if the comprehension impairment was caused by mental processes that occur after the immediate processing of the sentence, then if we were able to probe the comprehension processes as they occur in real time, minimizing the WM requirements, we might see evidence for intact comprehension. Further, since we hypothesized that the comprehension impairment might be due to reduced WM resources, we independently measured WM. Comparing on-line vs. off-line comprehension showed that AD subjects had difficulty responding accurately to the off-line (grammaticality judgment) task despite showing normal grammaticality effects in the on-line cross-modal naming study with the same sentences.

The explanation for this dissociation may lie in the role of working memory in language comprehension. The significant correlation between off-line grammaticality judgment and measures of working memory suggests that performance on these two tasks rely on the same pool of cognitive resources. In contrast, processing grammatical information on-line, at least in these measures and in this sample, appears not to be affected by memory deficits.

## GENERAL DISCUSSION

The data presented here have implications for (1) methodological issues of language comprehension assessment, (2) theoretical discussions of the nature and cause(s) of language deficits in AD, and (3) practical issues of intervention with AD patients.

First, the data show that the type of task used to assess comprehension greatly influences the results. There is now a significant literature comparing off-line and on-line language performance in impaired populations, particularly chronic aphasic patients (e.g., Tyler, 1992a). In general, on-line, implicit tasks tend to reveal more preserved knowledge than suggested by the results of off-line tasks (Tyler, 1992a). This phenomenon appears to be quite general, showing up across several patient populations and across different aspects of language (semantics, syntax, morphology) (Nebes, Boller, & Holland, 1986; Small, Andersen, & Kempler, 1997; Tyler, 1988; 1992; Tyler, Ostrin, Cooke, & Moss, 1995; Tyler et al., 1997). This study adds another line to this unfolding story: in contrast to well-documented off-line grammatical comprehension impairments in AD (e.g., Grober & Bang, 1995), grammatical processing deficits are not evident in this sample of AD patients when on-line methods are used to assess processing of basic sentence structure (see also Small, Andersen & Kempler, 1997).

Second, we can ask: what do our off-line and on-line data show us about the nature and causes of sentence comprehension impairments in AD? One preliminary interpretation of impaired off-line vs. intact on-line performance is that syntactic processing in AD is essentially intact and the observed off-line deficits are due to other, non-syntactic impairments. We have presented evidence that deficits in working memory correlate with the off-line sentence comprehension impairments, suggesting that memory deficits might cause the off-line comprehension problems.

A causal relationship between working memory impairments and language deficits is consistent with other research in both normal and impaired populations. For example, Just and Carpenter (1992) have argued that individual differences in linguistic working memory capacity underlie differences in linguistic performance within the normal population. This approach has been extended to the aphasic population; Miyake, Just, and Carpenter (1995) suggest that aphasic patients suffer from severe working memory limitations, which are the cause of the patients' comprehension deficits. However, the precise relationship between language and working memory is far from clear: Caplan, Waters, and Martin (Martin, 1995; Caplan & Waters, 1995; Waters & Caplan, 1996; Waters & Caplan, 1997) have argued that there is no unitary linguistic working memory that underlies all linguistic processes, but rather separate working memories that serve individual processing modules. Waters and Caplan (1996) have further argued that the linguistic working memory measures used by Just and Carpenter tap only

postinterpretive processes and do not correlate with on-line language comprehension (but see Just, Carpenter & Keller, 1996).

Our research bears on this debate in several ways. The working memory tasks used here, and in particular the digit- and month-ordering tasks which were designed to tap the linguistic working memory processes that are crucial for language comprehension, clearly distinguished the AD and control groups. These data confirm the substantial verbal working memory impairments in the AD population. Our working memory measures correlated with off-line language comprehension performance, but not with on-line performance. This is consistent with Waters et al.'s (1994; Waters and Caplan, 1996) position that correlations between working memory and language performance reflect the importance of working memory for postinterpretive processes, not linguistic processes per se. The lack of correlation between the linguistic working memory performance and on-line tasks of language comprehension underscores the possibility that the cognitive processes involved in verbal working memory tasks do not play a crucial role in sentence processing as it is taking place.

However, given other data in the literature and some limitations of these data, several caveats and possibly a more cautious interpretation are necessary. First, normal grammatical processing in our on-line paradigm does not necessarily mean that AD patients can adequately process sentences as they hear them, or that syntax is fully normal in AD. We have only shown that AD patients can respond normally when trying to integrate adjacent words in short sentence fragments. These findings would have to be expanded to include various and more challenging sentence structures before we conclude that on-line sentence processing is intact in AD.

Second, despite earlier findings of intact syntactic ability in AD (Kempler, Curtiss & Jackson, 1987; Schwartz, Marin & Saffran, 1979), there is now convincing evidence from other research that syntactic impairments, although subtle, do appear in AD. At least 2 studies have found that while AD subjects do not generally produce grammatical errors, their language production is not grammatically normal. Kemper et al. (1993) analyzed written sentences produced by 92 patients with AD and those produced by 276 control subjects. They found that even though the language produced by the demented group was grammatical, it was also syntactically simpler in that it contained fewer embeddings. Bates, Marchman, Harris, Wulfeck & Kritchevsky (1995) reported a similar finding: in describing a film that frequently elicits passive structures in normal subjects, AD subjects used fewer passives overall, and used more get-passives (as opposed to be-passives) than younger or age-matched controls. Again, although the AD patients did not make obvious grammatical errors they did use a restricted set of grammatical forms and used them in different circumstances than normal controls. Bates et al. argued that this difficulty accessing certain grammatical forms is comparable to other language problems in AD. Just as AD patients have a difficult time

accurately naming objects, and often come up with a description instead of a label, AD patients have a difficult time accessing the preferred grammatical forms. Their explanation for these findings in both lexical semantics and grammar is that “AD patients suffer from a progressive deterioration of controlled processes, including difficulty in suppressing an expected response (i.e., controlled inhibition), and difficulty in generating and/or deciding among a new set of alternatives that are not obvious in context (i.e. controlled excitation)” (Bates et al., 1995, p. 531). In addition to these subtle syntactic changes in AD, a few studies have found that AD patients do make grammatical errors. Altmann, Andersen & Kempler (1993), and Altmann (1998) have analyzed both spontaneous and elicited speech of AD patients and found a significant number of morpho-syntactic errors, indicating that even in a relatively unconstrained on-line task, requiring little or no working memory, AD patients do evidence syntactic deficits. One finding from Altmann’s work supports the Bates et al. contention that syntactic and semantic errors are not independent from one another: the number of syntactic errors correlated significantly with measures of semantic impairment. It is possible then, that difficulty in grammatical and semantic processing may both reflect a more general disturbance in language processing, and language deficits in AD may not be limited to one or the other domain.

These data suggest that, while AD patients may produce relatively intact syntactic forms in unrestricted contexts, when challenged, their grammar shows signs of impairment. This type of explanation may also account for the pattern of results described here: AD patients performed in the impaired range on all tasks where they were required to activate and maintain a semantic representation (sentence-picture picture matching, grammaticality judgment, and WM ordering tasks). In contrast, AD patients performed normally on those tasks where we measured automatic access to very high frequency, and therefore easily accessible structures (e.g., subject-verb agreement) without requiring them to reflect on sentence form or meaning. Our use of frequent grammatical structures and those that required only very local processing was intentional, and provided the greatest contrast with off-line comprehension tasks. However, these stimuli may have masked subtle syntactic processing impairments. Because these stimuli required grammatical integration only over several words, they did not require full integration of grammatical and semantic information into a sentential representation. Furthermore, only access to very frequent structures were tested. Processing of frequent and local grammatical constructions would be likely preserved even in patients with subtle grammatical processing impairments. Accordingly, we suspect that we may see syntactic impairments, even on-line, if we probe (1) sentences that contain greater distances between dependent constituents as with subject-verb agreement where the subject and the verb are separated by an independent modifying clause, and (2) less common grammatical structures such as passive voice sentences.

One final caveat is in order regarding the relationship between WM and off-line sentence processing. The underlying impairment responsible for the correlation between off-line and WM tasks could be characterized in several ways. We could propose, as others have, that the impaired off-line and WM tasks all require working memory, and performance on them suffers due to the memory impairment in AD. However, a correlation does not make a cause. It is equally possible that our measures of verbal memory and off-line language processing are correlated because they all require activation of semantic representations, and AD patients are impaired in accessing and deriving semantic representations. It should also be considered that a distinction between the constructs of verbal working memory and semantic activation may be more a theoretical convenience than anything else. On this view, deriving a semantic representation from a sentence and re-ordering a list of months or digits both require accessing and maintaining a semantic representation, and that is why they correlate. The reason that neither the off-line nor WM tasks appear associated with on-line grammatical processing is that a full semantic representation need not be activated or maintained in our on-line measure of grammatical processing.

Finally, these data have practical implications for conversing with Alzheimer patients. Our results suggest that the memory problems characteristic of AD cause, or at least aggravate, sentence comprehension deficits. Therefore, sentences that put extra demands on memory are likely to exacerbate the comprehension problems for this group of patients. Conversely, strategies that minimize memory demands are likely to enhance comprehension. As predicted, strategies such as repetition and paraphrase, which are thought to compensate for memory impairments, have been shown to improve comprehension in AD (Small, Kemper & Lyons, 1977). However, other strategies that are commonly recommended may actually increase demands on memory. For instance, several resources for caregivers recommend slowing speech rate in order to increase comprehension (e.g., Gwyther, 1985; Mace & Rabins, 1991). While this may ameliorate comprehension deficits for some patients, the beneficial effect will probably vary depending on the cause and severity of the comprehension impairment. On the positive side, slow speech is often more clearly articulated than fast speech and therefore may be more audible and intelligible. This is likely to benefit patients with presbycusis (hearing loss due to aging). However, stretching out speech over time will potentially interfere with comprehension for patients with severe working memory deficits, who will have to maintain verbal information longer in working memory. This hypothesis has been supported by recent research by Small and colleagues (Small, Andersen & Kempler, 1997) who demonstrated the different effects that slow speech can have in a small sample of AD patients. In that research, slow speech rate was beneficial for an AD patient with mild WM deficits, but detrimental for another AD patient with more severe WM impairments.

In summary, we have shown that sentence comprehension performance in AD depends greatly on how comprehension is measured. Measurement techniques that require postinterpretive processing such as picture-pointing and grammaticality judgment are likely to show deficits. We have attributed these impairments to deficits in working memory rather than deficits in syntactic ability. Importantly, insofar as we have measured sentence comprehension without postinterpretive or working memory demands in an on-line paradigm, AD patients appear to process grammar normally. Although on-line syntactic processing may be intact for the simple and frequent sentence structures investigated here, language comprehension in daily interaction will undoubtedly appear impaired because of other cognitive (i.e., memory) deficits. Any language based intervention for communication deficits in AD should therefore focus on techniques that compensate for memory deficits rather than focus on linguistic structure.

## REFERENCES

- Altmann, L. P. 1998. *Effects of working memory and semantic impairments on the speech of Alzheimer's disease patients*. Doctoral dissertation, Department of Linguistics, University of Southern California.
- Altmann, L. P., Andersen, E. S., & Kempler, D. 1993. *Reevaluating syntactic preservation in Alzheimer's disease*. Presented at the Academy of Aphasia, October, Tucson.
- Bates, E., Marchman, V., Harris, C., Wulfeck, B., & Kritchevsky, M. 1995. Production of complex syntax in normal aging and Alzheimer's disease. *Language and Cognitive Processes*, **10**,(5), 487–539.
- Bayles, K. A., & Kaszniak, A. W. 1987. *Communication and cognition in normal aging and dementia*. Boston: Little, Brown and Company.
- Bayles, K. A., Tomoeda, C. K. & Trosset, M. W. 1992. Relation of linguistic communication abilities of Alzheimer's patients to stage of disease. *Brain and Language*, **42**, 454–472.
- Caplan, D. & Waters, G. 1995. Aphasic disorders of syntactic comprehension and working memory capacity. *Cognitive Neuropsychology*, **12**(6), 637–649.
- Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. 1993. *PsyScope*. Software developed at Carnegie Mellon University, Department of Psychology.
- Daneman, M. & Carpenter, P. A. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, **19**, 450–466.
- Emery, O. B. 1983. Language in senile dementia of the Alzheimer type. *Psychiatric Journal of the University of Ottawa*, **8**(4), 169–178.
- Emery, O. B. 1985. Language and Aging. *Experimental Aging Research*, **11**, 3–60.
- Emery, O. B. 1988. Language and memory processing in senile dementia of the Alzheimer's type. In L. L. Light & D. M. Burke (Eds.), *Language, memory and aging*. New York: Cambridge University Press. Pp. 221–243.
- Folstein, M. F., Folstein, S. E., McHugh, P. R. 1975. Mini mental state: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–98.
- Grober, E. & Bang, S. 1995. Sentence comprehension in Alzheimer's disease. *Developmental Neuropsychology*, **11**, 95–107.

- Gwyther, L. P. 1985. *Care of Alzheimer's patients: A manual for nursing home staff*. American Health Care Association & the Alzheimers Disease and Related Disorders Association.
- Huff, F. J. 1988. The disorder of naming in Alzheimer's disease. In L. L. Light & D. M. Burke (Eds.), *Language, memory and aging*. New York: Cambridge University Press. Pp. 209–220.
- Huff, F. J., Mack, L., Mahlmann, J., & Greenberg, S. 1988. A comparison of lexical-semantic impairment in left hemisphere stroke patients and Alzheimer's disease. *Brain and Language*, **34**, 262–278.
- Just, M. A. & Carpenter, P. A. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122–149.
- Just, M. A., Carpenter, P. A., & Keller, T. A. 1996. The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, **103**, 773–780.
- Kemper, S., La Barge, E., Ferraro, R., Cheung, H., & Storandt, M. 1993. On the preservation of syntax in Alzheimer's disease. *Archives of Neurology*, **50**, 81–86.
- Kempler, D. 1988. Lexical and pantomime abilities in Alzheimer's disease. *Aphasiology*, **2**, 147–159.
- Kempler, D., Curtiss, S. & Jackson, C. 1987. Syntactic preservation in Alzheimer's disease. *Journal of Speech and Hearing Research*, **30**, 343–350.
- Kempler, D. 1995. Language changes in dementia of the Alzheimer type. In Lubinski (Ed.), *Dementia and communication: Research and clinical implications*. San Diego, CA: Singular Publishing Group. Pp. 98–114.
- Mace, H. L., & Rabins, P. V. 1991. *The 36-hour day: A family guide to caring for persons with Alzheimer's disease, related dementing illnesses, and memory loss in later life* (rev. ed.). Baltimore, MD: The Johns Hopkins University Press.
- Marslen-Wilson, W. D., & Tyler, L. K. 1975. Processing structure of sentence perception. *Nature*, **257**, 784–786.
- Marslen-Wilson, W. D., Tyler, L. K., & Koster, C. 1993. Integrative processes in utterance resolution. *Journal of Memory and Language*, **32**, 647–666.
- Martin, R. C. 1995. Working memory doesn't work: A critique of Miyake et al.'s capacity theory of aphasic comprehension deficits. *Cognitive Neuropsychology*, **12**(6), 623–636.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. 1984. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, **34**, 939–944.
- Miyake, A., Carpenter, P. A. & Just, M. A. 1995. Reduced resources and specific impairments in normal and aphasic sentence comprehension. *Cognitive Neuropsychology*, **12**(6), 651–679.
- Morris, R. G. & Baddeley, A. D. 1988. Primary and working memory functioning in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology*, **10**, 279–296.
- Nebes, R. D., Boller, F., & Holland, A. 1986. Use of semantic context by patients with Alzheimer's disease. *Psychology and Aging*, **1**, 261–269.
- Rochon, E., Waters, G. S., & Caplan, D. 1994. Sentence comprehension in patients with Alzheimer's disease. *Brain and Language*, **46**, 329–349.
- Schwartz, M., Marin, O., & Saffran, E. 1979. Dissociations of language function in dementia: A case study. *Brain and Language*, **7**, 277–306.
- Small, J. A., Kemper, S., & Lyons, K. 1977. Sentence comprehension in Alzheimer's disease: Effects of grammatical complexity, speech rate, and repetition. *Psychology and Aging*, **12**(1), 3–11.

- Small, J. A., Andersen, E. & Kempler, D. 1997. The effects of working memory capacity on understanding rate-altered speech. *Aging, Neuropsychology, and Cognition*, **4**, 126–139.
- Tomoeda, C. K., Bayles, K. A., Boone, D. R. & Kaszniak, A. W., & Slauson, T. J. 1990. Speech rate and syntactic complexity effects on the auditory comprehension of Alzheimer patients. *Journal of Communication Disorders*, **23**(2), 151–161.
- Tyler, L. K. 1988. Spoken language comprehension in a fluent aphasic patient. *Cognitive Neuropsychology*, **5**, 375–400.
- Tyler, L. K. 1992a. *Spoken language comprehension: An experimental approach to normal and disordered processing*. Cambridge, MA: MIT Press.
- Tyler, L. K. 1992b. The distinction between implicit and explicit language function: Evidence from aphasia. In M. Rugg & D. Milner (Eds.), *Neuropsychology of consciousness*. Academic Press.
- Tyler, L. K., Kamiloff-Smith, A., Voice, K., Stevens, T., Grant, J., Davis, M., Howlin, P. & Udwin, O. 1997. Do individuals with Williams syndrome have bizarre semantics? Evidence for lexical organisation using an on-line task. *Cortex*, **33**, 515–527.
- Tyler, L. K. & Marslen-Wilson, W. D. 1977. The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, **16**, 683–692.
- Tyler, L. K., Ostrin, R. K., Cooke, M., & Moss, H. E. 1995. Automatic access of lexical information in Broca's aphasia: Against the automaticity hypothesis. *Brain and Language*, **48**, 131–162.
- Waters, G. & Caplan, D. 1996. The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **49A**, 51–79.
- Waters, G. & Caplan, D. 1997. Working memory and on-line sentence comprehension in patients with Alzheimer's disease. *Journal of Psycholinguistic Research*, **26**, 4, 377–400.
- Waters, G., Caplan, D. & Hildebrandt, N. 1991. On the structure of verbal short-term memory and its functional role in sentence comprehension: Evidence from neuropsychology. *Cognitive Neuropsychology*, **8**, 81–126.
- Waters, G., Caplan, D., & Rochon, E. 1995. Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology*, **12**(1), 1–30.
- Wechsler, D. 1987. *Wechsler memory scale—Revised*. New York: Psychological Corporation.