

# Processing of fast speech by elderly listeners

Esther Janse<sup>a)</sup>

*Utrecht Institute of Linguistics OTS and Max Planck Institute for Psycholinguistics Nijmegen,  
Janskerkhof 13, 3512 BL Utrecht, The Netherlands*

(Received 8 April 2008; revised 21 January 2009; accepted 25 January 2009)

This study investigates the relative contributions of auditory and cognitive factors to the common finding that an increase in speech rate affects elderly listeners more than young listeners. Since a direct relation between non-auditory factors, such as age-related cognitive slowing, and fast speech performance has been difficult to demonstrate, the present study took an on-line, rather than off-line, approach and focused on processing time. Elderly and young listeners were presented with speech at two rates of time compression and were asked to detect pre-assigned target words as quickly as possible. A number of auditory and cognitive measures were entered in a statistical model as predictors of elderly participants' fast speech performance: hearing acuity, an information processing rate measure, and two measures of reading speed. The results showed that hearing loss played a primary role in explaining elderly listeners' increased difficulty with fast speech. However, non-auditory factors such as reading speed and the extent to which participants were affected by increased rate of presentation in a visual analog of the listening experiment also predicted fast speech performance differences among the elderly participants. These on-line results confirm that slowed information processing is indeed part of elderly listeners' problem keeping up with fast language. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3082117]

PACS number(s): 43.71.Lz [MSS]

Pages: 2361–2373

## I. INTRODUCTION

When people age, listening to speech can become an effortful task. Frequent complaints of elderly listeners are that they have trouble listening to speech against a background of noise or competing speech, or that their interlocutors speak too fast, or do not articulate clearly enough. The high-frequency peripheral hearing loss typically found in a large sample of elderly people (Willott, 1991) can be held as the primary cause of the speech perception problems. Additional age-related causes are central auditory processing impairments and (general) cognitive decline (information processing speed or attentional problems). The central auditory problems have been reported to be caused by peripheral hearing loss: hearing loss can impair auditory functions as frequency and temporal resolution (Glasberg and Moore, 1986; Sommers and Humes, 1993; George *et al.*, 2007). Note, however, that there is also overlap between the “central auditory” and “cognitive” accounts (Jerger *et al.*, 1991). Either way, the central auditory account seems to be a bridge between whatever is received as auditory input and how it is processed. Humes (2005) noted, on the basis of data from a large group of elderly listeners, that performance on a battery of auditory processing measures was systematically related to individual differences in cognitive, rather than auditory, function. For recognition of time-compressed words, hearing loss accounted for the greatest part of the variance, but there was a significant correlation between (non-verbal) IQ and the ability to perceive time-compressed words as well. Similar relations between auditory processing impairments and higher cognitive-level impairments can be found in the do-

mains of aphasia (Fink *et al.*, 2006) and dyslexia (Tallal, 1980; Wright *et al.*, 1997; Witton *et al.*, 1998, 2002; and Rosen, 2003 for an overview).

One approach to tease apart the relative contributions of age-related changes in hearing and cognitive factors to the increased problems elderly adults may have with speeded speech is to include several groups of listeners (Gordon-Salant and Fitzgibbons, 1993, 1995, 1999, 2001): young listeners with and without hearing loss and elderly listeners with and without hearing loss. In this way, the effects of aging and of hearing loss can be investigated separately. Younger adults with hearing loss also show substantial decrements in recognizing time-compressed speech compared to age-matched listeners with normal hearing. Gordon-Salant and Fitzgibbons (2001), Dubno *et al.* (1987), and Turner *et al.* (1997) attributed these problems to hearing impaired listeners' capacity to use brief acoustic cues. Nevertheless, Gordon-Salant and Fitzgibbons (1993, 2001) also showed that the problems in recognizing time-compressed speech that are attributed to age are independent of those attributed to hearing loss.

One other approach to evaluate the contribution of age-related hearing and cognitive factors to the increased problems elderly have with speeded speech is to speed up some parts of the signal more than others (Gordon-Salant and Fitzgibbons, 2001; Schneider *et al.*, 2005), or to allow listeners to “regain lost time” by introducing silent intervals at strategic points in the time-compressed speech material. In this latter way, listeners get time to catch up and process the information just presented at a high rate (Wingfield *et al.*, 1999). Even though allowing more time to catch up processing helped both elderly and young listeners, the elderly listeners, unlike the young, could not be brought back to base-

<sup>a)</sup>Electronic mail: E.Janse1@uu.nl

line (uncompressed) performance. These results could be taken to suggest that both auditory and cognitive factors contribute to elderly listeners' problems with fast speech. However, [Schneider et al. \(2005\)](#) argued for an auditory-only explanation, based on their results showing that the method of time compression matters. They referred to the findings of [Gordon-Salant and Fitzgibbons \(2001\)](#) that older adults find it especially difficult to deal with selective time compression of consonants because identification of consonants may depend on rapid formant transitions. Where most studies have used methods of time compression that involve removing speech segments without regard to their informational content, [Schneider et al. \(2005\)](#) used a selective method of time compression mainly affecting pauses or steady-state portions of vowels. This latter method of time compression thus presents the listener with a high rate of information content, while producing minimal acoustic distortion of the speech signal. If speech is speeded in such a selective way, elderly listeners are not affected more by an increase in rate than younger listeners ([Schneider et al., 2005](#)). In this view then, an inability of the older adult's auditory system to cope with speed-induced stimulus degradation thus suffices to explain the problem, leaving little or no room for cognitive slowing.

Given the difficulty to pinpoint the role of central auditory processing problems, a third approach has tried to account for the age group  $\times$  speed issue by investigating whether age-related cognitive slowing problems are modality-specific or not. In order to decide between central auditory and general cognitive explanations for elderly listeners' problems with fast speech, [Humes et al. \(2007\)](#) investigated whether cognitive processing is affected more in the auditory modality than in other modalities. [Humes et al. \(2007\)](#) compared speeded-speech performance to a visual text-based analog of their auditory test to investigate whether elderly participants were also more affected by increased presentation rate than younger participants when reading sentences. If amodal cognitive decline underlies performance on both reading and listening, performance should be strongly correlated because of this common underlying factor. Correlations across modalities had been found in people's ability to "fill up the gaps:" the results obtained with the text reception threshold test as an analog of the speech reception threshold test showed significant correlations between a participant's ability to identify auditory sentences in noise and written sentences masked by a bar pattern ([Zekveld et al., 2007](#); [George et al., 2007](#)). As found by [Zekveld et al. \(2007\)](#), [Humes et al. \(2007\)](#) found correlations of performance for similar tasks across modalities. Importantly, however, with respect to presentation rate, [Humes et al. \(2007\)](#) found that older adults, regardless of hearing status, were more affected by the speeded speech than young adults, whereas there were no differences between groups regarding the effect of rate of visual presentation. This latter result agreed with results by [Spehar et al. \(2004\)](#) who found equal effects of speeding for older and younger adults in a test where words had to be identified on the basis of the visual signal only. The absence of a differential rate effect in reading for the elderly in [Humes et al. \(2007\)](#) is surprising given the claim that aging slows down the cumulative dura-

tion of all processes involved in visual word recognition (ranging from basic visual analysis of letter strings to activation of lexical presentations) with a factor of approximately 1.5 ([Myerson et al., 1992](#)). [Myerson et al. \(1992\)](#) argued for this slowing factor on the basis of a meta-analysis of (visual) lexical decision studies. Since a similar performance ratio of older and younger adults had been found for naming studies ([Balota and Duchek, 1988](#)), [Myerson et al. \(1992\)](#) argued for general lexical slowing in aging. In other words, age not only affects sensory decoding of language (either in the auditory or visual domain) but also affects lexical activation patterns. Likewise, at the higher level of sentence comprehension, aging has been shown to delay semantic integration ([Federmeier et al., 2002, 2003](#)). Nevertheless, studies on aging and speech perception have often focused on elderly listeners' greater use of sentence context as a means to compensate for their poorer hearing ([Nittrouer and Boothroyd, 1990](#); [Pichora-Fuller et al., 1995](#); [Sommers and Danielson, 1999](#)). However, using context to fill up the gaps is a cognitive mechanism requiring mental resources ([McCoy et al., 2005](#)). Thus, when we aim to investigate the *relative* contributions of hearing and cognitive factors to elderly adults' problem of keeping up with a fast rate of speech, allowing elderly participants unlimited time to come up with a response may blur our conception of why perceptual processing itself has become more effortful and on why (cognitive) repair strategies are necessary.

Summing up, the literature overview above shows that it has been relatively difficult to establish which factors determine one's ability to keep up with a fast speech rate. The role of (peripheral) auditory decline seems relatively clear: having to decode an impoverished speech signal requires more time and effort, which becomes all the more problematic when speech rate is fast. However, the direct relation between non-auditory factors, such as cognitive slowing, and fast speech performance has been much more difficult to demonstrate. In order to obtain a better understanding of elderly listeners' increased difficulty with time-compressed speech, the present study took an on-line approach. Rather than focusing on the outcome of the recognition process (in terms of recall or correct identification of sentences), the present study focuses on speech processing itself. Investigating just the result of speech processing might obscure which processes led to this result. Particularly if decreased information processing rate is one of the potential underlying causes, it makes sense to tap into speech processing as the speech signal unfolds and to look at processing time, in combination with the result of processing.

Therefore, in the present study, an on-line speech processing approach was chosen to investigate how quickly listeners would be able to detect target words in running speech. Elderly and young listeners were presented with fast speech at two rates of time compression and were asked to detect pre-assigned target words as quickly as possible. An age group  $\times$  speed interaction was anticipated. The aim of the present study was to establish, within the group of elderly listeners, which factors (and to what extent) determine an individual elderly adult's ability to keep up with a fast speech rate. To this end, elderly listeners were recruited with

varying amounts of (age-induced) hearing loss. It was established how well individual hearing acuity and a number of non-auditory measures each predict elderly participants' word detection performance. The non-auditory measures involved a general measure of information processing speed [the digit symbol substitution (DSS) task] and two measures of the elderly participants' reading speed. Importantly, a number of studies have failed to find a link between elderly listeners' speech performance in taxing conditions and DSS score (Tun, 1998; Sommers and Danielson, 1999), thus challenging the impact of age-related general slowing on speech processing. This failure to find a relation might be due to the fact that most studies have used relatively short sentences (often from the revised sentence perception in noise materials) and have allowed listeners' unlimited response time to reproduce the presented sentence. The present focus on processing time may clarify whether general information processing speed, as measured by DSS performance, predicts fast speech performance in elderly participants.

Individual reading performance, as a measure of information processing rate specific to *language*, was expected to be a better predictor of fast speech performance than the more general DSS measure. If activation of lexical representations is slowed by age (due to general cognitive slowing), this should show up as a differential effect of increased rate of presentation on *reading* performance of the elderly participants as well (an age group  $\times$  speed interaction in reading). Two complementary measures of individual reading speed were selected to investigate which measure was a better predictor of one's difficulty keeping up with fast speech rates: either self-preferred reading rate (in a self-paced reading experiment) or a measure of the effect of increased rate in a word detection study in which two fast reading rates were imposed on the participants. This latter word detection experiment was a visual analog of the word detection study with fast speech: similar compression rates were used in the two sensory modalities (listening and reading) to investigate whether there would be a shared underlying component in the effect of increased speed. In both reading studies (visual word detection and self-paced reading), presentation of sentences was done in a word-by-word fashion to make reading more like listening with respect to the transient nature of speech. Obviously, word-by-word reading is far from normal or typical reading where readers are able to look back and look ahead. Nevertheless, word-by-word reading resembles speech processing more than a more typical way of reading where all words are present at the same time.

The present study was set up to test the hypothesis that both auditory and non-auditory measures predict elderly listeners' relative difficulty with fast speech rates. Hearing acuity may be the most important one in determining individual difficulty in keeping up with fast rates of speech. However, measures of general cognitive slowing and of reading rate were also expected to be significant predictors of how well one can keep up with fast rates of speech. In other words, this on-line study was expected to show that non-auditory measures indeed contribute to elderly listeners' difficulty in keeping up with fast speech rates.

## II. AUDITORY WORD DETECTION STUDY

### A. Time compression

In line with many previous studies on fast speech and aging (Tun *et al.*, 1992; Gordon-Salant and Fitzgibbons, 1993, 1995, 1999, 2001; Vaughan and Letowski, 1997; Tun, 1998; Wingfield *et al.*, 1999; Schneider *et al.*, 2005), the present study focuses on artificially time-compressed speech, and not on naturally produced fast speech. The motivation for this choice is that artificial speeding leaves articulation clarity intact, whereas in naturally produced fast speech, higher rates are (almost inevitably) accompanied by greater slurring. Given that naturally produced fast speech was shown to be less intelligible than artificially time-compressed speech (Janse, 2004), and given the literature on "speaking clearly for the hard of hearing" (Picheny *et al.*, 1985, 1986, 1989; Uchanski *et al.*, 1996), artificially speeded speech was used to avoid further complication from slurred or unclear speech.

As mentioned in the Introduction, Schneider *et al.* (2005) attributed elderly listeners' problems with time-compressed speech mainly to their inability to cope with the acoustic artifacts of the time-compression technique used in most studies. They concluded this on the basis of a comparison between the sampling technique (in which short, e.g., 10-ms, segments are deleted periodically) and a more refined technique in which only steady-state portions of the signal were removed (i.e., pauses or gaps between words or portions of steady-state vowels). However, if no age differences are found with the more refined technique of time compression, this could be due either to the absence of acoustic artifacts in, e.g., consonant transitions (as argued by Schneider *et al.*, 2005) or by the non-linear (and thus less severe) way of speeding the sentence. Compressing pauses more than the remaining speech is a form of non-linear time compression that will always lead to better intelligibility than can be obtained with linear compression because the speech itself is less compressed (Janse, 2003, and compare the results obtained with time-compression algorithm Mach 1; Covell *et al.*, 1998). The results of the study of Schneider *et al.* (2005) do not rule out the possibility that an age  $\times$  speed interaction may be found in terms of perceptual effort, or at more taxing time-compression rates.

The present study compares two methods of time compression: the sampling method as a rather crude method of deleting speech material that involves audible artifacts and the pitch-synchronous-overlap-add (PSOLA) method. PSOLA (Moulines and Charpentier, 1990), because of its high-quality time-scale manipulation, has become the technique incorporated in most speech editing programs (COOLEEDIT, PRAAT, and AUDACITY). First, a pitch detection algorithm places markers at consecutive pitch periods (also in the unvoiced portions). The signal is decomposed into separate but overlapping (Hanning) windows, with window maxima at each pitch mark location. Window length is usually twice the length of one pitch period (yielding overlap in consecutive windows). Time compression is accomplished by deleting pitch periods. The crucial asset of PSOLA time compression is that in constructing the new waveform, the



speech signal of one descending window ramp is added to that of the next ascending window ramp (with, e.g., one intervening window deleted). This means that the information contained in the deleted pitch period is averaged across now overlapping windows, thus avoiding discontinuities and spectral jumps (cf. also Letowski and Poch, 1996 and Vaughan and Letowski, 1997) and thus avoiding nasty acoustic artifacts. Any acoustic distortion can be anticipated to be more detrimental the poorer the hearing level of the listener. Inclusion of these two methods of time compression provides a comparison of their perceptual consequences for elderly listeners. Importantly, time compression was applied linearly with both time-compression techniques: the two techniques differed in whether or not they involved audible discontinuities.

Two rates of time compression were chosen: 1.5 times the original rate (compression to 67%) and 2 times the original rate (compression to 50%). These two rates were chosen on the basis of the results of a pilot listening experiment with elderly listeners. Since this pilot group of listeners, having a mean age of 80 years, hardly thought the experimental task challenging when the materials were presented at 1.5 times the normal rate, and since mean age was expected to be lower for the elderly participants in the present experiment, the 1.5 times normal-rate condition formed the baseline condition against which the fastest rate (2 times normal rate) could be compared.

Time compression was achieved by applying either PSOLA or sampling compression to the original speech materials in the sound-editing software package PRAAT. The sampling version of the 1.5 times rate condition (67%) was made by deleting every third 10-s fragment. The sampling version of the 2 times normal-rate condition (50%) was achieved by deleting every other 10-ms fragment. The sampling method yielded an audibly discontinuous sound, due to waveform discontinuities and spectral jumps at points at which now adjacent fragments were joined.

## B. Participants

The elderly participants were recruited via *Hoger Onderwijs Voor Ouderen* (“Higher Education for the Elderly”): an organization linked to several universities in the Netherlands that provides academic courses (on a variety of topics) for people of over 55 years of age. People who had subscribed for a course received an information letter in which they were asked to participate in the study. Given the academic level of the courses, the elderly group was a good match to the young (student) group in terms of education level.

There were 40 elderly participants, aged between 65 and 84, with a mean age of 70 (SE=0.7). Of the elderly participants, 18 held a university degree and 12 had finished a college of higher education (polytechnic). The 20 young participants were all students of Utrecht University (aged between 17 and 26, with a mean age of 21). All participants were native speakers of Dutch and they all received a small payment (students) or present (elderly) for their participation.

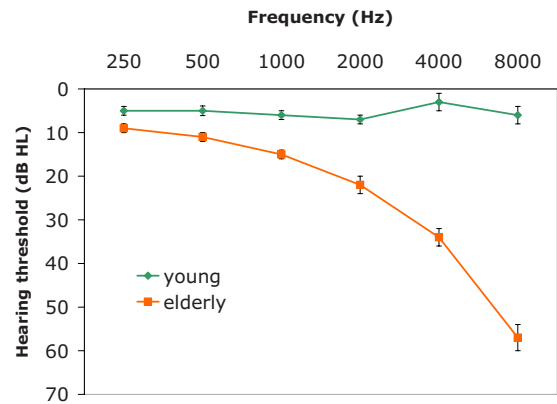


FIG. 1. (Color online) Mean pure-tone thresholds (better ear) for both age groups. Error bars reflect one standard error.

Some of the elderly participants had a hearing aid (only 4 out of 40). They were asked to remove it when participating in the listening experiment. Hearing levels of all participants were measured with a portable Maico ST20 audiometer in a silent booth of the Utrecht Institute of Linguistics Phonetics Laboratory. Figure 1 displays the mean threshold values (in decibels) measured in the better ear for both age groups.

Visual acuity was important in all studies because participants had to read target words from the computer screen (in both tested modalities). Obviously, visual acuity was particularly important in the two reading studies. Participants were asked to bring their glasses to the laboratory, if they had any, particularly the glasses they would use for looking at a computer screen. No further test was administered to measure participants' visual acuity. Large and easily legible fonts were used to minimize the chance that any subtle visual acuity differences affected the results. All participants also performed the DSS test (subpart of the *Wechsler Adult Intelligence Test*, 2004). Scores on the DSS test have been shown to exhibit strong correlations with measures involving perceptual speed or processing speed (Salthouse, 2000; Hoyer et al., 2004). Mean substitution time needed per symbol was 1.83 s/symbol (SE=0.06) for the elderly and 1.35 s/symbol (SE=0.03) for the young participants. This should be corrected for motor speed (writing time), which was 0.93 s/symbol (SE=0.06) for the elderly and 0.71 s/symbol (SE=0.03) for the young participants. The corrected coding time (substitution time minus writing time) was then 0.90 s/symbol (SE=0.05) for the elderly and 0.64 s/symbol (SE=0.04) for the young participants.

Vocabulary knowledge of all participants was tested by way of the vocabulary subpart of the Groningen Intelligentie Test (Luteijn and van der Ploeg, 1983). Out of 20 multiple choice synonym questions, mean score correct was 18 (SE=0.2) for the elderly and 15 for the young (SE=0.4) participants. Better vocabulary scores for the elderly are often found in studies in which these two age groups are compared.

## C. Material

The four conditions (two rates  $\times$  two compression types) were distributed over 120 test sentences according to a Latin

square design. This design required four different stimulus lists to which the participants were randomly assigned.

The Dutch language material consisted of 120 disyllabic nouns as targets for detection: 60 with lexical stress on the initial and 60 with stress on the final syllable. The targets were embedded in sentences that varied from 7 to 15 words in length. The sentences were news bulletin sentences that were changed such that the noun would fit the sentence. The target words were possible but not fully predictable continuations of the sentence, e.g., *He did not pay enough attention to the traffic around him*, and *But there too the supply was increasingly more extensive and expensive* (target words underlined). The target's position in the sentence ranged from the second word in the sentence to the last word. Position in the sentence was not manipulated systematically. Additionally, the sentence materials comprised 60 overlap filler sentences in which the target had word-initial phoneme overlap with one of the words in the sentence (e.g., target being *cirke* "circle" and sentence contains the word *circus*). There were also 40 miss filler sentences in which the target simply did not occur. One young male native speaker of Dutch read out all the materials at a normal speech rate (mean rate was 5.7 syllable/s). He was asked to place a sentence accent on the target word in the target sentences. The recording was made with an Audio-Technica AT 8410 microphone on a digital audio tape. All sentences were stored as separate files and downsampled to 32 kHz. Overall loudness of the audio files was equalized to 72 dB. Time compression was applied to the original recordings as discussed in Sec. II A.

#### D. Procedure

Testing was done in the Utrecht Institute of Linguistics Phonetics Laboratory in silent booths. Participants first performed the auditory word detection experiment. Then the diagnostic part was administered: the audiogram, the DSS test, and the vocabulary test. Then there was a short break after which they participated in the reading (word detection) experiment. Lastly, they performed the self-paced reading test.

In the auditory word detection study, the target words were presented visually in a large black font (Helvetica 80) on the computer screen 1200 ms before the auditory sentence started. The black target word was centered on the computer screen against a light gray background and remained on the screen during sentence presentation. Participants were instructed to respond as fast and as accurately as possible once they detected the pre-assigned visual target in the auditory sentence by pressing a button on the button box (with their dominant hand). Following sentence offset, participants still had 2 s to respond if they had not already done so, before the new visual target for the next sentence was presented. None of the participants reported feeling rushed by the pace of the experiment.

The auditory materials were presented binaurally over closed headphones (Beyer Dynamic DT 250) at a mean level of 81 dB SPL (for all listeners). Listeners were first presented with a practice part after which they could still ask questions if anything was unclear. They were then presented with the

TABLE I. Correct word detection rates in the different conditions for both age groups.

	Compressed to 1.5×normal rate (67%)		Compressed to 2×normal rate (50%)	
	PSOLA	Sampling	PSOLA	Sampling
Young	99	100	99	98
Elderly	100	100	98	92

material, blocked by speech rate (first the sentences at 1.5 times the original rate, then the even faster sentences). Within each speech rate block, items made with either type of compression were presented in random order. Order of presentation of test and filler sentences was randomized for each participant.

As a check whether participants had actually paid attention to the content of the sentences, a sentence recognition test was administered to them immediately after the listening study. Participants were presented with a list of 16 sentences: 8 of which had been presented in the listening study and 8 similar sentences which had not been presented. They were asked to tick the sentences they thought they had just heard. Elderly participants on average correctly remembered 6.1 (SE=0.2) out of 8 sentences; the young participants had an average score of 5.75 (SE=0.2). The mean number of falsely recognized sentences was 0.58 (SE=0.1) out of 8 for the elderly and 0.20 (SE=0.1) for the young participants.

#### E. Results

The overall aim of the present study was to investigate, within a group of elderly listeners, which auditory and non-auditory measures determine one's ability to keep up with a fast speech rate. Before the results of the listening study will be related to background measures and performance in the reading experiments, the results of the listening study are presented and discussed as such for the two age groups to show that the increased speech rate indeed affected the elderly listeners more than the young listeners. Thus, an age group×rate interaction was anticipated. Furthermore, an age group×compression type interaction was anticipated: the type of time compression involving acoustic artifacts was expected to be more detrimental for the elderly listeners given their generally poorer hearing level.

A word detection response was counted as correct when the participant had pressed the button after target onset (and before the time out). Responses that were given within 200 ms from target onset were also excluded (for the young listeners this last criterion excluded 8 out of 2379 observations and 2 out of 4667 observations for the elderly listeners). In Table I, correct detection rates of both participant groups are given for the different test conditions.

Table I shows that correct detection rates were always higher than 92%. False response rates to the overlap fillers were 5% and 4% in the elderly listener group at the fast (1.5 times normal rate) and fastest rate of speech (2 times normal rate), respectively. False response rate to the filler sentences which did not contain a word related to the target (the miss

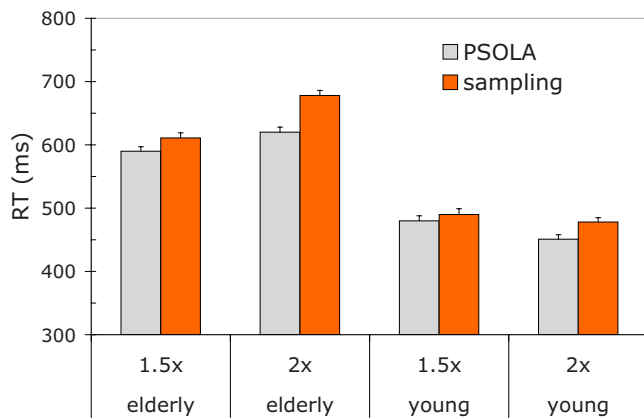


FIG. 2. (Color online) Mean response time (from target word onset) in the different rate conditions for both age groups. Error bars reflect standard errors.

condition) was 1% at both compression rates. False response rates for the young listeners were somewhat higher: 10% at both compression rates for the overlap fillers and 5% at both rates in the miss condition. These hit and false alarm rates indicate that the elderly listeners did not press the button randomly, nor did the younger listeners: collapsed over rates and conditions,  $d'$  (correct hits, relative to false alarms in overlap and miss conditions) for the elderly listeners was 3.73 and 4.50 for the young listeners.

Detection time was measured from target word onset. Mean average detection time in the different conditions is presented in Fig. 2.

The correct  $RT$  observations ( $N=7036$ ) were logtransformed. Rather than by means of two analyses of variance, one having participants as a random effect and one with items as a random effect, mixed-effect modeling was used to analyze the data. The mixed-effect method allows one to acknowledge that multiple crossed random factors (participants and items) affect performance in psycholinguistic designs as the one used here (Pinheiro and Bates, 2000). Similar mixed-effect approaches were taken in acoustic measurement studies of speech corpus data in which speaker and item are multiple random factors (Pluymaekers *et al.*, 2005; Ernestus *et al.*, 2006; Quené, 2008). Mixed-effect modeling is argued to be superior to repeated measures or univariate analysis of variance (ANOVA) in a number of tutorial articles (Baayen *et al.*, 2002; Quené and van den Bergh, 2004, 2008; Baayen *et al.*, 2008).

The log  $RT$  data were analyzed here to test for the effects of age group, rate, compression type (PSOLA/sampling), and stress position and possible interactions between them. For all significant effects and interactions, estimated coefficients ( $\beta$ 's), with standard errors for  $\beta$  in brackets, and  $p$ -values are reported.<sup>1</sup> The results showed a main effect of age group [estimated coefficient = 0.092(0.029),  $p < 0.001$ ]: the elderly participants were generally slower than the young. There was a main effect of rate [ $\beta = -0.025(0.009)$ ,  $p < 0.005$ ], and more importantly, there was a significant age  $\times$  rate interaction [ $\beta = 0.045(0.010)$ ,  $p < 0.001$ ]. As can be seen in Fig. 2, whereas the young listeners responded faster in the 2 times normal-rate condition

than in the 1.5 times normal-rate condition, the reverse is true for the elderly. It is important to note that if one measures  $RT$ 's from target word *offset*, rather than from target *onset*, all participants have increased  $RT$ 's when speech rate is increased further (measured from target word offset, and collapsed over compression types, the young participants have a mean  $RT$  of 199 ms in the 1.5 times normal rate and 249 ms in the 2 times normal-rate condition, whereas the elderly participants go from 314 to 432 ms in the fastest condition). There was no overall effect of compression type [ $\beta = 0.003(0.009)$ ,  $n.s.$ ]. However, there was a significant interaction between rate and compression type [ $\beta = 0.030(0.012)$ ,  $p < 0.05$ ]: the  $RT$  advantage of PSOLA over the sampling method was larger at the fastest (2 times original) rate than at the intermediate rate (1.5 times original). However, there was no interaction between age and compression type, nor a three-way interaction between age, compression type, and rate. Lastly, the effect of stress position of the target items was not significant overall, nor its interaction with age group [ $\beta = -0.014(0.010)$ ,  $n.s.$ ].

A very similar pattern of significant main effects and interactions was found when  $RT$  results were analyzed from target offset, rather than from target onset.

Accuracy of the responses in the different conditions was analyzed as well (note that accuracy was high overall). This was done with mixed-effect modeling for binomial data (responses being either correct or false). Quené and van den Bergh (2008) showed that mixed-effect logistic models provide a more conservative and appropriate method to analyze binomial data than the traditional practice of computing proportions correct over subjects or items. The effects of age group, rate, compression type, and stress position were tested and their interactions (estimated coefficients are reported, with their standard errors in brackets, and  $p$ -values). The mixed-effect model showed a significant effect of age group [the elderly having lower hit rates than the young:  $\beta = -1.439(0.382)$ ,  $p < 0.001$ ]. The main effect of rate [ $\beta = -1.954(0.503)$ ,  $p < 0.001$ ] was significant: detection accuracy was decreased at the fastest speech rate. There was no main effect of compression type [ $\beta = 0.160(0.673)$ ,  $n.s.$ ], but the interaction between rate and compression type was significant [ $\beta = -1.771(0.714)$ ,  $p < 0.05$ ]. Performance of both age groups was more affected by the acoustic artifact at the faster than at the intermediate (1.5 times normal) rate of speech. None of the other effects or interactions (including the interaction between age group and rate) was significant.

Thus, in the detection rate analysis and in the response time analysis, the effect of compression type was more disrupting at the faster rates, and this was found for both age groups.

The two reading studies will now be presented first, before the fast speech results are related to reading performance and to the background measures.

### III. VISUAL WORD DETECTION STUDY

#### A. Participants

The participants in the two reading studies (visual word detection study and self-paced reading) were the same as



those who participated in the listening study. Due to technical failure during one test session, reading performance of 1 of the 40 elderly participants who had completed the listening experiment was not available. The results reported below are based on performance of 40 elderly listeners (thus, one replacement participant was included from the same elderly group).

## B. Materials

One major complication in comparing rate in the auditory and visual modalities is that auditory presentation rate can always be related to “normal” or original speech rate: the rate at which the speaker uttered the sentence. In the visual modality, normal reading rate is up to each individual reader. On the basis of several studies on reading performance of elderly adults with sequential (word-by-word) presentation methods (Stine, 1990; Speranza *et al.*, 2000; Hartley *et al.*, 1994), normal presentation duration was estimated to be 400 ms per word. This estimation was on the safe side since this rate was to be speeded up.

The visual word detection study had an analogous set-up and similar sentence materials as the auditory word detection study. Sixty bisyllabic target words, always morphologically simplex nouns, were embedded in sentences in which they were not completely predictable from the sentence context. As in the listening experiment, half of the target items had initial stress; the other half had final stress. The sentence material was presented word-by-word: each individual word was displayed on the screen equally long. Given that normal reading rate was estimated to be 400 ms per word, the two rate conditions were made by speeding up presentation duration to 267 (67% rate condition) or 200 ms (50% rate condition) per word. Sentences were 8–19 words long. None of the words in the sentences had more than 12 letters. The two presentation rate conditions were distributed over the 60 test sentences according to a Latin square design. This design required two different stimulus lists to which participants were randomly assigned.

As in the auditory experiment, there were also filler trials: overlap fillers containing a noun that showed word-initial overlap with the target noun and miss fillers not containing a word that (partly) matched the target. The ratio of overlap and miss fillers to test sentences was identical to that in the auditory experiment: in addition to the 60 test sentences, there were 40 overlap trials and 26 miss trials.

## C. Procedure

As in the auditory word detection experiment, participants were first presented with a visual target word, 1000 ms before the first word of the sentence was shown. The target word had a different font (Courier bold which has serifs; 50 point) and color (dark green) than the words of the sentence, and was in upper case, unlike the words of the sentence. This was done to make the target word as distinctly different as possible from the sentence such that participants would really have to read the sentence and not be able to detect the target word purely on the basis of superficial “image” matching. The sentence words were in a black Helvetica 50 point

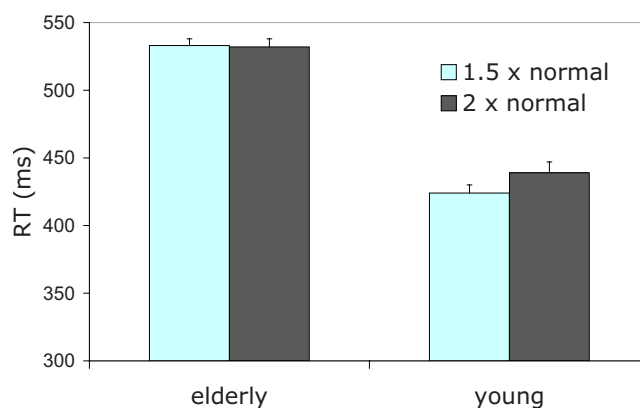


FIG. 3. (Color online) Mean response time in visual word detection experiment at both rates of presentation for both age groups (error bars reflect standard errors).

letter (this font has no serifs) on the screen: the words appeared in a white box against a light gray screen background. The target word remained on the screen during sentence presentation, clearly distinct from and under the box in which the words of the sentence were presented. Participants had 2 s following sentence offset to respond before the new visual target for the next sentence was presented. They were instructed to respond as fast and as accurately as possible once they detected the pre-assigned target in the sentence by pressing a button on the button box (with their dominant hand). They were informed that there would be sentences not containing the target word or with words similar to the target word. Participants were first presented with a practice part of eight sentences (containing both four test and four filler sentences in which the target would not appear: two misses and two with an overlapping word) after which they could still ask questions if anything was unclear. They were then presented with the material, blocked by speech rate (first the moderately fast sentences, then the faster sentences). Order of presentation of test and filler sentences was randomized for each participant.

## D. Results

As in the auditory study, the increased rate of visual presentation was expected to have a more pronounced effect on the elderly than on the young participants. However, the two rates of visual presentation in the word detection study were hardly reported as challenging by the elderly participants. As in the listening study, response time was computed from target onset. Target detection times are given in Fig. 3 for both age groups and both presentation rates (1.5 and 2 times the normal rate, or compression to 67% and 50% of the normal duration, respectively). Correct detection rates were 99% for the elderly (at both presentation rates) and 100% for the young participants (at both rates).

The (correct) detection times ( $N=3571$ ) were logtransformed and analyzed with linear mixed-effect models to test the effects of age group, presentation rate, and stress position of the target item. As in Sec. II E, estimated coefficients are reported, plus standard error for the estimate between brackets, and  $p$ -values. Elderly participants generally responded

more slowly than the young listeners [age group effect:  $\beta = 0.099(0.024)$ ,  $p < 0.001$ ], but there was no overall effect of rate of presentation [ $\beta = 0.005(0.007)$ , *n.s.*]. Responses to initially stressed items (word-initial stress being the default stress pattern in Dutch) were overall faster than to finally stressed items [main effect of stress position  $\beta = 0.029(0.010)$ ,  $p < 0.005$ ], and this effect was slightly more pronounced for the young than the elderly listeners [ $\beta = -0.023(0.009)$ ,  $p < 0.05$ ].

There was no specific hypothesis concerning the effect of stress position for the different age groups. Nevertheless, these reading results show that the elderly participants are less affected by stress position of the lexical item than the young participants. Increased exposure over the years to words that do not match the default stress pattern might explain this result.

The age group  $\times$  presentation rate interaction was not significant [ $\beta = 0.002(0.009)$ , *n.s.*]: if anything, the elderly participants as a group seemed to show less influence of the increased presentation rate than the young (cf. Fig. 3). None of the other interactions reached significance.

#### IV. SELF-PACED READING

##### A. Participants

Participants were the same as those who participated in the auditory word detection study and in the visual word detection study (cf. Sec. III A).

##### B. Stimuli and procedure

In the self-paced reading experiment, participants were asked to read a short folk story (about a monkey) at their own preferred pace. With each press on a response button of a button box, the next word of the sentence would appear (or the first word of the next sentence). Even though each word disappeared before the next word was presented, the sentence was presented “from left to right:” each new word appeared in the position in which it would have been if the sentence had been presented as a whole at once. The dependent variable is reading time per word, or the pause between consecutive button presses. The words were presented in a white Helvetica font (size 30) against a black background (rectangle in which the sentence would fit). This sentence box was positioned just above the middle of the screen.

Participants received an oral instruction how to perform the task. By way of a practice set, the first sentence pair introduced the task to the participants once more: “You are going to read a short story now.” “You decide what the tempo will be by pressing the button on the button box.” If the participants had no more questions, they proceeded to the test part. The short story consisted of ten sentence pairs (the introductory practice sentence pair excluded), each of which consisted of two sentences. These sentences varied in length between 7 and 18 words.

##### C. Results

Overall slower (i.e., longer) reading times were expected for the elderly participants than for the young participants. In

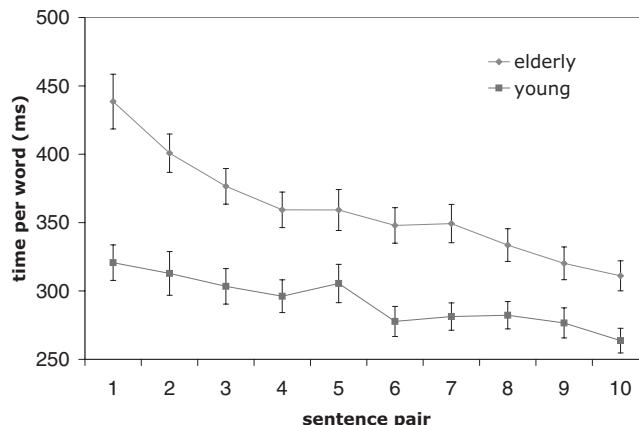


FIG. 4. Mean reading time per word per sentence pair for both age groups (error bars represent 1 SE).

Fig. 4 mean reading time per word is displayed for the two age groups as a function of sentence pair number. Because the story makes more and more sense toward the end, mean reading time per word shows a general downward trend. An ANOVA (by subjects) was carried out on these data because mixed-effect models (with both subjects and items as random factors) did not seem appropriate here. The (logtransformed) reading time per word data showed a main effect of age group [ $F(1,58) = 4.85$ ,  $p = 0.032$ ] and a main effect of sentence pair number [ $F(9,49) = 43.07$ ,  $p < 0.001$ ]. As Fig. 4 suggests, the interaction between age group and sentence pair is also significant [ $F(9,49) = 6.21$ ,  $p < 0.001$ ]: the difference in reading time between young and elderly listeners becomes smaller over the course of the short story.

The purpose of this self-paced reading experiment was to investigate a relationship between self-preferred reading rate and performance in the listening study. Therefore, each individual participant’s reading pace to be entered as a predictor for fast speech performance was not computed over the entire story, but was restricted to the first two sentences because these set up a new semantic context frame. The later sentences introduce new content as well, but also refer back to concepts and persons introduced earlier. This fact that the first two sentences of the story introduce more new persons and concepts than later sentences makes them more similar to the sentence materials in the two word detection studies where every sentence is completely unrelated to the previous one and sets up its own new context. Mean reading time per word (computed over the first two sentence pairs) was 317 ms per word for the young adults (SE=10) and 419 ms for the elderly adults (SE=10, but note that 4 out of 160 observations were discarded because they were more than three SEs removed from the mean). An ANOVA on (logtransformed) reading times showed that this reading time difference between the age groups was significant [ $F_1(1,58) = 13.67$ ,  $p < 0.001$ ;  $F_2(1,3) = 155.00$ ,  $p = 0.001$ ]. Computed over the entire story, mean reading time per word was 292 ms for the young adults (SE=4) and 359 ms for the elderly adults (SE=4).

#### V. COMBINED RESULTS

The following three reading measures were considered potential predictors for fast speech performance: partici-



TABLE II. Correlation matrix with correlation coefficients (Pearson's) between all predictor pairs for the group of 39 elderly participants. Asterisk notation “\*”:  $p < 0.05$ .

	Hearing acuity	Age	Digit symbol	Vocab	Self-paced reading	log RT (visual word detection)	Rate effect (visual word detection)
Hearing acuity		0.27	-0.17	-0.05	-0.10	0.31	0.35*
Age	0.27		-0.08	0.25	0.14	0.11	0.12
Digit symbol	-0.17	-0.08		-0.19	0.10	-0.08	0.04
Vocabulary score	-0.05	0.25	-0.19		0.25	0.09	-0.16
Self-paced reading	-0.10	0.14	0.10	0.25		0.05	0.15
Overall log RT in visual word detection study	0.31	0.11	-0.08	0.09	0.05		0.34*
Rate effect in visual word detection	0.35*	0.12	0.04	-0.16	0.15	0.34*	

pants' self-paced reading time (log time per word), overall target detection time (log) in the visual word detection study, and the effect of increased reading rate. Even though there was no overall effect of increased presentation rate in the visual word detection study, there were individual differences in the effect of presentation rate: the effect was computed for each of the elderly participants by dividing their mean (log) response time in the fastest presentation rate by their mean log RT in the slower presentation rate. As said, due to technical failure, this combination of reading and listening data applied to the data of 39 out of 40 elderly individuals. The other background measures were age, vocabulary score, hearing acuity, and corrected DSS score. For each listener, mean hearing loss in the better ear was computed by averaging pure-tone averages over 1, 2, and 4 kHz. Vocabulary performance was entered as the proportion of correct responses. Age was entered after subtraction of the minimum age of 65 years (because values of all other factors and predictors in the model are estimated for the minimum age value which would be far outside the actual age range if minimum age is not subtracted; cf. Snijders and Bosker, 1999 for the issue of centralization of predictor values). With respect to DSS test performance, each individual's corrected coding time (time required to recode one symbol) was entered.

Correlations among the different background measures were computed within the group of 39 elderly to address the issue of collinearity. If they are highly correlated, it does not make sense to consider them jointly in a multiple regression model of elderly participants' fast speech performance. Table II presents a correlation matrix with Pearson correlation coefficients between all predictor pairs (one asterisk denoting that  $p < 0.05$ ).

Given the weak (if any) correlations among these background measures, they were all entered into the analyses. The listening results were analyzed by using linear mixed-effect models with subject and item as crossed random effects (Pinheiro and Bates, 2000; Quené and van den Bergh, 2004; Bates and Sarkar, 2005). An additional advantage of mixed-effect models, over the advantage that multiple random effects can be considered in one analysis, is that both categorical and discrete predictors can be included in a single model (e.g., a manipulated factor, such as compression type, and individual characteristics, such as hearing acuity). Several models were fitted to the binomial detection accuracy data (using logistic regression, valid correct and incorrect

responses made up 4678 observations), and several models were fitted to the detection time data (correct responses only,  $N=4549$ ). It is important to note that in linear mixed-effect models, the order in which the predictors are entered does not influence their predictive value. The models evaluate whether coefficients are significantly different from zero in a model that contains all other predictors.

### A. Detection rate

In Table III, an overview is given of four tested models. Model 0 is the simplest linear mixed-effect model which only contains as fixed factors the manipulated factors compression rate, compression type, and their interaction, and takes participant and item as random effects (as will be the case in all later models as well). In model 1, hearing acuity is added to the fixed part see how its inclusion improves the previous model. In model 2, self-paced reading time is added to model 1, and lastly, digit symbol coding time is added to the previous model in model 3. Table III also contains an evaluation of the four consecutive models in terms of goodness of fit. This is done by the likelihood ratio test, rather than by a computation of  $R^2$  for each of the models because in mixed-effect models different sources of variation are modeled jointly (fixed and random effects). If  $R^2$  were to be computed, it would not only cover the variance explained by the fixed effects (as is normally the case in multiple regression analysis, e.g., Jerger *et al.*, 1991) but also that of the random effects. Table III also shows how inclusion of the background measure predictors in the later, more complex models leaves less and less random between-subject variance (relative to model 0).

The log likelihood ratio test takes log likelihood, a measure of the model's goodness of fit, for the simpler model (containing less predictors) and compares it with the log likelihood for the more complex model with more parameters. The difference between the two log likelihoods, multiplied by 2, follows a chi-square distribution with the difference in number of parameter as number of degrees of freedom. In other words, the output of this likelihood ratio test determines whether inclusion of more parameters is justified. In Table III, each later model is evaluated relative to the previous model, and each model has a better fit than the previous one. Asterisks denote which of the coefficients are significant within that model.

This model comparison shows that model 3 contains

TABLE III. Comparison of several mixed-effect models of the accuracy data: estimated parameters ( $\beta$  coefficients) for the fixed part, variance of the two random effects, and evaluation characteristics (significant codes: “\*\*\*”:  $p < 0.001$ , “\*\*”:  $p < 0.01$ , “\*”:  $p < 0.05$ , and “.”:  $p < 0.1$ ).

		Model 0	Model 1	Model 2	Model 3
Fixed effects	Compression rate	-2.26***	-2.23***	-2.25***	-2.27***
	Compression type	-0.22	-0.22	-0.24	-0.25
	Compression rate $\times$ compression type	-1.48.	-1.50.	-1.48.	-1.48.
	PTA		-0.06***	-0.06***	-0.07***
	SPR (log reading time)			-3.23*	-4.27*
	Digit symbol coding time				-1.05*
Random	Participant variance	1.06	0.50	0.38	0.27
	Item variance	5.43	5.46	5.54	5.59
Evaluation	Log likelihood	-413.08	-404.46	-401.55	-399.37
	Chi-square (compared to previous model)		17.242	5.84	4.35
	$p$ -deviance (compared to previous model)		<0.001	<0.05	<0.05

predictor values that show significant effects over and above the manipulated factors. The other background measures (age, vocabulary score, mean log  $RT$  in reading, and the rate effect in reading) turned out not to predict fast speech performance.<sup>2</sup> Interaction terms between design factors and predictor variables were also added, but this did not improve the model.

Thus, the model with the best fit, model 3, showed the following effects (as in previous discussions of linear mixed-effect results for binomial data, estimated coefficients are reported plus their standard errors, and significance levels). Compression rate had a significant main effect on detection rate [estimated coefficient (henceforth  $\beta$ ) is negative, meaning that the higher the speech rate, the lower the detection rate, compared to 1:  $\beta = -2.270(0.604)$ ,  $p < 0.001$ ], but the main effect of compression type was not significant [ $\beta = -0.250(0.750)$ , *n.s.*]. The interaction between compression rate and compression type approached significance [ $\beta = -1.477(0.797)$ ,  $p = 0.064$ ]. Hearing acuity [pure tone average (PTA)] significantly influenced detection rate: the more hear-

ing loss, the lower the detection rate [ $\beta = -0.067(0.012)$ ,  $p < 0.001$ ]. In addition, mean reading time per word (logtransformed, self-paced reading test) significantly influenced detection rate: the longer the reading time per word, the lower the detection rate [ $\beta = -4.270(1.339)$ ,  $p < 0.01$ ].<sup>3</sup> Lastly, performance in the DSS test (as a measure of information processing rate) also significantly influenced performance: the longer one took to recode a digit to a symbol, the lower the detection rate [ $\beta = -1.050(0.471)$ ,  $p < 0.05$ ]. Thus, hearing acuity, self-preferred reading rate, and information processing rate (in this order of relative importance) predicted success of target detection in fast speech.

## B. Response time

Several linear mixed-effect models were fitted to the log  $RT$  data of the fast speech study (4547 observations). The following measures did not predict response time (nor interacted with any of the factors): vocabulary score, age, and hearing acuity. Table IV lists the characteristics of three

TABLE IV. Comparison of several mixed-effect models of the log  $RT$  data: estimated parameters ( $\beta$  coefficients) for the fixed part, variance of the two random effects, and evaluation characteristics.

		Model 0	Model 1	Model 2
Fixed effects	Compression rate	0.02***	0.02***	-0.81***
	Compression type	0.01*	0.01*	0.01*
	Compression rate $\times$ compression type	0.03***	0.03***	0.03***
	Digit symbol coding time		0.002	-0.01
	log $RT$ in visual word detection study		0.89***	0.89***
	Rate effect in visual word detection		-1.09	-1.48
	Self-paced reading time		0.22.	0.22.
	Compression rate $\times$ digit symbol coding time			0.03**
	Compression rate $\times$ rate effect in visual detection			0.81***
	Random	Participant variance	0.01	0.007
Item variance		0.003	0.003	0.003
Evaluation	Log likelihood	3478	3491	3501
	Chi-square (compared to previous model)		24.42	20.73
	$p$ -deviance (compared to previous model)		<0.001	<0.001

models: again moving from the simplest model (model 0) that only contains the interaction between the design factors, to a model that contains more predictors (model 1), to a model that contains interactions between factors and predictors (model 2).

The evaluation part of Table IV shows that adding background measures (model 1) improved the model, relative to the simplest model (model 0). The asterisk notation shows that model 1 is better than model 0 only because two, rather than four, background measures (namely, log *RT* in reading and the self-paced measure) were significant: the other two background measures were not significant predictors in model 1. The latter two background measures turn out to be important in their interaction with the compression rate effect in model 2, however. One should also note that the value of the compression rate coefficient was positive in models 0 and 1, meaning that generally, response times became longer in the faster rate condition (as was shown in Fig. 2). The direction of the rate effect is reversed in model 2: now, the faster the speech rate, the shorter the response times ( $\beta = -0.809$ ), as was also seen for the young adults. Thus, the best-fitting model (model 2) showed a significant main effect compression rate on log *RT* [ $\beta = -0.809(0.227)$ ,  $p < 0.001$ ], and a main effect of compression type [ $\beta = 0.011(0.004)$ ,  $p < 0.05$ ], indicating that response times were longer in the sampling conditions than in the PSOLA time-compressed conditions. This effect was stronger at the faster speech rate, as shown by a significant interaction between rate of speech and compression type [ $\beta = 0.032(0.006)$ ,  $p < 0.001$ ].

Mean detection *RT* in the visual word detection experiment was also a significant predictor: the longer one's detection time in reading, the longer one's detection time in listening [ $\beta = 0.892(0.179)$ ,  $p < 0.001$ ]. This reflects the shared components in the two analogous studies: first in terms of basic motor speed (how fast can this individual press the response button) and second in terms of speed of language processing and decision making. Self-paced reading time was another predictor that was very close to significance: the longer the reading time, the longer the detection time in listening [ $\beta = 0.218(0.119)$ ,  $p = 0.067$ ].<sup>4</sup> There was no overall effect of digit symbol coding time on *RT*'s, but it did interact with the compression rate effect [ $\beta = 0.027(0.010)$ ,  $p < 0.01$ ]. Whereas listeners generally have shorter response times in the fastest rate condition (relative to the 1.5 times normal-rate condition, cf. the negative coefficient for compression rate), this effect is the more counteracted, the longer time one needed to recode a digit to a symbol. A similar interaction was seen between compression rate and the rate effect in the visual word detection experiment, which in itself did not have a main effect on detection *RT*. Again, whereas *RT*'s generally get shorter in the fastest speech condition, this effect is counteracted more for those participants with a larger rate effect in reading [ $\beta = 0.810(0.227)$ ,  $p < 0.001$ ]. This indicates that participants who responded relatively slowly at the faster rate of visual presentation did not speed up in the listening study either when presented with the faster rate of speech.

Thus, the model fitted to the response times mainly showed interesting predictors for rate effect size: the extent

to which participants were affected (i.e., sped up or slowed down) by increased rates of speech. Importantly, hearing acuity was not related to increased fast speech difficulty as measured by response time, but keep in mind that hearing level was the most important measure in explaining detection success. The model did show that increased difficulty with faster rates of speech was predicted by (a) the effect of presentation rate in the visual word detection experiment and (b) by information processing rate (DSS task performance).

## VI. DISCUSSION AND CONCLUSION

The present study was set up to investigate why elderly listeners are more affected than young listeners when they are presented with fast, or more particularly, time-compressed speech. This is an old question that has given rise to a large number of studies (Gordon-Salant and Fitzgibbons, 1993, 1995, 1999, 2001; Letowski and Poch, 1996, Tun, 1998, Wingfield *et al.*, 1999; Schneider *et al.*, 2005). Even though many studies have acknowledged the contributions of hearing and cognitive factors to this age  $\times$  rate interaction, it has been relatively difficult to quantify their relative contributions. The idea behind the present study was that if age-related general cognitive slowing and decreased information processing rate may be (partly) responsible for elderly listeners' increased difficulty with time-compressed speech, it makes sense to focus on this effortful process itself, rather than on just the result of it. Allowing elderly participants unlimited time to come up with a response may blur our conception of why perceptual processing itself has become more effortful. Moreover, the use of an on-line measure yields complementary accuracy and response time data.

The present study showed significant correlations between, on the one hand, hearing acuity and fast speech performance and, on the other hand, between individual presentation rate effects in word-by-word reading and in speech. Hearing loss clearly plays a primary role in explaining elderly listeners' problems with fast speech, as was evident from the detection rate analysis. Having to decode an impoverished speech signal requires more effort and thus takes more time, which becomes all the more problematic when speech rate is fast.

Note that the present data replicated the finding of Humes *et al.* (2007) that overall performance of the elderly participants was not affected by rate of visual presentation. Clearly, increased presentation rate affects cognitive processing more in the auditory modality than in the visual modality. This ties in with the suggestion by Humes *et al.* (2007) that high-frequency hearing loss is associated with central auditory problems, such as poorer spectral and temporal resolution, over and above the loss of audibility. In their results, high-frequency hearing loss turned out to be a significant predictor of fast speech performance, even after spectral shaping had been applied to compensate for loss of audibility. Nevertheless, despite the fact that overall performance of the elderly participants was not affected by rate of presentation, the present focus on processing time has resulted in two important findings. Even though correlations in performance across modalities have been found before (Zekveld *et al.*,



2007; Humes *et al.*, 2007), the present study is the first to show correlations between *rate effects* across modalities. Furthermore, the present results provide support for a direct relationship between individual differences in processing speed (DSS) and fast speech perception (cf. Sommers, 2005). It is not surprising that performance in the two reading tasks, being language measures, turned out to be better predictors than the more general cognitive measure of information processing rate (DSS).

The combination of correlations between fast speech performance and self-paced reading rate (detection accuracy analysis) and between individual rate effects in the two analogous word detection studies (*RT* analysis) provides support for an underlying problem with fast language processing: language being a particularly complex type of information to process. At first sight, it seems that the present study's "fast rates" of visual presentation were not fast or challenging enough. Note, however, that normal reading rate for the elderly listeners was estimated at 400 ms per word, which turned out to be remarkably close to elderly participants' mean self-preferred reading time (419 ms per word) in the first sentences of the folk story.

Self-paced reading may have given a more reliable insight into individual speed of language processing than the visual target detection study. It seems unlikely that participants ignored the content of the sentence while detecting pre-assigned target words: the target item that elicited the fastest detection time may have been more predictable from its preceding sentence context than average (*All of a sudden tapped he with a pencil on the window*). However, there was no check on whether all participants actually remembered some of the visually presented sentences afterward, which leaves the possibility that (some) participants only scanned the sentences for an orthographic match to the target. Even though self-paced reading had the better predictive power of the two for fast speech performance (cf. the individual detection rate analysis), the related effects of speeding in the visual and auditory modalities provide support that there is a "general cognitive" component to the age  $\times$  rate interaction in listening to fast speech.

This brings us back to the claim by Schneider *et al.* (2005) that elderly listeners' problems with speeded speech should be attributed mainly to their inability to cope with the acoustic artifacts involved in methods of time compression, rather than to cognitive slowing. The present results did not show a clear age group  $\times$  compression type interaction, but Schneider *et al.* (2005) may be right that, in general, any acoustic artifact is more disrupting for elderly than young listeners. Yet, the rate correlations across modalities provide solid evidence that there is more to the age  $\times$  rate interaction than acoustic artifacts. Future research could investigate the potential benefits of non-linear over linear time compression, given their finding that elderly listeners were not more affected if speeding was accomplished by only removing silences and steady-state portions of the signal.

In conclusion, the present study has provided direct evidence that auditory and cognitive factors underlie elderly listeners' problems keeping up with fast rates of speech. Hearing loss, combined with central auditory problems that

may be linked to high-frequency hearing loss, is certainly the primary cause of the problem. Nevertheless, be it at a more subtle level, general information processing problems also affect rate of language processing in auditory and non-auditory domains.

## ACKNOWLEDGMENTS

The Dutch Organisation for Scientific Research (NWO) is gratefully acknowledged for funding this research (Grant No. 275-75-004). I would like to thank Sieb Nooteboom for his comments on earlier drafts of this study and Mirjam Ernestus and Hugo Quené for their statistical advice.

<sup>1</sup>*p*-values (obtained by generating 10 000 simulations of the data sample) were computed with an ancillary function *pvals.fnc* defined in the LAN-GUAGER package. This function can be applied to the fitted model (<http://cran.r-project.org>, see Baayen, 2008).

<sup>2</sup>When raw age was in the model (instead of age minus 65) it was not a significant predictor either.

<sup>3</sup>A somewhat weaker, but still significant, predictor of detection rate is individual self-paced reading time per word averaged over the entire story, rather than averaged over only the first two sentences of the story [ $\beta = -2.899(1.187)$ ,  $p = 0.015$ ].

<sup>4</sup>If self-paced reading time is averaged over the entire story, rather than over the first two sentences, it is not a significant predictor [ $\beta = 0.119(0.098)$ , *n.s.*].

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics* (Cambridge University Press, Cambridge).

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). "Mixed-effects modeling with crossed random effects for subjects and items," *J. Mem. Lang.* **59**, 390–412.

Baayen, R. H., Tweedie, F. J., and Schreuder, R. (2002). "The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon," *Brain Lang.* **81**, 55–65.

Balota, D. A., and Duchek, J. M. (1988). "Age-related differences in lexical access, spreading activation, and simple pronunciation," *Psychol. Aging* **3**, 84–93.

Bates, D., and Sarkar, D. (2005). "The lme4 library," available online at <http://lib.stat.cmu.edu/R/CRAN/> (Last viewed April, 2008).

Covell, M., Withgott, M., and Slaney, M. (1998). "Mach1: Nonuniform time-scale modification of speech," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Seattle, pp. 493–496.

Dubno, J. R., Dirks, D. D., and Schaefer, A. B. (1987). "Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition," *J. Acoust. Soc. Am.* **81**, 1940–1947.

Ernestus, M., Lahey, M., Verhees, F., and Baayen, R. H. (2006). "Lexical frequency and voice assimilation," *J. Acoust. Soc. Am.* **120**, 1040–1051.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., and Kutas, M. (2002). "The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study," *Psychophysiology* **39**, 133–146.

Federmeier, K. D., Schwartz, T. J., Van Petten, C., and Kutas, M. (2003). "Sounds, words, sentences: Age-related changes across levels of language processing," *Psychol. Aging* **18**, 858–872.

Fink, M., Churan, J., and Wittmann, M. (2006). "Temporal processing and context dependency of phoneme discrimination in patients with aphasia," *Brain Lang.* **98**, 1–11.

George, E. L. J., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., and Houtgast, T. (2007). "Auditory and nonauditory factors affecting speech reception in noise by older listeners," *J. Acoust. Soc. Am.* **121**, 2362–2375.

Glasberg, B. R., and Moore, B. C. (1986). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," *J. Acoust. Soc. Am.* **79**, 1020–1033.

Gordon-Salant, S., and Fitzgibbons, P. (1993). "Temporal factors and speech recognition performance in young and elderly listeners," *J. Speech Hear. Res.* **36**, 1276–1285.

Gordon-Salant, S., and Fitzgibbons, P. (1995). "Recognition of multiply

- degraded speech by young and elderly listeners," *J. Speech Hear. Res.* **38**, 1150–1156.
- Gordon-Salant, S., and Fitzgibbons, P. (1999). "Profile of auditory temporal processing in older listeners," *J. Speech Lang. Hear. Res.* **42**, 300–311.
- Gordon-Salant, S., and Fitzgibbons, P. J. (2001). "Sources of age-related recognition difficulty for time-compressed speech," *J. Speech Lang. Hear. Res.* **44**, 709–719.
- Hartley, J. T., Stojack, C. C., Mushaney, T. J., Annon, T. A. K., and Lee, D. W. (1994). "Reading speed and prose memory in older and younger adults," *Psychol. Aging* **9**, 216–223.
- Hoyer, W. J., Stawski, R. S., Wasylyshyn, C., and Verhaeghen, P. (2004). "Adult age and digit symbol substitution performance: A meta-analysis," *Psychol. Aging* **19**, 211–214.
- Humes, L. E. (2005). "Do 'auditory processing' tests measure auditory processing in the elderly?," *Ear Hear.* **26**, 109–119.
- Humes, L. E., Burk, M. H., Coughlin, M. P., Busey, T. A., and Strauser, L. E. (2007). "Auditory speech recognition and visual text recognition in younger and older adults: Similarities and differences between modalities and the effects of presentation rate," *J. Speech Lang. Hear. Res.* **50**, 283–303.
- Janse, E. (2003). *Production and Perception of Fast Speech* (Landelijke Onderzoeksschool Taalwetenschap, Utrecht).
- Janse, E. (2004). "Word perception in fast speech: Artificially time-compressed vs. naturally produced fast speech," *Speech Commun.* **42**, 155–173.
- Jerger, J., Jerger, S., and Pirozzolo, F. (1991). "Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly," *Ear Hear.* **12**, 103–109.
- Letowski, T., and Poch, N. (1996). "Comprehension of time-compressed speech: Effects of age and speech complexity," *J. Am. Acad. Audiol.* **7**, 447–457.
- Luteijn, F., and van der Ploeg, F. A. E. (1983). *Handleiding Groninger Intelligentietest (Manual Groningen Intelligence Test)* (Swets and Zeitlinger, Lisse, The Netherlands).
- McCoy, S. L., Tun, P. A., Cox, C. L., Colangelo, M., Stewart, R. A., and Wingfield, A. (2005). "Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech," *Q. J. Exp. Psychol. A* **58**, 22–33.
- Moulines, E., and Charpentier, F. (1990). "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones," *Speech Commun.* **9**, 453–467.
- Myerson, J., Ferraro, F. R., Hale, S., and Lima, S. D. (1992). "General slowing in semantic priming and word recognition," *Psychol. Aging* **7**, 257–270.
- Nittrouer, S., and Boothroyd, A. (1990). "Context effects in phoneme and word recognition by young children and older adults," *J. Acoust. Soc. Am.* **87**, 2705–2715.
- Picheny, M. A., Durlach, N. I., and Braidia, L. D. (1985). "Speaking clearly for the hard of hearing: I. Intelligibility differences between clear and conversational speech," *J. Speech Hear. Res.* **28**, 96–103.
- Picheny, M. A., Durlach, N. I., and Braidia, L. D. (1986). "Speaking clearly for the hard of hearing: II. Acoustic characteristics of clear and conversational speech," *J. Speech Hear. Res.* **29**, 434–446.
- Picheny, M. A., Durlach, N. I., and Braidia, L. D. (1989). "Speaking clearly for the hard of hearing: III. An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech," *J. Speech Hear. Res.* **32**, 600–603.
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.* **97**, 593–608.
- Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-Effects Models in S and S-Plus* (Springer, New York).
- Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005). "Lexical frequency and acoustic reduction in spoken Dutch," *J. Acoust. Soc. Am.* **118**, 2561–2569.
- Quené, H. (2008). "Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo," *J. Acoust. Soc. Am.* **123**, 1104–1113.
- Quené, H., and van den Bergh, H. (2004). "On multi-level modeling of data from repeated measures designs: A tutorial," *Speech Commun.* **43**, 103–121.
- Quené, H., and van den Bergh, H. (2008). "Examples of mixed-effects modeling with crossed random effects and with binomial data," *J. Mem. Lang.* **59**, 413–425.
- Rosen, S. (2003). "Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything?" *J. Phonetics* **31**, 509–527.
- Salthouse, T. A. (2000). "Aging and measures of processing speed," *Biol. Psychol.* **54**, 35–54.
- Schneider, B. A., Daneman, M., and Murphy, D. R. (2005). "Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing?" *Psychol. Aging* **20**, 261–271.
- Snijders, T. A. B., and Bosker, R. J. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (Sage, London).
- Sommers, M. S. (2005). in *The Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (Blackwell, Oxford), pp. 469–493.
- Sommers, M. S., and Danielson, S. M. (1999). "Inhibitory processes and spoken word recognition in young and elderly adults: The interaction of lexical competition and semantic context," *Psychol. Aging* **14**, 458–472.
- Sommers, M. S., and Humes, L. E. (1993). "Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners," *J. Acoust. Soc. Am.* **93**, 2903–2914.
- Spehar, B., Tye-Murray, N., and Sommers, M. (2004). "Time-compressed visual speech and age: A first report," *Ear Hear.* **25**, 565–572.
- Speranza, F., Daneman, M., and Schneider, B. A. (2000). "How aging affects the reading of words in noisy backgrounds," *Psychol. Aging* **15**, 253–258.
- Stine, E. A. L. (1990). "On-line processing of written text by younger and older adults," *Psychol. Aging* **5**, 68–78.
- Tallal, P. (1980). "Auditory temporal perception, phonics and reading disabilities in children," *Brain Lang.* **9**, 182–198.
- Tun, P. A. (1998). "Fast noisy speech: Age differences in processing rapid speech with background noise," *Psychol. Aging* **13**, 424–434.
- Tun, P. A., Wingfield, A., Stine, E. A. L., and Mecsas, C. (1992). "Rapid speech processing and divided attention: Processing rate vs. processing resources as an explanation of age effects," *Psychol. Aging* **7**, 546–550.
- Turner, C. W., Smith, S. J., Aldridge, P. L., and Stewart, S. L. (1997). "Formant transition duration and speech recognition in normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **101**, 2822–2825.
- Uchanski, R. M., Choi, S. S., Braidia, L. D., Reed, C. M., and Durlach, N. I. (1996). "Speaking clearly for the hard of hearing: IV. Further studies of the role of speaking rate," *J. Speech Hear. Res.* **39**, 494–509.
- Vaughan, N. E., and Letowski, T. (1997). "Effects of age, speech rate, and type of test on temporal auditory processing," *J. Speech Lang. Hear. Res.* **40**, 1192–1200.
- Wechsler Adult Intelligence Test* (2004). 3rd ed., Dutch version (Harcourt Test, Amsterdam).
- Willott, J. F. (1991). *Aging and the Auditory System: Anatomy, Physiology and Psychophysics* (Thompson, San Diego).
- Wingfield, A., Tun, P. A., Koh, C. K., and Rosen, M. J. (1999). "Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech," *Psychol. Aging* **14**, 380–389.
- Witton, C., Stein, J. F., Stoodley, C. J., Rosner, B. S., and Talcott, J. B. (2002). "Separate influences of acoustic AM and FM sensitivity on the phonological decoding skills of impaired and normal readers," *J. Cogn. Neurosci.* **14**, 866–874.
- Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., Stein, J. F., and Green, G. G. R. (1998). "Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers," *Curr. Biol.* **8**, 791–797.
- Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., and Merzenich, M. M. (1997). "Deficits in auditory temporal and spectral resolution in language-impaired children," *Nature (London)* **387**, 176–178.
- Zekveld, A. A., George, E. L. J., Kramer, S. E., Goverts, S. T., and Houtgast, T. (2007). "The development of the text reception threshold test: A visual analogue of the speech reception threshold test," *J. Speech Lang. Hear. Res.* **50**, 576–584.