Effects of stimulus variability and adult aging on adaptation to time-compressed speech

Julie D. Golomb,[a] Jonathan E. Peelle, and Arthur Wingfield[b]
Volen National Center for Complex Systems, Brandeis University, Waltham, Massachusetts 02454

(Received 4 February 2006; revised 13 September 2006; accepted 1 January 2007)

With as few as 10–20 sentences of exposure, listeners are able to adapt to speech that is highly distorted compared to that which is encountered in everyday conversation. The current study examines the extent to which adaptation to time-compressed speech can be impeded by disrupting the continuity of the exposure sentences, and whether this differs between young and older adult listeners when they are equated for starting accuracy. In separate sessions conducted one week apart, the degree of adaptation was assessed in four exposure conditions, all of which involved exposure to the same number of time-compressed sentences. A continuous exposure condition involved presentation of the time-compressed sentences without interruption. Two alternation conditions alternated time-compressed speech and uncompressed speech by single sentences or groups of four sentences. A fourth condition presented sentences that were separated by a period of silence but no uncompressed speech. For all conditions, neither young nor older adults' overall level of learning was influenced by disruptions to the exposure sentences. In addition, participants' performance showed reliable improvement across the first and subsequent sessions. These results support robust learning mechanisms in speech perception that remain functional throughout the lifespan.
 [DOI: 10.1121/1.2436635]

PACS number(s): 43.71.Es, 43.71.Lz, 43.70.Dn [MSS] Pages: 1701–1708

I. INTRODUCTION

Speaking rates vary considerably in the course of everyday listening. Although speech rates typically range from 140 to 180 words per minute (wpm), speakers reading from a prepared script often exceed these rates (Stine *et al.*, 1990). Even a single talker will vary his or her speaking rate considerably within a given conversation (Miller, Grosjean, and Lomanto, 1984). These variations in speech rate influence, among other things, listeners' perception of phonemic boundaries (Miller, Aibel, and Green, 1984; Miller and Liberman, 1979), which play a key role in speech recognition. To effectively process speech, listeners must therefore be able to adjust to changes in speech rate extremely rapidly. When variations in speech rate are minimal, as is typically the case in natural speech, listeners can accomplish this task without apparent effort. However, when speech rates are extremely fast, adaptation becomes considerably more difficult.

Despite the challenge posed by rapid speech rates, listeners are able to adapt to speech time compressed to rates that are substantially faster than that to which they are accustomed. Studies consistently show that exposure to 10–20 sentences of time-compressed speech is sufficient to significantly improve listeners' comprehension of such materials (Dupoux and Green, 1997; Pallier *et al.*, 1998; Sebastián-Gallés *et al.*, 2000). This perceptual learning is a robust phenomenon that is largely preserved in healthy older adults (Peelle and Wingfield, 2005) in spite of other perceptual (Schneider, 1997) and cognitive (Wingfield and Stine-Morrow, 2000) changes.

Perceptual learning is often characterized as the shifting of attention toward cues that are relevant for a task and away from those which are irrelevant (Francis *et al.*, 2000; Goldstone, 1998; Nosofsky, 1986). For example, Francis *et al.* (2000) used category-level feedback to train listeners to distinguish speech cues that varied across two dimensions: (1) the frequency spectrum occurring during burst onset, and (2) the origin of the formant transitions. Different groups of listeners were provided feedback that emphasized the frequency or burst components of the stimuli. Listeners' judgment scores reflected this feedback, with decisions about stimuli with conflicting burst and formant information following the dimension on which they were trained. In addition, listeners made less use of cues present in the dimension on which they were not trained. These results support a mechanism of perceptual learning of speech sounds that involves a reallocation of attention during the learning process.

Given the above-mentioned view of perceptual learning, it is reasonable to think that adaptation to time-compressed speech occurs as listeners recalibrate phonemic boundaries to accommodate the much more rapid speech rate. An outstanding question is whether the redirection of attentional focus necessary for this adjustment can be disrupted if listeners hear speech at two speech rates that are far removed from each other. If increasing attention to cues salient at one speech rate comes at a cost to attention to cues at a second rate, alternating between the two rates over the course of an adaptation period should produce shifts in attention that hinder learning.

[a]Current affiliation: Interdepartmental Neuroscience Program, Yale University, New Haven, Connecticut 06520.
[b]Electronic mail: wingfield@brandeis.edu

An alternate line of research has examined stimulus variability as a means of increasing the overall learning and generalizability of phonetic training (e.g., Barcroft and Sommers, 2005; Lively *et al.*, 1993). Logan *et al.* (1991) trained Japanese listeners on identification of /r/ and /l/, a discrimination that previous training paradigms had little success in teaching. The authors effected a significant improvement in identification scores by employing training words that varied in placement and context of the consonant sounds (word-initial and final positions, singleton and cluster environments, and in intervocalic positions), spoken by five talkers. Follow-up studies using a single talker for training found that listeners improved in the /r/-/l/ distinction for that talker, but that this learning did not generalize to novel talkers (Lively *et al.*, 1993). These results indicate that, at least in some contexts, variability in training stimuli may actually be beneficial to listeners' perceptual learning. This presumably occurs because, rather than merely forming associations with a single idiosyncratic token of a phoneme, with sufficient variability listeners are able to define a perceptual space that includes both a prototype and several variants (e.g., Kuhl, 1991).

Dupoux and Green (1997) asked whether adaptation to time-compressed speech might be affected by interrupting a series of time-compressed sentences with presentation of sentences heard at a different speech rate. The authors presented listeners with 10 sentences time-compressed to 38% of their original duration followed by 5 uncompressed sentences, and a final 5 sentences again compressed to 38% of their original duration. Perceptual learning was assessed by asking listeners, following each sentence, to report as many words as possible from that sentence. Listeners' recall accuracy improved over the course of the time-compressed sentences, but showed a moderate drop for the two time-compressed sentences immediately following the uncompressed speech. However, listeners' accuracy quickly regained (and surpassed) the level achieved at the end of the initial 10-sentence set of compressed sentences. Dupoux and Green (1997) concluded that there was a short-term, local adjustment to the uncompressed speech that caused the drop in performance, but that perceptual learning persisted through a longer-term component, as evidenced by the rapid recovery.

The initial drop in accuracy following a change in speech rate reported by Dupoux and Green (1997) suggests that listeners' attentional focus may have indeed shifted toward speech cues salient at the slower speech rate, hindering their performance on the time-compressed sentences. It is possible that, had there been more interruptions between the exposure sentences, a summation of the small performance drops associated with each interruption would have resulted in a significant decline in overall learning. This would be consistent with attentional processes that tune in to the current or local speech rate at the expense of other speech rates. Alternatively, it may be the case that such narrow local attention is not necessary to adapt to speech at a new rate. The primary goal of the current study was to investigate whether frequent shifts in speech rate might have a significant impact on perceptual learning.

A second goal of this experiment was to examine the effect of shifts in speech rate on perceptual learning in older adults. There are many reasons one would expect age to adversely affect adaptation to rapid speech. First, older adults typically have poorer hearing acuity than do young adults (Morrell *et al.*, 1996), with this decline in auditory acuity contributing to poorer performance on a variety of language tasks (Sommers, 1997; Wingfield *et al.*, 2006). Accompanying these changes in peripheral hearing are age-related declines in auditory temporal processing ability. In gap detection tasks, older adults typically require longer gaps between sounds than do young adults to correctly discriminate one from two tones (Schneider *et al.*, 1994; Strouse *et al.*, 1998; Schneider and Hamstra, 1999). Fitzgibbons and Gordon-Salant (1995) investigated listeners' temporal resolution ability by comparing discrimination thresholds for changes in the duration of tones embedded in sequences to those for the same tones presented in isolation. They found that older adults' discrimination ability was significantly worse for tones embedded in sequences compared to that of the young adults. These findings have been upheld in studies using modified natural speech sounds (Gordon-Salant *et al.*, 2006).

Above and beyond these age-related changes in sensory processing, older adults exhibit declines in several cognitive domains important for language processing. These include declines in attentional resources (McDowd and Shaw, 2000), slower speed of information processing (Salthouse, 1996), a diminished working memory capacity (Zacks *et al.*, 2000), and difficulty inhibiting irrelevant or distracting information (Tun *et al.*, 2002). The result is that older adults perform more poorly than their younger counterparts on a wide variety of language tasks, including exhibiting a greater difficulty processing time-compressed speech (Gordon-Salant and Fitzgibbons, 1993; Wingfield *et al.*, 1999, 2003).

In addition to auditory and linguistic processing, adaptation to rapid speech involves learning new acoustic representations of stored phonemic categories, and older adults exhibit declines on a number of tasks involving learning and memory. These include word generation, paired associates, cued recall, and free recall (Craik, 1977; Craik *et al.*, 1987; Kausler, 1994). There is also evidence from the episodic memory literature that supports age-related weakening of associations between items (Kahana *et al.*, 2002; Wingfield and Kahana, 2002). Although the literature on age-related changes in perceptual learning is limited, the available evidence would lead one to expect age-related deficits on this front as well. For example, older adults show significant impairments in prism adaptation (Fernandez-Ruiz *et al.*, 2000) and semantic category visual search tasks (Gilbert and Rogers, 1996; Rogers *et al.*, 1994).

To the degree that adaptation to time-compressed speech depends on the above-listed processes, such adaptation should be adversely affected in adult aging. Contrary to this expectation, it has recently been demonstrated that under ideal training conditions—that is, an uninterrupted series of training sentences in an otherwise quiet environment—older adults initially adapt to time-compressed speech in a manner comparable to young adults (Peelle and Wingfield, 2005). However, this does not rule out age-related declines in perceptual learning. In numerous other language paradigms, older adults can compensate to keep pace with young adults in easier conditions, but their performance falls off as difficulty is increased through perceptual or cognitive manipulations (e.g., Tun, 1998; Wingfield *et al.*, 2003). Older adults' potential difficulty with perceptual learning is also supported by age-related declines in the maintenance and transfer of perceptual learning with time-compressed speech (Peelle and Wingfield, 2005). It is therefore important to examine whether variations in speech rate exert a differential effect on older adults' adaptation.

A better understanding of the dynamics of perceptual learning in older age is vital for a number of reasons. In practical terms, as older adults have more difficulty understanding certain kinds of speech, they stand to benefit from well-designed training programs. To be effective, such programs will need to reflect a deep knowledge about the dynamics of perceptual learning and how easily it can be disrupted. Second, as indicated previously, existing research on perceptual learning in older adults is scant, particularly with respect to speech sounds. This study thus helps to address an important gap in our understanding of cognitive aging, particularly as it relates to language comprehension. Finally, given known age-related declines in cognitive resources, assessing older adults' performance on this task provides additional guidance in identifying cognitive systems that support perceptual learning. That is, age-related differences in task performance are likely attributable to a cognitive system which shows significant age-related change. Conversely, a lack of age differences would suggest that the processes do not rely on such systems.

In the current experiment, young and older adult listeners were presented with time-compressed sentences to assess perceptual learning in four conditions. In the first condition, time-compressed sentences were presented without any intervening uncompressed sentences, comparable to previous studies (Dupoux and Green, 1997; Pallier *et al.*, 1998; Sebastián-Gallés *et al.*, 2000; Peelle and Wingfield, 2005). In addition, adaptation was tested in three alternation conditions. Two of these conditions involved alternating compressed sentences with uncompressed sentences. Equal numbers of compressed and uncompressed sentences were used, but switches between the two rates occurred with different frequencies. In the final alternation condition, compressed sentences were alternated with a silent period to determine whether disruptions in exposure continuity, without experiencing sentences at a different speech rate, would influence learning.

II. METHOD

A. Participants

All participants were native English speakers and reported themselves to be in good health. The young participants (8 men and 8 women) consisted of university undergraduates aged 18–22 years ($M=19.6$, s.d.$=1.8$). They had a mean of 14.0 years of education at the time of testing (s.d.$=1.3$) and a mean WAIS-III vocabulary score (Wechsler, 1997) of 47.2 (s.d.$=5.2$). The older participants (6 men and 10 women) were healthy, community-dwelling adults aged 65–85 years ($M=72.4$, s.d.$=5.0$). They had a mean of 15.8 years of education (s.d.$=1.4$) and a mean WAIS-III vocabulary score of 49.1 (s.d$=7.2$). Both groups were thus well-educated, with the older group having an average of 1.8 more years of formal education [$t(30)=3.74, p<0.001$]. There was no difference in vocabulary score between the young and the older adults [$t(30)=0.84, n.s.$]. Participants were compensated with a small monetary sum for their involvement.

Audiometric testing was conducted to ensure that all participants had clinically normal hearing within the speech range. The mean pure-tone average [(PTA)—taken as the mean threshold for tones at 500, 1000, and 2000 Hz] in the better ear was 3.5 dB HL (s.d.$=3.6$) for the young adults and 13.7 dB (s.d.$=8.2$) for the older adults. Mean speech reception thresholds [(SRTs)—the intensity at which two-syllable words can be correctly identified 50% of the time] for the better ear were 4.4 dB HL (s.d.$=2.5$) for the young adults and 12.8 dB (s.d.$=8.4$) for the older adults. Although age differences in both PTA [$t(30)=4.54, p<0.001$] and SRT [$t(30)=3.89, p<0.001$] were significant, both groups fell within the range of clinically normal hearing, defined as a PTA of 25 dB HL or less in the speech frequency range (Hall and Mueller, 1997).

B. Stimuli

The stimuli consisted of 136 sentences, each of which contained 10 words (7 content words and 3 function words) and 14–16 syllables (e.g., *Sarah took her dirty work clothes to the neighborhood cleaners*). Sentences were recorded by a female native speaker of American English at a comfortable speech rate of approximately 200 wpm.

Because older adults are differentially affected by rapid speech (Gordon-Salant and Fitzgibbons, 1993; Wingfield *et al.*, 1999,2003), it would have been problematic to use a single speech rate for all participants: Speech that would have been appropriate for young adults would have proven too difficult for older adults, whereas slowing the speech rate to accommodate the older adults would not have posed a sufficient challenge for the young adult listeners. Different speech rates were therefore used for the young and older adults so that they would be approximately equated for starting accuracy level. These speech rates were chosen based on pilot data to equate both participant groups for overall accuracy levels prior to any adaptation.

Sound-editing software (SoundEdit, Macromedia Inc., San Francisco) was used to compress 96 of the stimulus sentences to 30% of their original duration (approximately 680 wpm) for the young adults and 40% of their original duration (approximately 510 wpm) for the older adults. The software uses a type of PSOLA algorithm for time compression, which compresses both speech and silence equally and preserves information from short sounds such as stop consonants and formant glides through temporal averaging. The remaining 40 sentences were left uncompressed. Sentences were presented binaurally over headphones at a comfortable

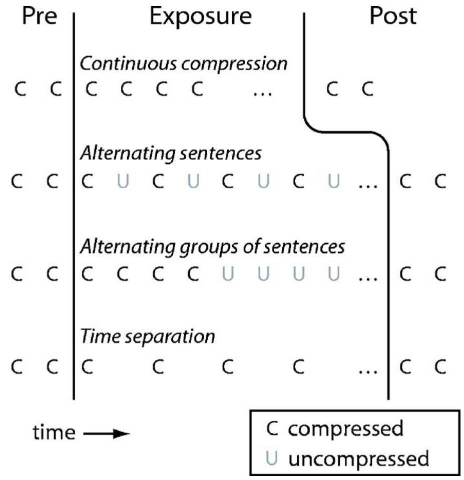

FIG. 1. Schematic illustration of the four experimental conditions employed. In all four conditions, two time-compressed sentences were presented before and after exposure to assess the effects of exposure on sentence intelligibility. All four conditions contained 20 time-compressed sentences during the exposure portion, with no interruption (*continuous compression*), alternating with uncompressed speech in single sentences (*alternating sentences*) or groups of four (*alternating groups of sentences*), or with a silent pause between compressed sentences (*time separation*).

listening level, determined for each participant, which remained unchanged for that participant throughout the experiment.

C. Procedure

The four presentation conditions used in this study are schematically displayed in Fig. 1. In each of the four conditions, the first two sentences (preexposure) and the last two sentences (postexposure) were always presented at the compressed rate. The sentences in between varied in composition according to condition. The *continuous compression* condition consisted of 20 time-compressed exposure sentences presented without any intervening uncompressed sentences and is comparable to previous studies involving perceptual adaptation to time-compressed sentences (Dupoux and Green, 1997; Pallier *et al.*, 1998; Sebastián-Gallés *et al.*, 2000; Peelle and Wingfield, 2005). The *alternating sentences* and *alternating groups of sentences* conditions consisted of presentation of 20 compressed exposure sentences and 20 uncompressed sentences alternating either every other sentence or in groups of four sentences, respectively. A fourth *time separation* condition consisted of 20 compressed exposure sentences presented with a 5 s pause inserted between them to approximate the spacing of the *alternating sentences* condition without presenting any uncompressed speech.

Participants were individually tested in four sessions. Each session involved one of the four above-described conditions, with the order of conditions counterbalanced across participants. The particular sentences used in each condition were also counterbalanced across participants, such that across all participants each sentence occurred equally often in each condition. Sessions were separated by 7–8 days to minimize possible carryover in learning from previous sessions.

Participants were instructed to recall verbatim as much of each sentence as possible immediately following its presentation. They were encouraged to guess if they were unsure, and they were informed that there would be no penalty for words recalled incorrectly or in the wrong order. No feedback was given regarding the correctness of responses. Participants pressed a key to indicate when they were finished recalling; a second keypress initiated presentation of the next sentence. Participants' spoken responses were recorded onto cassette tape for later analysis. The first session included a single time-compressed practice sentence to orient participants to the task; subsequent sessions included no practice.

D. Scoring

Sentence recall was scored as the number of correct content words out of seven possible for each sentence. Words with added or eliminated suffixes (e.g., -s, -ed, -ing) were counted as correct, as were verbs given in the wrong tense. If a participant recalled half of a compound word (e.g., "school" instead of "schoolwork") it was marked as half credit. Portions of noncompound words (e.g., "let" instead of "letters") were not given any credit. These scoring criteria are consistent with previous studies (Dupoux and Green, 1997; Pallier *et al.*, 1998; Sebastián-Gallés *et al.*, 2000; Peelle and Wingfield, 2005).

III. RESULTS

A. Effects of exposure condition

All analyses focused on listeners' comprehension of time-compressed sentences. Performance on any intervening uncompressed sentences, which occurred in the two alternating speech conditions, was near perfect and was not analyzed further. For each condition, listeners' performance on preexposure and postexposure sentences was compared to evaluate the magnitude of adaptation.

Recall accuracy for participants' pre- and postexposure sentences in all conditions is displayed in Fig. 2. As intended, by using a somewhat slower speech rate for the older adults than for the young adults, the preexposure levels were approximately equal across condition and age group. These preexposure data were submitted to a 4 (Exposure condition: continuous compression, alternating sentences, alternating groups of sentences, time separation) $\times$ 2 (Age: young, older) mixed design ANOVA, with exposure condition as a within-participants variable. The main effect of exposure condition was not significant $[F(3,90)<1]$, as was expected, given that for each session the preexposure sentences occurred before any experimental manipulations. Although older adults showed a slight advantage in starting accuracy levels due to the specific speech rates that were used, there was no significant effect of age $[F(1,30)<1]$, or an age by condition interaction $[F(3,90)<1]$. Thus, the speech rates were appropriately chosen to equate the two groups on starting accuracy.

Visual inspection of Fig. 2 shows that exposure to time-compressed speech improved listeners' accuracy scores in all conditions. To confirm this conclusion, the data were submitted to a 2 (Exposure: preexposure, postexposure) $\times$ 4 (Ex-

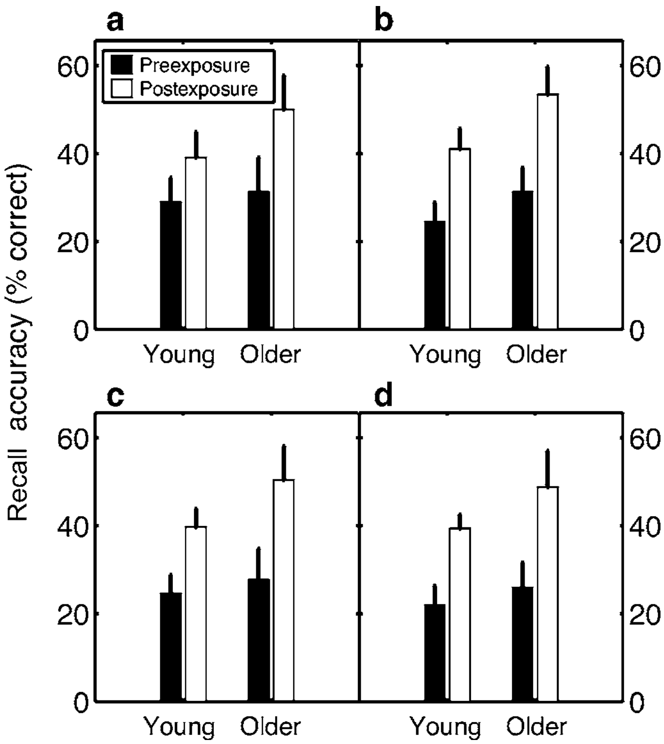

FIG. 2. Response accuracy for young and older listeners before (black bars) and after (white bars) exposure to time-compressed sentences as a function of exposure condition. (**a**) Continuous compression condition (no interruption); (**b**) alternating sentences condition (alternating single compressed and uncompressed sentences); (**c**) alternating groups of sentences condition (alternating compressed and uncompressed sentences in groups of 4); and (**d**) time separation condition (5 s between sentences). Error bars represent 1 standard error.

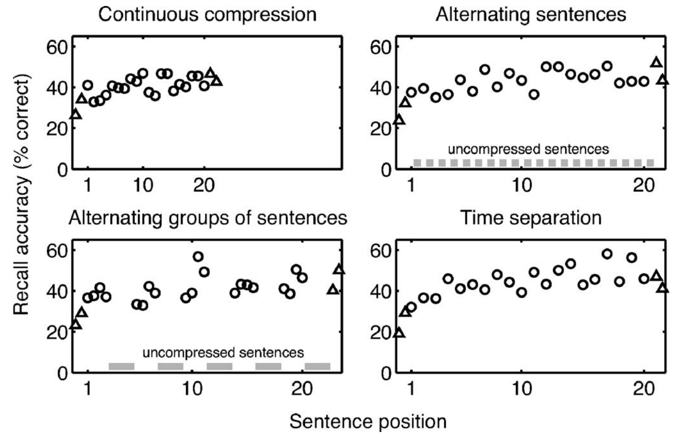

FIG. 3. Response accuracy for time-compressed sentences during each of four exposure conditions (circles), collapsed across young and older adults. The gray bars along the abscissa indicate when uncompressed sentences occurred for the alternating sentences and alternating groups of sentences conditions. The preexposure and postexposure sentences are replotted here for reference (triangles). Sentence position refers to the location of each time-compressed sentence within the exposure sentences.

posure condition: continuous compression, alternating sentences, alternating groups of sentences, time separation) $\times$ 2 (Age: young, older) ANOVA. There was a main effect of exposure $[F(1,30)=53.87, p<0.001]$, confirming that listeners' accuracy improved with exposure to the time-compressed speech. On average, the young group improved from 25.1% correct in the preexposure test to 39.8% correct in the postexposure test, an average improvement of 14.7%. The older group improved from an average of 29.0% correct in the preexposure test to 50.7% correct in the postexposure test, an average improvement of 21.7% correct. There was no effect of exposure condition $[F(3,90)<1]$, supporting the observation that participants' improvement was not affected by stimulus spacing or by speech rate variability. In addition, the older adults' performance did not differ significantly from that of the young adults, producing no main effect of age $[F(1,30)=1.41, n.s.]$. None of the interactions was significant (all F's <2, all p's>0.17).

It is conceivable that, although overall improvement due to exposure to time-compressed speech was equivalent across conditions, learning might still have been affected by the exposure condition manipulation. For example, participants' recall accuracy may have fluctuated more from trial to trial in the alternating conditions relative to the continuous compression condition due to the inclusion of uncompressed sentences, which have been shown to transiently reduce recall accuracy (Dupoux and Green, 1997). To determine whether this was the case, the exposure data were analyzed for each condition trial-by-trial. These exposure data, along with those for pre- and postexposure data, are displayed in Fig. 3 (the gray bars along the abscissa indicate the location of uncompressed sentences or groups of sentences in those conditions). To allow comparison across conditions, each exposure sentence was assigned a position number 1–20 based on the amount of exposure, regardless of any intervening sentences or pauses. Thus, equivalent sentence positions indicate that comparable exposure had occurred.

The exposure data were submitted to a 4 (Exposure Condition: continuous compression, alternating sentences, alternating groups of sentences, time separation) $\times$ 20 (Exposure sentence position: 1–20) ANOVA. As expected, recall accuracy improved with increased exposure, indicated by a main effect of exposure sentence position $[F(19,589)=2.75, p<0.001]$. Visual inspection of Fig. 3 might suggest that the alternating group condition resulted in slightly less steady learning, with brief drops in accuracy every four time-compressed sentences, corresponding to the inserted groups of uncompressed sentences. However, despite these minor fluctuations, there was no consistent effect of exposure condition on learning, evidenced by a lack of a main effect of exposure condition $[F(3,93)=1.43, n.s.]$. The results of the learning data analysis are thus consistent with our analysis of the pre- and postexposure sentences, and do not indicate any significant effect of exposure condition.

B. Effects of repeated exposure

Although sessions were conducted a minimum of 7 days apart in order to minimize carryover of learning from one session to another, the possibility that some carryover occurred could not be ruled out. Because the first set of analyses failed to reveal any effect of exposure condition, it was possible to analyze the data as a function of session, collapsing across exposure condition.

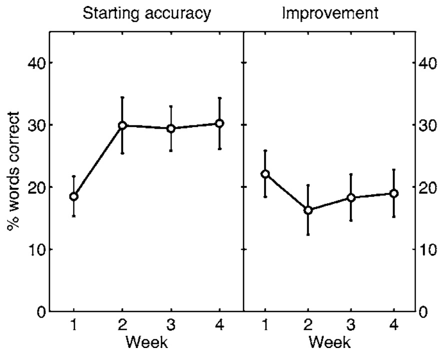

FIG. 4. Starting accuracy (left panel) and amount of improvement (right panel) for the four experimental sessions, collapsed across age and exposure condition. Error bars represent 1 standard error.

Participants' preexposure recall accuracy at the start of each of the four sessions is shown in the left panel of Fig. 4. Because there were no age differences in our analyses of exposure condition, data for young and older listeners have been combined when looking at session-to-session improvement. The lack of any age differences was also verified using uncorrected t-tests (all p's > 0.14).

Visual inspection of the left panel of Fig. 4 suggests that starting accuracy increased between the first session and subsequent sessions, which is consistent with retention of learning from one session to subsequent sessions. An ANOVA confirmed this difference in starting accuracy levels, showing a significant effect of session $[F(3,93)=5.11, p<0.05]$. Post-hoc t-tests confirmed that this effect was attributable to the significant increase in starting accuracy from week 1 to week 2 $[t(31)=3.08, p<0.01]$. There was no significant change between weeks 2–3 or 3–4 (both p's > 0.8).

In addition to examining listeners' starting accuracy levels, the question of whether total amount of improvement changed over time was addressed by analyzing improvement scores (difference between postexposure and preexposure scores) over sessions. These data are plotted in the right panel of Fig. 4. Visual inspection of the data suggests listeners' accuracy improved a comparable amount each week. Consistent with this observation, an ANOVA on these data failed to show a significant effect of session $[F(3,93)<1]$, indicating that the amount of improvement indeed remained stable across sessions. As with the starting accuracy data, t-tests for each week's accuracy levels relative to baseline confirmed that there were no differences in performance between young and older adults (all p's > 0.29).

IV. DISCUSSION

Although there are a number of possible ways to characterize perceptual learning, researchers in the speech literature have generally opted for models involving the allocation of attention to relevant stimulus dimensions (e.g., Francis and Nusbaum, 2000; Logan *et al.*, 1991; Nittrouer and Miller, 1997). In the current study we investigated whether wide fluctuations in speech rate would disrupt this attentional process. We found that interrupting time-compressed training sentences with brief pauses or uncompressed speech had no effect on perceptual learning. Furthermore, there were no differences in performance between young and older adult listeners.

Our results suggest that listeners' attentional allocation to a particular speech rate is not disrupted by the use of a second speech rate, as would be predicted from previous studies of learning with speech sounds (Francis, 2000). This finding is especially surprising given that the current experiment employed a fairly marked contrast, alternating time-compressed speech (510 or 680 wpm) with uncompressed speech (200 wpm). Because listeners take longer to adapt to speech that is further removed from unaltered speech (Dupoux and Green, 1997; Peelle and Wingfield, 2005), these extreme choices of speech rates should present a worst-case scenario for the listener.

This discrepancy might seem to imply that the paradigm of attentional allocation is incorrect, and that adapting to a new speech rate does not in fact require attentional resources. If this were the case, it would not be surprising if shifts in speech rate should fail to hinder learning. An alternative framework within which to view adaptation to time-compressed speech might be closer to that of stimulus imprinting, in which repeated exposure to novel stimuli lead to strengthened internalized traces of that stimulus (e.g., Logan, 1988; Palmeri *et al.*, 1993; Palmeri, 1997). Under such a view, learning would be governed most strongly by the amount of exposure to novel sounds, regardless of intervening sentences. However, this and other alternative frameworks seem to be poor fits for the generalizability shown in speech perception, and are not consistent with previous studies of phoneme discrimination.

A more likely explanation, then, is that natural speech represents a special case in speech processing. Although we used two very different speech rates, only the faster rate required adaptation by the listeners. As the normal speech rate did not require assimilation of any new information, it was not dependent on attentional resources, and thus did not take these resources away from the novel, faster speech rate. Restated more formally, the normal speech rate presumably contains phoneme prototypes (Kuhl, 1991). At the beginning of training the time-compressed speech tokens are too far away from the prototype to be correctly identified, but through training they are brought closer in perceptual space (Iverson and Kuhl, 1995). Within such a framework the prototypes remain relatively stable, and would not be disrupted by the small number of sentences employed in our training regime. Under this interpretation, attention is indeed directed toward relevant dimensions of the novel stimuli, but is not diverted when the uncompressed stimuli are presented. This description fits best with prior studies of phoneme discrimination and the current data.

Successful speech comprehension relies upon a number of more general cognitive processes, including auditory processing, working memory, attention, and executive control (e.g., Grossman *et al.*, 2003; Peelle *et al.*, 2004; Price *et al.*, 2005; Wingfield and Stine-Morrow, 2000). Given the notable age-related declines in the cognitive abilities presumed to be necessary for perceptual learning, the lack of age differences in the current study is particularly striking. Two possible explanations exist for this difference. First, it is possible that portions of the language processing system function in a mostly autonomous and encapsulated manner, and that because of their importance and continuous use are relatively resistant to age-related cognitive declines seen in the larger language-processing network. A second possibility is that, although there are age-related declines in cognitive systems important for perceptual learning, older adults are able to compensate for these declines through increased expertise (Wingfield and Stine-Morrow, 2000) and neural recruitment (Cabeza, 2002; Wingfield and Grossman, 2006). Regardless of the explanation, the current results support the hypothesis that, within the temporal limits of the aging auditory system, the initial stage of adaptation to time-compressed speech is preserved in adult aging (Peelle and Wingfield, 2005).

The design of the current study does not allow us to rule out an alternate interpretation of our results. Specifically, the two alternating conditions contained an additional 20 uncompressed sentences. It is conceivable that the alternation did in fact hinder performance, but that this effect was obscured by a commensurate *increase* in performance due to the additional numbers of stimuli employed in these two conditions. Previous studies have indicated the adaptation to time-compressed speech is not simply due to general task practice (Dupoux and Green, 1997; Peelle and Wingfield, 2005). However, because all of the stimuli were spoken by the same speaker, it is possible that there was adaptation to speaker-specific acoustic cues apart from speech rate aided by increased exposure to uncompressed sentences.

It is notable that both young and older participants' pre-exposure recall accuracy showed a significant increase between the first and second weeks. Although the duration of perceptual adaptation has not yet been fully specified (see Janse, 2003), the current data provide strong evidence that learning persists for at least one week. It is also striking that no significant age differences were observed in week-to-week retention. It has been previously demonstrated that older adults' learning reaches an asymptote after 20 time-compressed sentences, whereas young adults' performance continues to improve (Peelle and Wingfield, 2005). The current results indicate that if training is conducted in multiple sessions, older adults' performance can improve beyond this initial limit and keep pace with that of their younger counterparts. The breaks between weekly sessions may be necessary to combat fatigue or to allow time for memory consolidation.

Finally, these results have practical implications for training regimes in both experimental and real-world situations in which acoustically altered speech stimuli are encountered (e.g., Davis *et al.*, 2005; Rosen *et al.*, 1999; Shannon *et al.*, 1995; Vongphoe and Zeng, 2005). First, the precise timing of exposure sentences is not important for effective learning. Second, interrupting training sentences with uncompressed speech (e.g., conversation with the researcher) will not influence adaptation.

The current study demonstrates that adaptation to time-compressed speech is not hindered by brief pauses between training sentences or the insertion of uncompressed sentences. Most important, the lack of differences in young and older adults' performance points to learning processes that are highly resistant to age-related cognitive decline, or that such declines are effectively countered in older adults by compensatory processes.

ACKNOWLEDGMENTS

The authors acknowledge support from NIH Grant Nos. AG04517 and AG019714 from the National Institute on Aging (A.W.) and fellowship F31 DC006376 from the National Institute on Deafness and other Communicative Disorders (J.P.). We also gratefully acknowledge support from the W. M. Keck Foundation.

Barcroft, J., and Sommers, M. S. (**2005**). "Effects of acoustic variability on second language vocabulary learning," Stud. Second Lang. Acquis. **27**, 387–414.

Cabeza, R., (**2002**). "Hemispheric asymmetry reduction in older adults: The HAROLD model," Psychol. Aging **17**, 85–100.

Craik, F. I. M. (**1977**). "Age differences in human memory," in *Handbook of the Psychology of Aging*, edited by J. E. Birren and K. W. Schaie (Von Nostrand Reinhold, New York), pp. 384–420.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., and Anderson, N. D. (**1987**). "Patterns of memory loss in three elderly samples," Psychol. Aging **2**, 79–86.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., and McGettigan, C. (**2005**). "Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences," J. Exp. Psychol. Gen. **134**, 222–241.

Dupoux, E., and Green, K. (**1997**). "Perceptual adjustment to highly compressed speech: Effects of talker and rate changes," J. Exp. Psychol. Hum. Percept. Perform. **23**, 914–927.

Fernandez-Ruiz, J., Hall, C., Vergara, P., and Diaz, R. (**2000**). "Prism adaptation in normal aging: Slower adaptation rate and larger aftereffect," Cognit. Brain Res. **9**, 223–226.

Fitzgibbons, P. J., and Gordon-Salant, S. (**1995**). "Age effects on duration discrimination with simple and complex stimuli," J. Acoust. Soc. Am. **98**, 3140–3145.

Francis, A. L., Baldwin, K., and Nusbaum, H. C. (**2000**). "Effects of training on attention to acoustic cues," Percept. Psychophys. **62**, 1668–1680.

Gilbert, D. K., and Rogers, W. A. (**1996**). "Age-related differences in perceptual learning," Hum. Factors **38**, 417–424.

Goldstone (**1998**). "Perceptual Learning," Annu. Rev. Psychol. **49**, 585-612.

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

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Barrett, J. (**2006**). "Age-related differences in identification and discrimination of temporal cues in speech segments," J. Acoust. Soc. Am. **119**, 2455–2466.

Grossman, M., Cooke, A., DeVita, C., Lee, C., Alsop, D., Detre, J., Gee, J., Chen, W., Stern, M. B., and Hurtig, H. I. (**2003**). "Grammatical and resource components of sentence processing in Parkinson's disease: An fMRI study," Neurology **60**, 775–781.

Hall, J., and Mueller, G. (**1997**). *Audiologist Desk Reference* (Singular, San Diego).

Iverson, P., and Kuhl, P. K. (**1995**). "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling," J. Acoust. Soc. Am. **97**, 553–562.

Janse, E. (**2003**). *Production and Perception of Fast Speech* (Landelijke Onderzoekschool Taalwetenschap, Utrecht, The Netherlands).

Kahana, M. J., Howard, M., Zaromb, F., and Wingfield, A. (**2002**). "Age dissociates recency and lag-recency effects in free recall," J. Exp. Psychol. Learn. Mem. Cogn. **28**, 530–540.

Kausler, D. H. (**1994**). *Learning and Memory in Normal Aging* (Academic, San Diego).

Kuhl, P. (**1991**). "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not," Percept. Psychophys. **50**, 93–107.

Lively, S. E., Logan, J. S., and Pisoni, D. B. (**1993**). "Training Japanese listeners to identify English /r/ and /l/. II. The role of phonetic environment and talker variability in learning new perceptual categories." J.

Acoust. Soc. Am. **94**, 1242–1255.

Logan, G. D. (**1988**). "Toward an instance theory of automatization," Psychol. Rev. **95**, 492–527.

Logan, J. S., Lively, S. E., and Pisoni, D. B. (**1991**). "Training Japanese listeners to identify English /r/ and /l/. A first report," J. Acoust. Soc. Am. **89**, 874–886.

McDowd, J. M., and Shaw, R. J. (**2000**). "Attention and aging: A functional perspective," in *The Handbook of Aging and Cognition*, 2nd ed, edited by F. I. M. Craik and T. A. Salthouse (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 221–292.

Miller and Liberman (**1979**). "Some effects of later-occurring information on the perception of stop consonant and semivowel," Percept. Psychophys. **25**, 457-465

Miller, J. L., Aibel, I. L., and Green, K. (**1984**). "On the nature of rate-dependent processing during phonetic perception," Percept. Psychophys. **35**, 5–15.

Miller, J. L., Grosjean, F., and Lomanto, C. (**1984**). "Articulation rate and its variability in spontaneous speech: A reanalysis and some implications," Phonetica **41**, 215–225.

Morrell, C. H., Gordon-Salant, S., Pearson, J. D., Brant, L. J., and Fozard, J. L. (**1996**). "Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level," J. Acoust. Soc. Am. **100**, 1949–1967.

Nittrouer, S., and Miller, M. E. (**1997**). "Predicting developmental shifts in perceptual weighting schemes," J. Acoust. Soc. Am. **101**, 2253–2266.

Nosofsky, R. M. (**1986**). "Attention, similarity, and the identification-categorization relationship," J. Exp. Psychol. Gen. **115**, 39–57.

Pallier, C., Sebastián-Gallés, N., Dupoux, E., Christophe, A., and Mehler, J. (**1998**). "Perceptual adjustment to time-compressed speech: A cross-linguistic study," Mem. Cognit. **26**, 844–851.

Palmeri, T. J. (**1997**). "Exemplar similarity and the development of automaticity," J. Exp. Psychol. Learn. Mem. Cogn. **23**, 324–354.

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (**1993**). "Episodic encoding of voice attributes and recognition memory for spoken words," J. Exp. Psychol. Learn. Mem. Cogn. **19**, 309–328.

Peelle, J. E., McMillan, C., Moore, P., Grossman, M., and Wingfield, A. (**2004**). "Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: Evidence from fMRI," Brain Lang **91**, 315–325.

Peelle, J. E., and Wingfield, A. (**2005**). "Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech," J. Exp. Psychol. Hum. Percept. Perform. **31**, 1315–1330.

Price, C., Thierry, G., and Griffiths, T. (**2005**). "Speech-specific auditory processing: Where is it?," Trends Cogn. Sci. **9**, 271–276.

Rogers, W. A., Fisk, A. D., and Hertzog, C. (**1994**). "Do ability-performance relationships differentiate age and practice effects in visual search?," J. Exp. Psychol. Learn. Mem. Cogn. **20**, 710–738.

Rosen, S., Faulkner, A., and Wilkinson, L. (**1999**). "Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants," J. Acoust. Soc. Am. **106**, 3629–3636.

Salthouse, T. A. (**1996**). "The processing-speed theory of adult age differences in cognition," Psychol. Rev. **103**, 403–428.

Schneider, B. A. (**1997**). "Psychoacoustics and aging: Implications for everyday listening," Journal of Speech-Language Pathology and Audiology **21**, 111–124.

Schneider, B. A., and Hamstra, S. J. (**1999**). "Gap detection thresholds as a function of tonal duration for younger and older listeners," J. Acoust. Soc. Am. **106**, 371–380.

Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., and Lamb, M. (**1994**). "Gap detection and the precedence effect in young and old adults," J. Acoust. Soc. Am. **95**, 980–991.

Sebastián-Gallés, N., Dupoux, E., Costa, A., and Mehler, J. (**2000**). "Adaptation to time-compressed speech: Phonological determinants," Percept. Psychophys. **62**, 834–842.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.

Sommers, M. S. (**1997**). "Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment," J. Acoust. Soc. Am. **101**, 2278–2288.

Stine, E. L., Wingfield, A., and Myers, S. D. (**1990**). "Age differences in processing information from television news: The effects of bisensory augmentation," J. Gerontol. **45**, 1–8.

Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (**1998**). "Temporal processing in the aging auditory system," J. Acoust. Soc. Am. **104**, 2385–2399.

Tun, P. A. (**1998**). "Fast noisy speech: Age differences in processing rapid speech with background noise," Psychol. Aging **13**, 424–434.

Tun, P. A., O'Kane, G., and Wingfield, A. (**2002**). "Distraction by competing speech in young and older adult listeners," Psychol. Aging **17**, 453–467.

Vongphoe, M., and Zeng, F.-G. (**2005**). "Speaker recognition with temporal cues in acoustic and electric hearing," J. Acoust. Soc. Am. **118**, 1055–1061.

Wechsler, D. (**1997**). *Wechsler Adult Intelligence Scale* (Psychological Corporation, New York).

Wingfield, A., and Grossman, M. (**2006**). "Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging," J. Neurophys. **96**, 2830-2839.

Wingfield, A., and Kahana, M. J. (**2002**). "The dynamics of memory retrieval in older adulthood," Can. J. Exp. Psychol. **56**, 187–199.

Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, L. C. (**2006**). "Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity," J. Am. Acad. Audiol **17**, 487–497.

Wingfield, A., Peelle, J. E., and Grossman, M. (**2003**). "Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults," Aging, Neuropsychology, and Cognition **10**, 310–322.

Wingfield, A., and Stine-Morrow, E. A. L. (**2000**). "Language and speech," in *The Handbook of Aging and Cognition*, 2nd ed., edited by F. I. M. Craik and T. A. Salthouse (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 359–416.

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.

Zacks, R. T., Hasher, L., and Li, K. Z. H. (**2000**). "Human memory," in *The Handbook of Aging and Cognition*, edited by F. I. M. Craik and T. A. Salthouse (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 293–357.