

Title	Exploring auditory aging can exclusively explain Japanese adults age-related decrease in training effects of American English /r/-/l/
Author(s)	Kubo, Rieko; Akagi, Masato
Citation	Proceedings of Meetings on Acoustics, 19: 060078
Issue Date	2013-06-02
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/12327
Rights	Copyright (C) 2013 Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America. The following article appeared in Proceedings of Meetings on Acoustics, 19, 2013, 060078 and may be found at http://dx.doi.org/10.1121/1.4800151
Description	



Proceedings of Meetings on Acoustics

Volume 19, 2013

<http://acousticalsociety.org/>**ICA 2013 Montreal****Montreal, Canada****2 - 7 June 2013****Speech Communication****Session 2aSC: Linking Perception and Production (Poster Session)**

2aSC34. Exploring auditory aging can exclusively explain Japanese adults' age-related decrease in training effects of American English /r/-/l/

Rieko Kubo* and Masato Akagi***Corresponding author's address: JAIST, Nomi, 9231292, Ishikawa, Japan, rkubo@jaist.ac.jp**

Age-related decrease in training effect was shown by training of American English /r/-/l/ contrasts on Japanese speakers. This study examined whether the decrease can be explained exclusively by auditory aging, or other, compensatory cognitive processing should be taken into account. Japanese speakers aged 60's participated the experiment. Hearing threshold and spoken word perception test of participants' first language were used to estimate their auditory aging. The word perception test was composed of low-familiar, high-familiar words and mono syllables. The audiograms showed low threshold at high frequencies. The result of the perception test showed that low intelligibility for phonemes with high frequency or short duration, and confusion between contracted sounds and basic sounds. These were particular for low-familiar words and mono syllables. These results suggest that participants had auditory aging emerging as high frequency loss and time-/frequency-resolution degradation. Nonetheless, the acoustic features to distinguish /r/ and /l/ have long duration, low frequencies and wide frequency distance which are supposed to be unaffected by these auditory aging. The effect of word familiarity suggested compensatory cognitive processing involved. These suggest that age-related decrease cannot be explained exclusively by auditory aging, compensatory cognitive processing should be taken into account.

Published by the Acoustical Society of America through the American Institute of Physics

INTRODUCTION

In speech communication, listeners perceive the speech signals that speakers produced. The speech communication sometimes fails due to mismatches between the underlying phonological systems of the speaker and listener. For example, Japanese speakers have difficulties identifying American English /r/ and /l/. Generally speaking, adults have more difficulties to learn second-language (L2) phonological system than younger generation.

Age-related inhibitory effects to acquire L2 phonological system have been discussed. It has been pointed out that first language (L1) phonological system which acquired in early development involves in [1][2]. That is, once the L1 phonological system was established, it is so robust that it would resist alteration or re-expansion to include new sounds from the L2.

Since long time had passed after the acquisition of L1, adults, including young, middle-aged and older, supposed to have the similar degree of difficulty in learning L2 phonological system. However, perceptual training on adults ranged from 20s to 60s [3][4] showed that adults improved identifying new L2 phonemes and that the amount of training effects were not constant across the age groups. The training effect declined gradually with age, while their L1 phonological system supposed to be established well. It is also shown that older adults' success learning American English /r/-/l/ contrasts varied across the phonetic environment in the word, i.e., they improved the identification accuracy of final-position stimuli (i.e., poor-pool), but not initial cluster (i.e., bright-blight). The relative difficulties of the phonetic environments remained constant across age groups, however, the proportion changed with age so that 20s improved identification accuracy of all phonetic environment almost equally. The learning difficulties shown by older adults corresponded to identifying difficulties by Japanese speakers. It suggested that adults became more influenced by language-specific perception with age, not only in perception but also in learning.

These findings suggested that the appropriate training is able to modify learners' phonological system even after L1 acquisition, L1 phonological system involved learning of L2 phonemes as has been pointed out, and that adults' learning of L2 phonemes was inhibited by L1 phonological system more strongly with age even after learners supposed to have almost equal established L1 phonological system.

Our research aims to account for the age-related inhibitory effects on L2 phonological learning.

Firstly, we hypothesized that age-related inhibitory factor other than L1 acquisition is involved in the success acquiring L2 phonological system. Then we also hypothesized that L1 phonological system is involved in the processing of speech sounds in learning and inhibits learning as has been pointed out, and that the degree of the inhibition caused by L1 phonological system changes with age. Although changes in upper level in human information processing are assumed in these hypotheses, several age-related changes in speech processing could emerge in speech processing process. As a first step, we examined whether age-related changes in lower level interact with the experimental result of adults' training [4] or not, to support our hypotheses.

The most well-known characteristic of age-related changes in speech processing is auditory aging, i.e., threshold elevation for high-frequency sounds, declines in temporal- and frequency-resolutions. These auditory aging may affect the perception of the acoustic features of /r/ - /l/ and resulted in the age-related decrease in the training effect. But also it is well known that auditory speech processing is a complex process shared between bottom-up sensory mechanisms and top-down cognitive control. Listeners compensate for insufficient acoustic features by their k knowledge (e.g., lexical knowledge) to represent the phonemes and understand the speech.

Japanese speakers do not pay attention to F3 differences that are critical to the /r/-/l/ distinction. The listeners supposed to perceive the physical attributes such as frequency, however, they do not have appropriate acoustic features to correctly represent the phonemes. In this instance, they need to compensate by their acquired knowledge. Acquired L1 phonological system could be used as acquired knowledge. The findings of age-related changes in compensation for insufficient acoustic features by listeners' acquired knowledge might lead to a presumption that compensation based on L1 phonological system changes with age. In other words, age-related changes in upper level of speech processing of unfamiliar sounds could be presumed to result in the age-related decrease in the training effects.

This paper investigates mainly 1) auditory aging fully account for aging effect on L2 learning or other, changes in upper processing should be taken into account, and also investigated 2) age-related changes in compensation emerge when insufficient acoustic features, or not.

Auditory aging, i.e., deficits in hearing threshold, temporal- and frequency-resolutions were investigated through a set of experiments consisted of measuring hearing threshold and speech perception test. And acoustic analyses of stimuli of speech perception test and /r/-/l/ were conducted to investigate whether adults with auditory aging could perceive the physical acoustic features of /r/-/l/ or not. Three groups, young, middle-aged and older adults were

tested to investigate that adults' hearing ability decrease consistently with age as the training effects decreased consistently with age, and whether auditory aging could fully explain the decrease the training effects or not. Two word-familiarity ranks and monosyllabic words were used to investigate the compensation by acquired knowledge.

METHODS

Participants

Three groups of native speakers of Japanese were participated (TABLE 1). Young and middle-aged adults were recruited from JAIST whereas older adults were recruited from the local community. None of them reported any hi any speaking, hearing, or reading problems.

TABLE 1. Participants

Age group	Age (average) [year]	N (Male, Female)
Young adults	22 - 28 (23.9)	9 (9, 0)
Middle-aged adults	34 - 49 (38.6)	8 (3, 5)
Older adults	62 - 67 (64.0)	7 (4, 3)

Stimuli

Speech Perception Test

Words were selected from the familiarity-controlled word lists 2003 (FW03), which consists of Japanese words. Low-familiarity words, high-familiarity words and monosyllabic words were selected from word-familiarity rank of 1.0 - 2.5, 5.5 - 7.0 and monosyllable respectively. Each word from low-familiarity and high-familiarity consisted of four syllables, and each syllable was V or CV (V: Vowel, C: Consonant). Each monosyllabic word consisted of one syllable, V or CV. Two hundreds low-familiarity words, 200 high-familiarity words and monosyllabic words were selected for the test session. The responses for 127 low-familiarity words, 133 high-familiarity words and 100 monosyllabic words were analyzed since all participants completed. Finally, 5 vowels and 26 consonants were included in the analyzed responses. Although words were selected so that the number of each syllable was balanced as well as possible, it varied from 1 to 63 (average: 12.0). Three low-familiarity words and 2 high-familiarity words were selected for the practice session. The utterances by a female speaker (fto) for the test session and a male speaker (mya) for the practice session were normalized for peak intensity and presented to the participants.

Acoustic Analyses

1) Misidentified sounds and confused sounds and 2) identified sound were selected from the result of the speech perception test to investigate what extent adults had hearing deficits, and 3) Non-words contrasting English /r/ and /l/ were also selected to be compared with 1) and 2) to investigate the relative position of these three types of stimuli as for duration and frequency of spectral cue.

- 1) **Misidentified sounds and confused sounds** Confusions shared by a number of individuals were selected. For each misidentified sound, a word which consisted of the confused sound was selected from the database so that the syllable structures of the confused sound and misidentified sound were identical.
- 2) **Identified /i/ and /u/ in the low-familiarity word** The most important acoustic feature to distinguish /i/ and /u/ is the second formant (F2), expecting its range overlaps that of the third formant (F3) of /r-/l/. Words consisted of monosyllabic /i-/u/ or syllable-final /i-/u/ were selected from identified low-familiarity words in the speech perception test.
- 3) One pair of /i-/u/ for each manner of articulation was selected and the phonetic environments of vowels were identical.
- 4) **Non-words contrasting English /r-/l/** Non-words contrasting /r-/l/ at the 5 phonetic positions used in the perceptual training were recorded by a female American English speaker. One pair for each phonetic environment was selected for the analyses.

Procedures

All experiments were carried out in a soundproof room with a background noise of less than 21 dB. The audiogram measured hearing sensitivity at ten frequencies (125, 250, 500, 750, 1000, 2000, 3000, 4000, 6000, 8000 Hz). A Rion AA-72B audiometer was used. In the speech perception test, each participant heard the stimuli through a headphone (AKG K272HD) from a PC (ThinkPad x121e) via a headphone amplifier (Fostex HP-A3) and answered by typing the keyboard. Stimuli for practice were presented at 65.3 dB on the average, and then participants were required to modify the volume if needed. They could replay the stimulus as often as they wished. Low-familiarity and high-familiarity words were presented in a session whereas monosyllabic words were presented in a session. The stimuli were presented randomly.

RESULTS

Hearing Threshold

Hearing thresholds averaged across both ears and participants within each age group were shown in FIGURE 1. The young and middle-aged adults' thresholds were essentially normal. In contrast, older adults demonstrated hearing losses at higher frequencies.

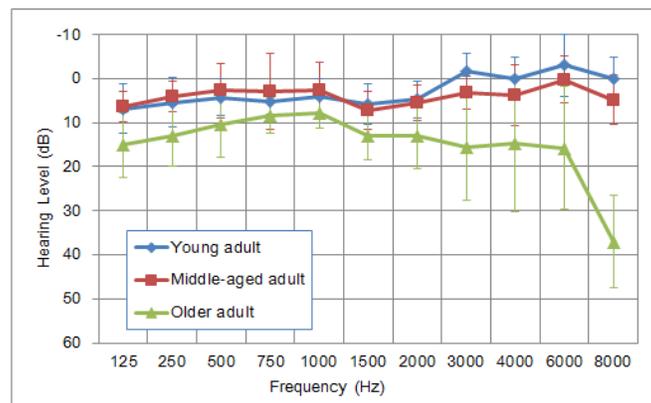


FIGURE 1. Average of hearing thresholds

The thresholds were submitted to a two-factor analysis of variance (ANOVA), where frequency was a within-subject variable, and age group (young, middle-aged and older) was a between-subject variable. The interaction between frequency and age group was significant [$F(11.36, 119.30)=5.40, p<.01$]. Results of a multiple comparison indicated that the threshold of older was significantly higher than those of young and middle-aged. On the other hand, there was no significant difference between young and middle-aged.

Speech Perception Test

Correct response rate for each syllable in the words was calculated, ignoring clear notational mismatches. Then they were averaged over the participants for age group, familiarity and syllable-initial sound (FIGURE 2). It is shown that older adults misidentified the several specific sounds. Participants in all age groups misidentified the specific words, the margin in the sound files of those words was so short that the sound file did not contain the whole speech sound. Older adult's accuracies for those words were lower than those of young and middle-aged.

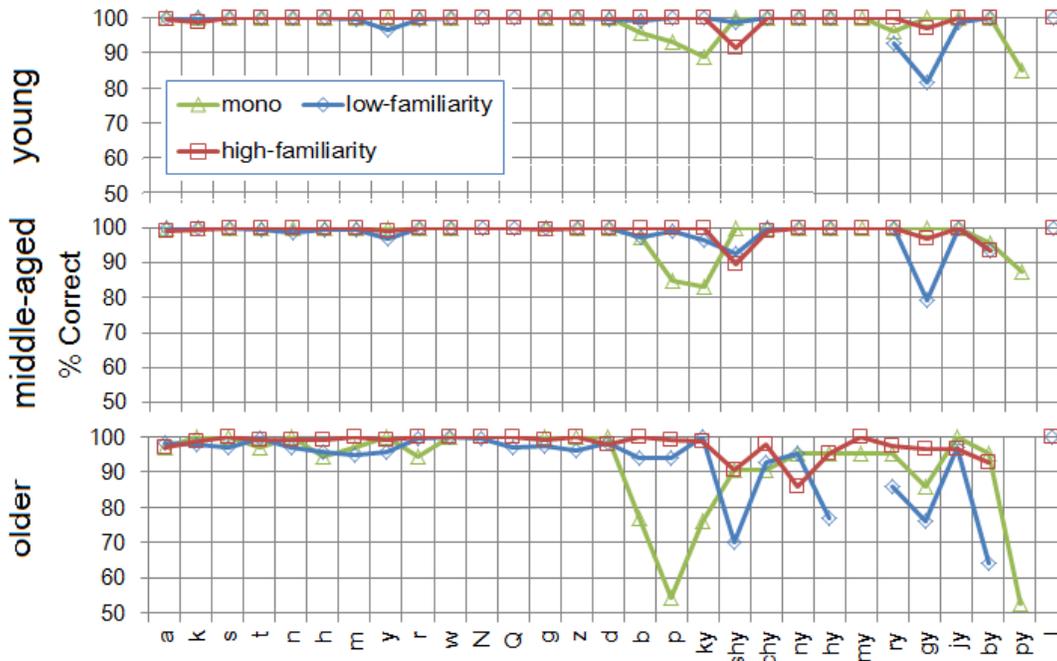


FIGURE 2. Accuracy by syllable-initial sound, age group and familiarity

Syllable-initial phonemes were classified into 10 groups based on the manner of articulation to find patterns of misidentification. The sound classifications were as follows; vowel(a, i, u, e, o), voiceless-stop (k, t, p), voiced-stop (g, d, b), voiceless-fricative (s, h, shy), voiced-fricative (z, j), affricate (ch, ts), nasal (n, m), contracted (ky, ny, hy, my, ry, gy, by, py), semi-vowel(w, y) and flap (r).

The accuracies were averaged for the sound classification and familiarity. The averaged accuracies were submitted to a two-factor analysis of variance (ANOVA), where sound classification was a within-subject variable, and age group (young, middle-aged and older) was a between-subject variable, separately for each familiarity.

As for low-familiarity words, the interaction between sound classification and age group were significant [$F(6.06, 63.68)=2.66, p<.01$]. Results of a multiple comparison indicated that the accuracy of older was significantly lower than those of young and middle-aged for voiced-stop and voiced-fricative, lower than that of young for voiceless-stop, nasal and contracted ($p < .05$).

As for high-familiarity words, the interaction between sound classification and age group was not significant [$F(18, 189)=1.60, NS$].

As for monosyllabic words, the interaction between sound classification and age group was significant [$F(18, 189)=4.22, p<.01$]. Results of a multiple comparison indicated that accuracy of older was significantly lower than those of young and middle-aged for voiceless-stop and contracted ($p < .05$).

In summary, older misidentified stops, fricatives, nasals and contracted when the stimuli were low-familiarity or monosyllabic words, suggesting they fail to perceive the acoustic features of these sounds. When the stimuli were high-familiarity words, there was no difference between age groups. There is no difference between young and middle-aged no matter what familiarity was.

Acoustic Analyses

Misidentified Sounds and Confused Sounds

Confusions shared by a number of individuals were selected and classified as follows; between stops, voiceless-voiced, stop to flap and between fricatives. Duration and vowel onset time (VOT) or spectral center of gravity (SCG) for each misidentified / confused sound was measured using Praat. Averaged values for each confusion were shown in Table 2. Nasals were not analyzed in this paper because of difficulties to track the formant movement.

The results showed that pairs of short duration (shorter than 80 ms), or voiceless and voiced which is characterized by VOT (smaller than ± 80 ms), or fricatives with high SCG (higher than 5000 Hz) were often confused.

TABLE 2. Misidentified Sounds and Confused Sounds in Speech Perception test

Misidentified Sound		Confused Sound	
Consonant	Duration (VOT) [ms]	Consonant	Duration (VOT) [ms]
Between stops			
p	71 (71)	k	36 (36)
Voiceless - Voiced			
p	67 (67)	b	40 (-24)
k	31 (31)	g	58 (-25)
p	17 (17)	b	11 (-44)
Stop to Flap			
p	14 (14)	r	40 (40)
b	8 (-51)	r	24 (24)
Consonant	SCG [Hz]	Consonant	SCG [Hz]
Between fricatives			
h	5037	s	7228
sh	5779	s	6350

Identified /i/ or /u/ in the Low-familiarity Words

Each vowel's duration and frequencies of F1 and F2 at the mid-point were measured and averaged for each vowel (Table 3). The result showed that the durations of these identified sounds were longer than those of misidentified short sounds, and F2 are lower than those of misidentified sounds with high-frequency.

TABLE 3. Identified Sounds in the Speech Perception test

Vowel	Duration (average) [ms]	F1 (average) [Hz]	F2 (average) [Hz]
/i/	86 -196 (121)	276 - 714 (392)	2484 - 2889 (2665)
/u/	100 - 149 (129)	312 - 462 (380)	1490 - 2123 (1956)

American English /r/-/l/

Each consonant's duration and frequency of F3 at the mid-point were measured and averaged for each consonant (Table 4). The result showed that the duration of these sounds was longer than those of identified sounds in the speech perception test, and that difference between cue frequencies (F3) is larger than that of identified sounds in the speech perception test (F2). F3 frequencies ranged at high frequency region when the utterances were produced by a female speaker. Nevertheless, cue frequency of /r/ was lower than or equal to cue frequencies of the identified sounds in the speech perception test.

TABLE 4. Duration and F3 frequency of /r/-/l/

Consonant	Duration (average) [ms]	F3 (average) [Hz]
/r/	139 - 302 (205)	2058 - 2188 (2130)
/l/	123 - 330 (226)	2878 - 3364 (3169)

Thus, it was shown that the cue frequencies of /r/-/l/ were lower, the duration of /r/-/l/ were longer and the differences between cue frequencies were wider than identified and misidentified sounds, suggesting that adults had capability to perceive the physical acoustic features of /r/-/l/ even when they had auditory aging.

DISCUSSIONS

It was demonstrated that older adults had difficulties to identify sounds with short duration or with spectral cue at high-frequency, suggesting that older adults were not able to perceive the specific speech acoustic features because of elevation threshold and impaired temporal-resolution.

The acoustic analyses revealed that the acoustic features of /r/ and /l/ ranged at the region they could perceive the signals. Even the older adults suffered from hearing losses, they should have perceived the acoustic features to distinguish /r/ and /l/ during the perceptual training. Thus, it is hard to say that adults, including older adults with auditory aging, were inhibited from perceiving the acoustic features, and also hard to say that auditory aging brought the difficulties to distinguish and learn/r/-/l/.

As for the interaction between hearing ability and age, the results of hearing threshold measurement and accuracies of speech perception test revealed that young adults and middle-aged adults had comparable hearing abilities whereas older adults had lower hearing abilities than those age groups. That is, as far as investigated in this paper, hearing ability does not interact with age as training effects did [4]. It is hard to link auditory aging with the decrease of the training effects directly.

Therefore, we can conclude that auditory aging does not fully account for aging effect on acquisition of /r/-/l/ contrast.

It is also shown that older adults compensated for insufficient acoustic features by their acquired lexical knowledge to represent phonemes. Since young and middle-aged had relatively high accuracies for all word-familiarity, suggesting they did not have auditory aging. They were able to perceive acoustic features. However, small differences in accuracies between familiarities can be seen in Figure 2, that is, they compensated when acoustic features were insufficient. The effect of word-familiarity which was seen in adults, typically older adults indicates that they have been experienced compensation for insufficient acoustic features by their acquired knowledge.

Although it is difficult to confirm age-related changes in compensation, it is partially confirmed that adults compensate for insufficient acoustic features by their acquired knowledge.

As stated previously, the learning difficulties shown by older adults concerning phonetic environment corresponded to identifying difficulties by Japanese speakers, suggesting that adults became more influenced by language-specific perception with age, not only in perception but also in learning. Further work is needed to assess whether the decrease in training effects was caused by age-related changes in compensation.

CONCLUSIONS

The present study confirmed that auditory aging does not fully account for aging effect on L2 learning, other changes in upper processing should be taken into account, and partially confirmed that adults compensated for insufficient acoustic features by applying their acquired knowledge.

REFERENCES

1. C. T. Best, "A direct realist view of cross-language speech perception", in *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange (Timonium, MD: York Press, 1995), pp. 171–204.
2. J. E. Flege, G. H. Yeni-Komshian, and S. Liu, "Age constraints on second-language acquisition", *Journal of Memory and Language* **41**, 78–104 (1999).
3. A. R. Bradlow, R. Akahane-Yamada, D. B. Pisoni, and Y. Tohkura, "Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production", *Attention, Perception, & Psychophysics* **61**, 977–985 (1999).
4. R. Kubo, R. Akahane-Yamada, and M. Akagi, "Effects of perceptual training on the ability of elderly adults of Japanese speakers to identify American English /r/-/l/ phonetic contrasts", *Proc. Mtgs. Acoust.* p. 124 (2012).