Development of Speaker Identification in Young Children

K. Saito*, K. Asakawa, Y. Shimura** and S. Imaizumi

1.Introduction

In the study of language development, there has been a great deal of research focused on the domains of sounds, word meaning and grammar. The research into these domains has put an emphasis on the child's ability to distinguish linguistically critical units or categories in each domain. Most developmental researchers selectively analyze linguistically relevant variations, and are apt to ignore other variations as if they were irrelevant noise. However, some attempts in the fields of speech science and social psychology have focused on one of these linguistically irrelevant variations, i.e., between and within speaker variation, as an important variable in adult communication. Speaker identification means that each speaker's voice is identified by the hearer through both processes of discriminating the difference with other speakers' voices, and identifing the commonality within one speaker's voice tokens. From the developmental point of view as well, speaker identification is important not only for communication development but for the socio-cognitive development of identification of person and his/her social categories (age, gender, etc.). From our view, a child's ability to manage between- and within-speaker variation and to identify speakers is supposed to be share the same fundamentalcategorization mechanism as phonetic or lexical categorization of linguistic units.

In this paper we will specifically examine how children differentiate each person's voice and what characteristics aid the child in making these distinctions. We will examine two characteristics, the social factor and the psychoacoustic factor, which seem to aid children in discriminating voices. (1) The degree of familiarity and the quality of the social relationship with the speaker can be expected to contribute to the child's ability to distinguish individual voices. (2) The quality of the voice can also expected to be important in voice distinction. For example high toned voices may be more easily distinguished than lower toned voices.

Our research purpose is to determine if (1) theredevelopmental differences in speaker identification ability; (2) whether quality of social relationships has an effect on speaker identification; (3) whether the quality of voice has an effect on speaker identification.

2.Method

Subjects: Nursery School Children. Each group included all children from one class: eight three-year-olds (6 boys, 2 girls) aged 3:2 - 4:1 (average 3:7); nineteen four-year-olds (8 boys, 11 girls) aged 4:3 - 5:2 (average 4:8); twenty five-year-olds (14 boys, 6 girls) aged 5:3 - 6:3 (average 5:8).

* Department of Art, Kokugakuin University, Tokyo, Japan
** Department of Education, Saitama University, Urawa-shi, Saitama, Japan

Stimulus Voices: Children's voices were recorded as they named fruit from a picture book (the words spoken were in Japanese: MIKAN, RINGO, BANANA, which mean orange, apple and banana, respectively). Then, the sets of test voices were made. Three words by each child were repeated three times in a list. For the three-year-olds this list included the voices of all the class members, but for the four- and five-year-olds the list did not include the entire class. The latter two lists included the class leader, the most talkative child, the least talkative child, the child with the most friends, the child with the least friends, the friend of the child who was listening, and the listener's own voice. We also added two voices (one boy, one girl) of the same age which were not known to the children. Each list was different for each listener, and the voices were played in random order.

Evaluation of social relationships: We examined the following factors by interviewing the nursery school teachers and children about what the nature of the following relationships in the class: 1, the class leader. 2, whether the children had a lot of friends or not 3, whether the children were talkative or not. 4, each child's most close friend and least close friends.

Evaluation of voice quality: Voice quality was measured by an acoustic analysis of the following features of the voice. 1, fundamental frequency (Fo); 2, word length ; 3, first, second, third formant frequencies.

Procedure of Identification: Each child heard the tape in a small nursery school room. When they heard each voice they were asked to match the voice with a picture from a set of photos of all the speakers.

3.Results

(1) Developmental differences: Figure 1 shows that the rate of speaker identification increases with age. The average rate of speaker identification was 20% for three-year-olds, 29% for four-year-olds and 49% for five-year-olds. Statistically, there were significant age differeneces (p<.001) except between three- and four-year-olds. (2) Figure 2 shows the effect of the quality of social relationships: Identification of class members' voices increased with age. On the other hand, Figure 2 shows that strangers' voices were discriminated significantly better by the three-year-olds than by the four- or five-year-olds. As for the effects of social relationships, the four- year-olds identified the least talkative boy and the boy with the fewest friends most easily. In contrast, the five-year-olds most easily identified the most talkative boys and girls. The five-year-olds did not recongnize those whom they had identified as their friend more easily. The leader of the class was not identified more than the others. (3) Quality of voice: We divided the 25 speakers into two groups: a highly identified group (9 speakers) and a lowly identified group (16 speakers). Each speaker's three words were acoustically analyzed for fundamental frequencies, word pitch patterns, word duration and vowel formants. Results showed that these qualities were not different between the highly identified and lowly identified groups.

4.Discussion

Development of general voice identification ability shows improvement with age, however, a former study (Saito et al 1994) showed that even university students' rate of speaker identification was only 40% correct. The university subjects were 16 members of the same class. This result was

below expectations. However, nursery teachers identified children's voices almost 100% of the time in the same experimental condition. Thus, the factors determining adults ability for speaker identification in an experimental situation remain an unanswered question. We feel there are some factors as yet undetermined which affect adults' ability to identify voices.

Figure 2 shows that the rate of identification of the listener's own voice increased with age, but at a slower pace than that for identifying other voices. Also, for four- and-five-year-olds identification of own voice was lower than for other voices. It is likely that the listener's tape recorded voice is unfamiliar to the listener. However, more importantly, this evidence makes us expect that there are different fundamental processes between identifying other voices and one's own voice in speaker identification . The effect of social relations on the ability of the listener to identify the speaker was also unclear. Acoustical features studied here were shown not to have a discernable affect on speaker identification. A more strictly controlled study is required.

The identification of a speaker in real-life situations is not only accomplished with auditory information but also with visual information. In these situations, we can use mouth movement which synchronizes with speech and also we can use the direction of the source of sound for speaker identification. Thus it is possible that the type of task in the present study may have resulted in difficulties in identifying the speakers not only for children but for adults. If we use cross-modal stimuli the identification rate may increase. A study of this kind should be carried out to clarify speaker identification development.

Acknowledgements

References

Akiyama,K.,Saito,K.,Shimura,Y.,and Imaizumi,S.(1995) "Youji ni okeru wasya ninsiki no hattatsu (1)," Paper presented at the 37th annual meeting of the Japanese Educational Psychological Association.

Saito,K.,Shimura,Y.,and Akiyama,K.(1994). "Youji no tasya no koe no ninsiki nouryoku no hattatsu," Paper presented at the 58th annual meeting of the Japanese Psychological Association. pp.277.
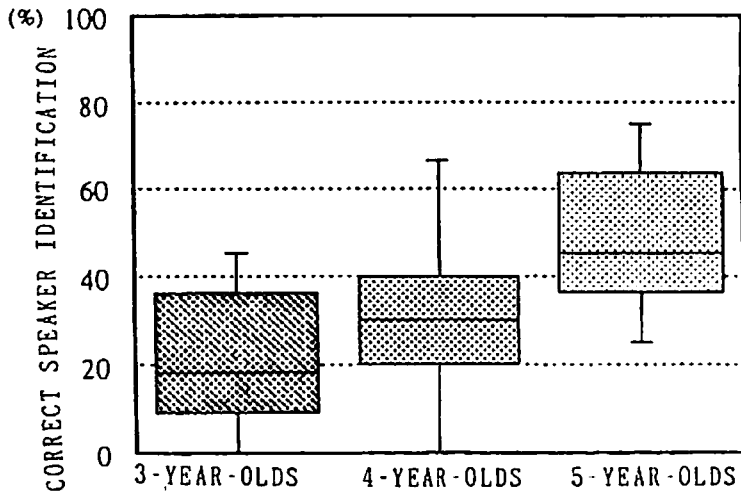
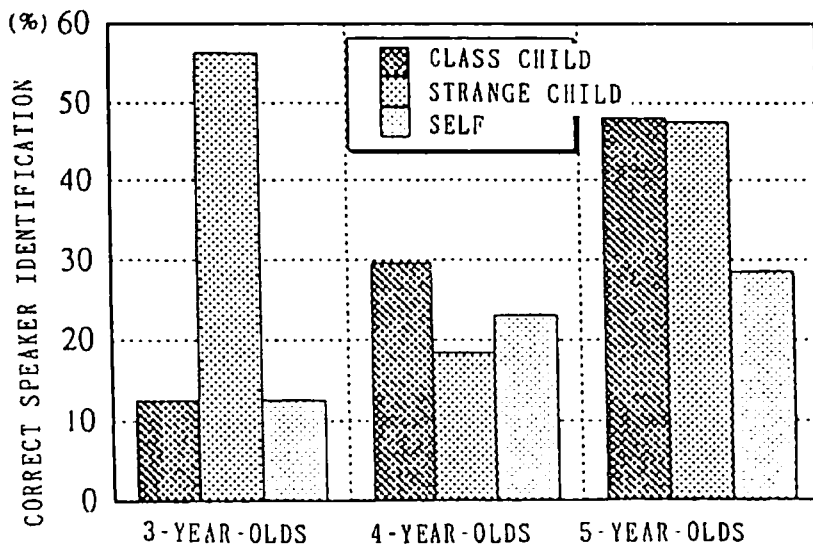Figure 1. Speaker identification in each age group.

Figure 2. Mean correct speaker identification for each speaker by age group.