



## AUDITION AND SPEECH PERCEPTION IN THE CHIMPANZEE

Shozo KOJIMA

Primate Research Institute, Kyoto University  
Kanrin, Inuyama, Aichi, 484 JAPAN

### ABSTRACT

The basic auditory functions and the perception of vowel-like vocal sounds (grunts) were studied in the chimpanzee. The chimpanzee showed a W-shaped auditory sensitivity function. The chimpanzee was more sensitive to 1-kHz and 8-kHz than 2-kHz and 4-kHz tones. It was found that the first formant was more important than the second formant for the perception of these vocal sounds. The chimpanzee did not vocalize vowel-like sounds [i] and [e]. The relationship between the perception of grunts and the auditory sensitivity function and between the perception and the production of grunts of the chimpanzee was discussed.

### I. INTRODUCTION

Of all primates, only humans have languages. Recent research on the learning of 'language' by chimpanzees and the vocal behaviors of monkeys in the field suggest that human languages are results of mutations and natural selection. Human languages are spoken languages which was the vocal-auditory channel. If we assume that vocal-auditory functions of early hominids were not so different from those of living apes, experiments that clarify differences in vocal-auditory functions between apes and humans may help us to understand the origin and evolution of human speech. Despite their obvious importance, there are only a few studies of the hearing and vocal behaviors of apes. In the present experiments, basic auditory functions and the perception and the production of vowel-like vocal sounds were studied.

### II. GENERAL METHOD

Subjects were three female chimpanzees (Pen, 9 years old, Popo, 7 years old and Pan, 5 years old).

The subjects sat in a chair and wore earphones (TDH-39). They faced a panel which contained a lamp and a key in a sound-attenuating room. Two reaction time tasks were employed, in which a trial was started by the illumination of the lamp. In one of the tasks (RT-I), pressing the key in the presence of the light started the foreperiod which lasted 2-6 s and then a pure tone was presented. The subjects were required to release the key as quickly as possible when they detected the tone. The maximum duration of the tone was 1 s. A reaction time of key release was measured. In the other task (RT-II), pressing the key in the presence of the light initiated repeated presentation of a background vowel, with a 1-s interstimulus interval. The vowel was repeated two to six times and then changed to another, target vowel. The subjects were trained as quickly as possible when they detected the change in vowels. Key release within 1 s or 1.5 s produced a reward. The reaction time of key release and detection rates were measured. In general, longer reaction times were required for discrimination between perceptually similar vowels. Stimulus conditions were described in each section.

### III. EXPERIMENT I: AUDITORY SENSITIVITY

Pen and Popo were used. The task used was RT-I. The frequencies of the pure tone were from 125 Hz to 32 kHz in 1-octave intervals as well as 24 kHz. Intensities were selected in 10-dB steps from -10 dB to 90 dB SPL (dB re 20 Pa) and changed randomly every eight trials. Tone signals were presented to the subject's right ear.

A curvilinear regression line was fitted to intensity-reaction time functions and equal reaction time contours were derived from them. The absolute threshold at each frequency was defined as the intensity associated with a mean reaction time of 800 ms. The chimpanzees were sensitive to 1-kHz and 8 kHz tones and showed a sensitivity loss to 2-kHz and 4-kHz tones. Thus, they showed a W-shaped auditory sensitivity function (see Fig 1). The chimpanzees were more sensitive to high frequencies but less sensitive to middle and low frequencies than humans.

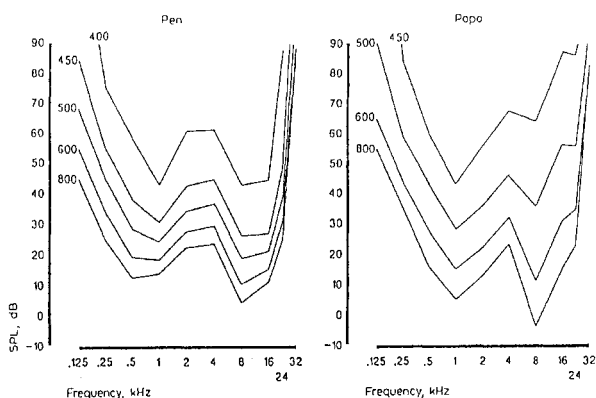


Fig. 1. Equal loudness (latency) contours for chimpanzees. The lowest line in each panel indicates an auditory sensitivity curve.

#### IV. EXPERIMENT II: DIFFERENCE THRESHOLDS

Popo and Pan were the subjects. The task was RT-II. The frequency of the standard (background) stimulus was from 0.5 kHz to 4 kHz in 1-octave intervals and its intensity was 70 dB SPL for measurements of frequency difference thresholds. Frequency differences between the standard and the comparison (target) stimulus were +5 Hz to +80 Hz (6 steps). For measurements of intensity difference thresholds, the frequency of the standard stimulus was 1 kHz and its intensity ranged from 50 dB to 90 dB SPL. Intensity differences between the standard and the comparison stimulus were between +1 dB and 8 dB SPL (6 steps). Difference thresholds were defined as the difference in frequency or intensity between the standard and comparison stimuli that was accompanied by a detection rate of 50 %.

Fig. 2 shows frequency and intensity difference thresholds. As shown, difference thresholds for frequency were about 10-15 Hz and were relatively small between .5 kHz and 2 kHz, but rose abruptly to 35-40 Hz at 4 kHz. The chimpanzees showed U-

shaped intensity-threshold functions, especially for intensity increment thresholds. Thresholds for discrimination of intensity increments were usually smaller than those for intensity decrements, when the intensity of the tones was low.

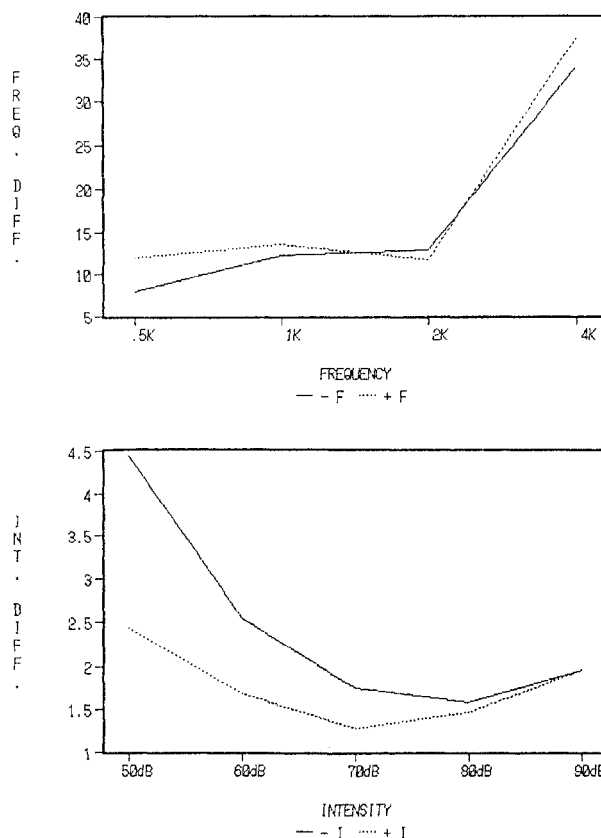


Fig. 2. Difference thresholds for frequency (upper panel) and those for intensity (lower panel).

#### V. EXPERIMENT III: AUDITORY EVOKED POTENTIALS

In this experiment, Japanese monkeys were used as subjects. Sine wave tone bursts of various frequencies were presented and auditory brainstem responses (ABR), whole nerve action potentials (AP) and cochlear microphonics (CM) were recorded. Frequencies of the tone bursts were from 0.5 kHz to 8 kHz in 1-octave intervals and 6 kHz. The tone burst was 20 ms in duration with 5 ms rise and fall times and presented 2048 times at a rate of 10/s. The intensity was between 20 dB and 90 dB SPL. The ABRs were recorded from a skin electrode placed on the vertex. A reference electrode was placed on the earlobe ipsilateral

to the headphone to which tone bursts were presented. The CM and the AP were recorded from a silver ball electrode attached to the wall of the external auditory meatus near the ear drum by the electrode paste. The reference electrode was placed on the earlobe ipsilateral to the recording electrode. The tone burst was presented 512 times at a rate of 10/s through a loudspeaker.

Fig. 3 shows auditory sensitivity functions based on the evoked potentials and behavior. The Japanese monkey showed a sensitivity loss at 6 kHz in all the sensitivity functions. The resonance frequency of the external ear canal was 6 kHz, which indicates that the origins of the sensitivity loss exist in somewhere between the eardrum and hair cells in the cochlea, that is, in the peripheral auditory system.

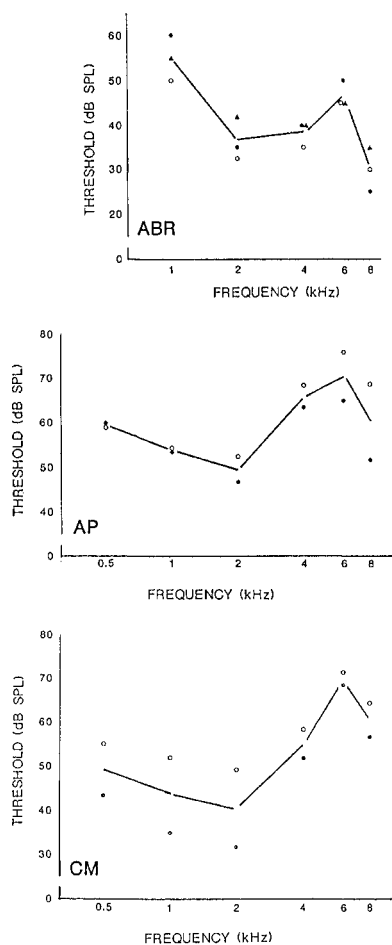


Fig. 3. Auditory sensitivity functions based on evoked potentials (ABR, AP, CM).

#### VI. EXPERIMENT IV: THE PERCEPTION OF VOWEL-LIKE VOCAL SOUNDS

Recently Kojima & Kiritani (1989) reported that the chimpanzee required long reaction times for discrimination of synthetic [i] from [u] and [e] from [o], that is, the chimpanzee need long latencies for discrimination between vowels based on differences in frequency of the second formant. To examine the generality of this result, the perception of vowel-like vocal sounds (grunts) of the chimpanzee was studied.

In this experiment, components of an original grunt were deleted by a digital filter and the subjects were required to discriminate filtered grunt from the original. The upper panel of Fig. 4 shows the sonograms of an original and filtered grunts. The original grunt [ORG] was heard as [o]. For one of the filtered sounds, components other than the first formant were deleted (in other words, only F1 was remained [F1]). For another filtered grunt, only the second harmonic sound was deleted [H2-]. The intensity of the original and filtered stimuli was matched. The task was RT-II in which the background voice was the original grunt and the

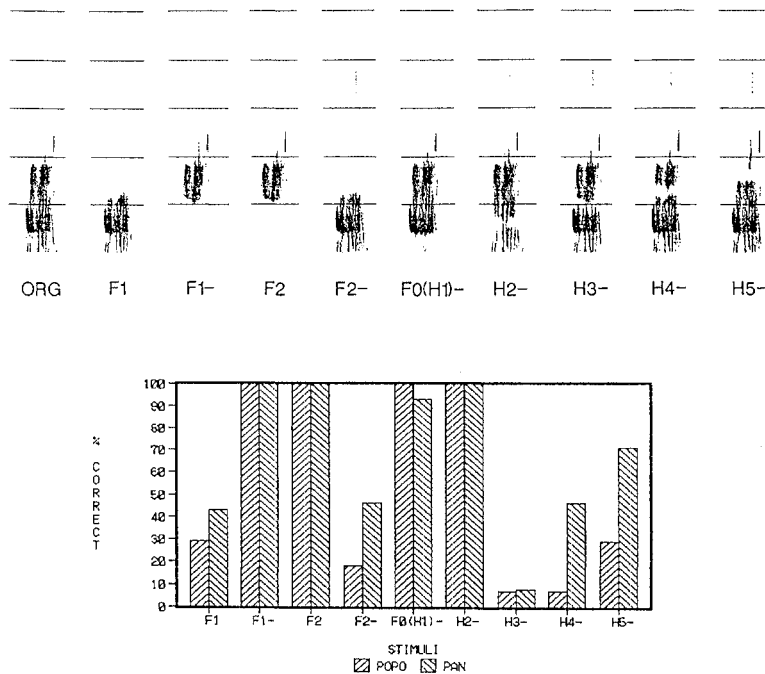


Fig. 4. Upper: The sonograms of the original grunt [o] and filtered grunts used in Experiment IV. Lower: The results of the discrimination test.

target voice was one of the filtered grunts.

The lower panel of Fig.4 shows the percentage of correct detections for the two chimpanzees. As to formants, when F1 was remained in the filtered grunt, as in the case of [F1] and [F2-], percent correct detections were low. On the other hand, when F2 was remained in the filtered grunt, as in the case of [F1-] and [F2], discrimination performance was perfect. This indicates that filtered grunts without F1 are perceptually different from the original grunt. In other words, consistent with the previous experiment, F1 is more important than F2 for the perception of the vowel-like vocal sound of the chimpanzee. About harmonic sounds, the performance was higher when F0(H1) [F0(h1)-] or H2[H2-] was deleted, which constitutes F1, than when H4[H4-] or H5[H5-] which constitutes F2, was deleted from the original grunt. Similar results were obtained when chimpanzee grunts, which were heard as [u] or [a], were used.

#### VII. EXPERIMENT V: THE PRODUCTION OF VOWEL-LIKE VOCAL SOUNDS

The author recorded vowel-like sounds (grunts) of a chimpanzee baby which were heard as [u], [o] or [a]. No grunts heard as [i] or [e] were recorded, for which the high position of the larynx and the small pharynx of the chimpanzee are responsible (Chiba and Kajiyama, 1941).

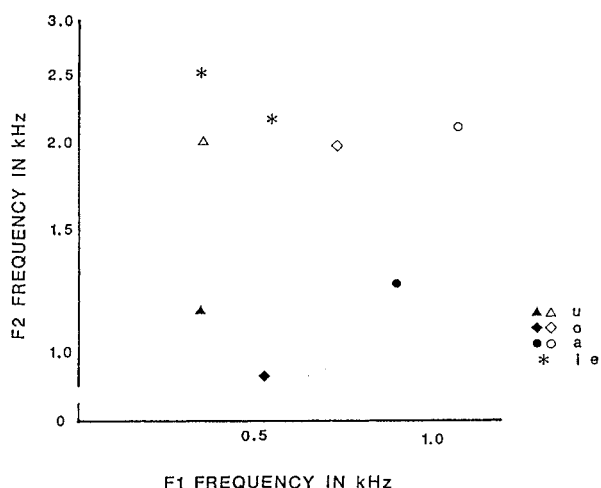


Fig. 5. Vowel-like vocal sounds of the chimpanzee (open symbols) and Japanese vowels (filled symbols and asterisks) in the F1-F2 plane.

Fig.5 shows grunts of the chimpanzee in the F1-F2 plane along with Japanese vowels. As shown in this figure, chimpanzee grunts have about the same F2 frequency and the frequency of F1 was different between the grunts. Because of the small vocal tract, the frequencies of F1 and F2 of chimpanzee grunts were higher than those of human vowels.

#### VIII. GENERAL DISCUSSION

The perception of vowels and vowel-like vocal sounds experiment indicate that F1 is more important than F2 for the perception of vowels and grunts for the chimpanzee (EXPERIMENT III). There may be two coexistent interpretations to these results. One interpretation is related to the basic auditory function of the chimpanzee. As shown in EXPERIMENT I, the chimpanzee has a W-shaped auditory sensitivity function. That is, the chimpanzee was more sensitive to 1 kHz and to 8 kHz than to 2 kHz and to 4 kHz (see Fig.1). The sensitivity loss at mid-frequencies may be responsible for these results, because the frequency of F2 often falls into this mid-frequency range. Difference thresholds for frequencies (F) may not be responsible for the results, because the value of F were about the same between 0.5 kHz and 2 kHz (EXPERIMENT II).

The other interpretation is related to the repertoire of vowel-like vocal sounds (grunts) of the chimpanzee. As shown in EXPERIMENT V, the chimpanzee baby rarely vocalized vowel-like sounds [i] and [e], to which the small pharynx is responsible. The close relationship between the production and the perception of vocal sounds suggests that it may be difficult for the chimpanzee to perceive vocal sounds [i] and [e].

These two interpretations are not mutually exclusive. Both the hearing and the production of vocal sounds in the chimpanzee may be responsible for the present and previous results.

#### REFERENCES

- Chiba, T. and Kajiyama, M. (1941) The vowel: its nature and structure. Tokyo: Tokyo-Kaiseikan.
- Kojima, S. (1990) Comparison of auditory functions in the chimpanzee and human. *Folia Primatol.*, in press.
- Kojima, S. and Kiritani, S. (1989) Vocal-auditory functions in the chimpanzee: Vowel perception. *Int. J. Primatol.*, 10, 199-213.