QUI EST QUI: COMMENT LES CHIMPANZES (PAN TROGLODYTES) ET LES HUMAINS APPARIENT UNE VOIX A LA PHOTOGRAPHIE DU VOCALISATEUR DANS UNE TACHE DE RECONNAISSANCE IDENTITAIRE.

WHO IS WHO: HOW CHIMPANZEES *(PAN TROGLODYTES)* AND HUMANS MATCH VOICE TO THE VOCALIZER'S PICTURE IN AN IDENTITY RECOGNITION TASK.

Laura Martinez[1] and Tetsuro Matsuzawa [1]
Language and Intelligence section- Primate Research Institute, Kyoto University.
Kanrin 41, Inuyama, 484-8506, Aichi, Japan.
Tel: 0568-63-0550 – Fax: 0568-62-2428
E-mail: laura@pri.kyoto-u.ac.jp

Résumé:
Les recherches antérieures menées chez le chimpanzé font état d'un lien étroit entre leur comportement vocal et la complexité de leurs relations sociales. En effet, les groupes de chimpanzés sont caractérisés par la variabilité des associations entre individus au sein d'un même groupe ainsi que par une fréquente migration entre différentes communautés, impliquant ainsi un grand nombre d'individus. Cette dispersion spatiale requiert nécessairement une reconnaissance non visuelle efficace de nombreux vocalisateurs. La présente étude a pour but de comprendre comment sont reliées entre elles les informations visuelles et auditives impliquées dans une telle catégorisation de l'identité. Dans une première expérience, s'appuyant sur une tâche d'appariement intermodal, les sujets devaient apparier les retransmissions de vocalisations émises par trois chimpanzés familiers avec des photographies les représentant. Parmi un groupe de 14 chimpanzés vivants en captivité, une femelle adulte appelée Pan, préalablement experte en tâches audio-visuelles, a obtenu 90% de bonnes réponses, tandis que deux autres femelles adultes, nommées Aï et Chloé, n'ayant eu aucune expérience préalable avec ce genre de tâche, ont obtenu 46 et 44% de bonnes réponses. Les sujets humains (n=11), tous très familiarisés avec ces chimpanzés, ont été testés dans les mêmes conditions et ont obtenus un score intermédiaire de 74% en moyenne. Dans les expériences en cours, nous manipulons plusieurs paramètres de la procédure afin de faciliter le processus d'acquisition chez les individus inexpérimentés. En parallèle, nous testons les trois enfants chimpanzés (âgés de 5 ans) pour identifier toute possible corrélation avec l'âge ou l'expérience individuelle. Dans de prochaines expériences, nous utiliserons des voix et visages d'humains comme stimuli afin d'évaluer de quelle façon la familiarité et « l'effet autre-espèce » peut influer sur la reconnaissance intermodale de l'identité.
Mots-clé : Chimpanzé. Vocalisation. Reconnaissance vocale individuelle. Appariement intermodal.

Abstract:
Previous studies have emphasized that the vocal behavior of chimpanzees is closely related to the complexity of their social relations, which involves frequent intra- and inter-group mixing among a large number of individuals. This spatial dispersion requires efficient non-visual recognition of the vocalizer. In this study, we attempt to investigate the relationship between auditory and visual categorization of identity through a cross-modal matching-to-sample task. In preliminary experiments, subjects were required to match playbacks of pant-hoot calls from 3 familiar chimpanzees with the corresponding vocalizers' facial pictures. Among a group of 14 captive chimpanzees, one adult female named Pan, an expert in audio-visual tasks prior to this study, obtained 90% of correct responses, while two other female chimpanzees named Ai and Chloe, with no prior experience in audio-visual task, scored 46 %

and 44%. Human subjects familiar to these chimpanzees (n=11) were tested under the same conditions. Their performances were located just between the expert and inexperienced chimpanzees, with 74 % correct on average (ranging from 54 to 88% among subjects). In ongoing experiments with one inexperienced chimpanzee, we are manipulating several parameters of the procedure to facilitate the acquisition process and besides, we are carrying out the task with three infant chimpanzees (5 years old) to look forward any possible correlation with age or individual experiences among subjects. In further experiment, we aim to use human voice and face as stimuli in order to assess how familiarity and "other-species effect" can affect this cross-modal identity recognition.
Key-words : Chimpanzee. Vocalization. Vocal individual recognition. Cross-modal matching.

Introduction:

The auditory perception and categorization of the world

Among all the animal kingdom, including humans, communication involves complex behavior in multiple sensory channels (Partan and Marler, 2005). Concerning auditory communication, many researches have focus on the existence of a conceptual understanding of the information provided by the conspecifics, especially in non-human primates (Cheney and Seyfarth, 1990). Among anthropoids, chimpanzees posses a broad and complex repertoire of vocalizations, notably varying with the context of emission (Goodall, 1986 ; Marler and Tenaza, 1977). Many studies have emphasized that this vocal behavior is characterized by an intra-group convergence and inter-population variability of calls structure (Marshall et al., 1999 ; Arcadi, 1996) and also by an acoustic distinctiveness of vocalizations among individuals (Mitani et al., 1996).

The social group and its living space in wild chimpanzees

Chimpanzees living in their natural environment have to deal with a visual environment partially occluded as well as with an important spatial dispersion between individuals described as a "fision-fussion" social organization (Goodall, 1986 ; Sugiyama, 1984 ; Sakura, 1994), implying an intra-group mixing and inter-group migration. This requires an efficient auditory understanding of the social environment (Crockford at al., 2004).

Previous experimental studies

Experimental studies carried out in monkeys have demonstrated that some neurons in the Marmosets are specialized in the encoding of acoustic pitch (Bendor and Wang, 2005); that Macaques Rhesus are able to establish the correspondence between auditory and visual components of calls (Ghazanfar and Logothetis, 2003), and moreover, that this monkeys have the capacity to learn auditory concepts (Wright et al., 1990). In chimpanzees, Kojima and his colleagues have initiated a broad amount of researches at the Primate Research Institute of Kyoto University concerning the auditory and vocal functions of these apes. After the proficient training of one chimpanzee, named Pan, they succeeded to teach her various auditory-visual matching-to-sample tasks (Hashiya and Kojima, 2001) including the discrimination of conspecifics (Hashiya,1999), objects, humans (for a review see Kojima, 2003). Pan is even able to match a specific vocalization with the corresponding vocalizing face (Izumi and Kojima, 2004). In humans, Patterson and Werker (2002) demonstrated that infants at the age of 4.5 months showed no evidence of matching face and voice on the basis of gender but were able to match them on the basis of the vowel.

Global perspective of research

This research aims to enhance our knowledge of chimpanzees' auditory cognition which remains mostly unknown. Among others, we intend to investigate the process of acquisition, development and long-term retention of the vocal identity recognition; to explore the acoustic understanding of social interactions in relation with the identity of the vocalizers and finally to asses the influence of "other-species effect" and familiarity in the classificatory process comparing humans and chimpanzees.

Main objectives of this first experiment

The first objective was to test the ability to recognize familiar individuals from their voice using a cross-modal matching-to-sample (MTS) procedure. For that purpose we compare the performance of chimpanzees and humans according to their prior experience. The second objective was to determine whether chimpanzees with no prior experience in audio-visual task can progressively learn to explicitly make the relation between a vocal sample and a visual target in a MTS task.

Objective 1:

Method

a) Subjects

Three adult female chimpanzees belonging to the same group participated to this experience. Pan (22 years old) was already an expert in audio-visual task since infancy, whereas Ai (28 y.) and Chloe (25y.) were experts in visual tasks since infancy but absolutely inexperienced in audio-visual task. The human participants were 11 persons being in daily contact with these chimpanzees (3 care-takers, 3 researchers and 5 students).

b) Stimuli

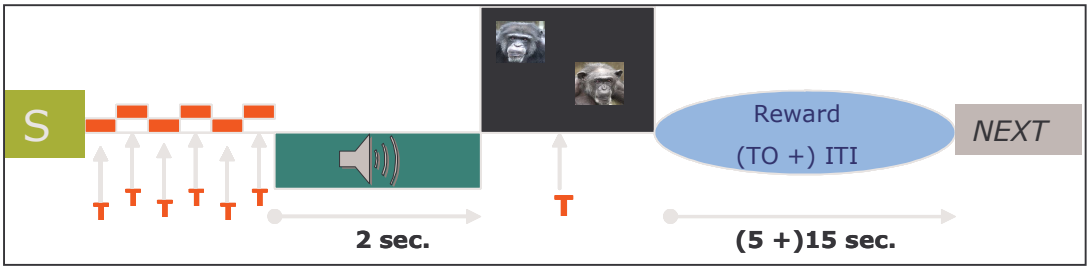
We used unique facial pictures of 3 familiar chimpanzees (Akira, Reiko and Ai) as visual stimuli (size: 10.5 x 10.5 cm), all presenting a frontal view of the individual with mouth close and direct gaze. As acoustic stimuli, we used 4 samples of "pant-hoot" vocalizations (length: 2 seconds) for each of the 3 familiar chimpanzees. All the calls were recorded in the chimpanzees' outdoor compound with a directional microphone and a DAT walkman (48 kHz sampling, 16 bit-precision).

c) Procedure

Two sessions of 24 trials were performed by each subject. Within a session, each audio stimulus were repeated twice, whereas each visual stimulus appeared 8 times.

The procedure is illustrated in figure 1.

Figure 1



Results

Pan's results (90% of correct responses) confirm her skillfulness even after several months without performing audio-visual task. Humans can recognize familiar chimpanzees from their voice but not with the highest accuracy of Pan (average of 74%). Non-expert chimpanzees show results at chance level (44 and 46 %). Median latency of response (between 1100 and 1400 ms.) reveals an intra-individual variability in the response time of human and chimpanzee subjects rather than an inter-individual difference. However, whereas Pan can discriminate the 3 individuals with 80% of accuracy, humans show this level of accuracy only for Akira but this difference it's not significant. Humans show longer latencies for the two individuals more difficult to discriminate. Moreover, Pan's responses were 100 % correct in the recognition of the 4th Ai's audio samples whereas humans showed their worst performance with these samples. Finally, the number of years humans spent in close contact with these chimpanzees doesn't seem directly correlated with their level of performance.

Objective 2:

Method

a) Stimuli

Twelve new pictures were introduced for each individual in the 3rd experimental condition (AV-3). They were subjected to the same rules than in objective 1.

b) Procedure

33 sessions carried out with Ai (792 trials) in order to intend to improve her performances.

We changed different parameters of stimuli and procedure set such as:

- Ai voice and picture removed from the targeted individuals (AV-2)
- Number of pant-hoot samples reduced to 1 and number of target pictures increased to 4 (AV-3)
- Introduction of Time Out (5 seconds) (AV-2)
- Introduction of pictures as visual sample
    - Simultaneous to the audio sample (VAV)
    - Simultaneous to the visual targets (AVV)

Results

The main result consists on Ai's no improvement above chance level in AV conditions. Across the sessions, she shifted between individual, picture and position preference. However, in the Visuo-Audio-Visual condition (VAV), she scored 87% of correct responses.

Discussion and further experiments:

Chimpanzees' voice-face matching-to-sample was not an easy task for human subjects, even after long years of close interaction with them. Inexperienced chimpanzee subjects didn't show a spontaneous understanding of the task .In this experience, the pitch of the chimpanzees' voice seems to be a determinant cue to identify the vocalizer for humans as well as for it has been already demonstrated in chimpanzees (Kojima, 2003; Mitani et al., 1996). Furthers acoustic analyses will be done to explore more precisely the acoustics cues pertinent in this work.

These earlier results do not permit to determine whether inexperienced chimpanzees can learn this difficult audio-visual task. Difficulties can be related to a neglect of audio stimulus or/and to a misunderstanding of the identity-matching nature of the task but VAV's condition results corroborate the disregard of audio stimulus even if it cannot confirm a true understanding of the task. Nevertheless, audio stimuli may influence the responses, at least implicitly. It remains the possibility that the acquisition of this task would be easier for infants, more flexible and not so strongly specialized in

visual tasks, than for adults. This set of experience need to be continued in order to deeply assess the presented objectives of research.

In further experiment, chimpanzee Pan and human subjects will be tested with the same task using human voice and face stimuli. Afterwards, new sets of stimuli including all the members of the group based on trial-unique presentation, will allow to generalize the recognition of chimpanzees from their voice. In the future, we aimed to take advantage of the great Pan's potentiality in performing cross-modal tasks to address related issues. Finally, with the inexperienced chimpanzees, especially with infants, we will keep trying to teach them the task.

References:

Arcadi A.C. (1996). Phrase structure of wild chimpanzee pant hoots: Patterns of production and interpopulation variability. *American Journal of Primatology* **39**:159-178.

Bendor D. and Wang X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature* **436**: 1161-1165.

Crockford C., Herbinger I., Vigilant L. and Boesch C. (2004). Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology* **110**: 221-243.

Ghazanfar A.A. and Logothetis N.K. (2003). Facial expressions linked to monkeys calls. *Nature* **423**: 937-938.

Goodall J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Belknap, Cambridge, Mass.

Hashiya K. and Kojima S. (2001). Acquisition of auditory-visual intermodal matching-to-sample by a chimpanzee (*Pan troglodytes*): Comparison with visual-visual intramodal matching. *Animal Cognition* **4**: 231-239.

Hashiya K. (1999). Auditory-visual intermodal recognition of conspecifics by a chimpanzee (*Pan troglodytes*). *Primate Res* **15**: 333-342.

Izumi A. and Kojima S. (2004). Matching vocalizations to vocalizing faces in a chimpanzee (*Pan troglodytes*). *Animal Cognition* **7**: 179-184.

Kojima S. (2003). *A search for the Origins of Human Speech: Auditory and Vocal Functions of the Chimpanzee.* Kyoto University Press, Kyoto and Trans Pacific Press, Melbourne.

Marler P. and Tenaza R. (1977). Signaling behavior of apes with special reference to vocalization. *In How Animals Communicate* (ed. by T.A. Sebeok), Indiana University Press, pp.965-1033.

Marshall A.J., Wrangham R.W. and Arcadi A.C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour* **58**: 825-830.

Mitani J.C., Gros-Louis J. and Macedonia J.M. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *International Journal of Primatology* **17**(4): 569-583**.**

Partan S.R. and Marler P. (2005). Issues in the classification of multimodal communication signals. *The American Naturalist* **66**(2): 231-245.

Patterson M.L. and Werker J.F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental child Psychology* **81**, 93-115.

Sakura O. (1994). Factors affecting party size and composition of chimpanzees (*Pan troglodytes verus*) at Bossou, Guinea. International Journal of Primatology **15**(2): 167-183.

Cheney D.L. and Seyfarth R.M. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. University of Chicago Press, Chicago and London.

Sugiyama Y. (1984). Population dynamics of wild chimpanzees between at Bossou, Guinea, between 1979-1983. *Primates* **25**: 391-400.

Wright A.A., Shyan M.R. and Jitsumori M. (1990). Auditory same/different concept learning by monkeys. *Animal Learning and Behavior* **18**(3): 287-294.

Acknowledgments:

We gratefully acknowledge all the staff of the Language and Intelligence section at The Primate Research Institute, especially Sana Inoue, Misato Hayashi and Etsuko Nogami for their precious help in conducting the experiments. Thanks to Akihiro Izumi for his inestimable guidance at the beginning of this study and to Professor Shozo Kojima for his pertinent advices. Thanks also due the staff of the Center for Human Evolution Modeling Research and for taking care of chimpanzees. This research is being financed by Grants from the Ministry of education, Science, Sports and Culture, Japan to T. Matsuzawa and also supported by the fellowship to L. Martinez by the same Ministry.