Interjections as a window to the others’ mental states: detection of “intention” of the act by adults and children.

Keywords: intention inference, Japanese interjection, development, imitation, prosody

Department of Behavior and Health Sciences
Yusuke Uto

Introduction

Understanding what others intend to do is important to communicate, cooperate and compete with them and a number of studies focused on development of the behaviors. We must use various cues to infer other’s intention which is not known to us. We have no way to know intention directly but we receive many signals to infer it. As visual signals for intention inference, even children and infants can use agency [1], gesture [2], gaze [3] to infer other’s intention. Also, previous studies revealed verbal markers were utilized as auditory signals of others’ intention [4,5].

In most of previous studies, auditory signals have syntactical structures. We might consider non-syntactical but stylized utterances such as interjections. Interjections are defined as non-elliptical signals which communicate speaker’s current mental states and they have no lexical items which are homophonic and semantically related with the interjections [6,7,8].

If interjections actually deliver speaker’s mental states including intention, we might consider what factors of interjection work on inference of others’ mind. These utterances might have no linguistic meaning but paralinguistic one [9].

In experiment 1 (in adults) and 2 (in children), I examined what interjections were used as cues for intention inference. In experiment 3 (in adults) and 4 (in children), I examined what factor of interjections were used as cues for intention inference.

Experiment 1

Thirty-six native Japanese students (20 females, mean age = 20.78) of Kyushu University took part in this experiment. I confirmed that they were raised primarily in Japanese-spoken environment.

Stimuli were the video clips in which an agent collided against an object. At the moment of the collision, one of 6 interjections, which were selected from Japanese vocabularies (“ach”, “ech”, “oh”, “ei”, “yah”, “toh”), or silence as a baseline, was presented. Each participant was randomly assigned to one of 3 stimulus patterns. After the each stimulus presentation, the participants were asked to answer 2 questions; “To what extent do you think intentionally the agent collided against the object?” and “To what extent do you think accidentally the agent collided the object?” (reverse code), using 7-point scale. The sum of the scores of 2 questions represented the participant’s response in that particular trial.

A two-way mixed-participant ANOVA with stimuli pattern and interjection revealed significant main effect of interjection ($F(4.16, 137.25) = 33.72, p < .001, \eta_p^2 = .51$). Either the main effect of pattern or the interaction between pattern and interjection was not significant (pattern: $F(2, 33) = 1.40, p = .26$; pattern and interjection: $F(8.32, 137.25) = 1.29, p = .25$). Multiple comparisons using Shaffer’s method revealed that “ach”, “ech” and “oh” were scored as less intentional than “ei”, “yah” and “toh”. Also, “ah” and “oh” were scored less intentional, and “ei” and “yah” were scored more intentional, intentional than in the silence condition (all $ps < .01$, Figure 1).
Experiment 2

Thirty-four Japanese-speaking preschoolers who go to the kindergarten in Fukuoka, Japan, took part in this experiment as the final sample (17 female, mean age = 67.38 months, range: 55 - 78 months).

Stimuli were the video clips in which a model appeared with a bear toy and built blocks. In collapse condition, she had the bear toy hit and collapse blocks. In jump condition, she had the bear toy jump over the blocks without collapsing. At the moment of collapse or jump, one of 2 interjection (“ach” and “ei”), which were selected based on the results of experiment 1. The participants of group 1 observed collapse condition with “ach” utterance and jump condition with “ei” utterance. The participants of group 2 observed collapse condition with “ei” utterance and jump condition with “ach” utterance. After video clips, participants were asked to imitate model’s behavior.

Fisher's exact test significantly associated group of participants with imitation (p = .039). The most of participants in group 1 imitated jump action, whereas the most of participants in group 2 imitate collapse action. In the other words, children imitated models' behaviors which were demonstrated with “ei” utterance (Table 1).

<table>
<thead>
<tr>
<th>Table 1. The number of children's choice in each group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 (Collapse act with “ei”)</td>
</tr>
<tr>
<td>Group 2 (Jump act with “ei”)</td>
</tr>
</tbody>
</table>

Experiment 3

Thirty native Japanese students (20 females, mean age = 19.77 years old) of Kyushu University took part in this experiment. I confirmed that they were raised primary in Japanese-spoken environment.

To examine what factor of interjection affect inference of intention for adults, intact and controlled interjections, produced by Praat ver.5.4.19 (Boersma & Weenink, 1992-2015), were used (Figure 2).

Stimuli were the video clips in which 2 female models did two kinds of task demonstrated. One task was ball and block task. In collapse condition, she had the ball hit and collapse blocks. In jump condition, she had the ball jump over the blocks without collapsing. The other task was ring and board task. The cork board had a perch on the upper part and a box on the lower part. In hang condition, she hung the bracelet on the perch. In drop condition, she dropped the bracelet into the box. At the moment of each action, one of 2 interjection (“ach” or “ei”) was presented.

All participants saw eight video clips. In the first half of the experiment, they observed video clips in which actions were differently pairing with controlled interjections such like experiment 2. In the last half, they observed video clips in which intact interjection were used instead of controlled ones. After they observed each video clip, they answered three questions. First and second questions were “to what extent do you think intentionally the model do the action?” for each action using 7-point scale. As intention scores of each interjection, the sums of the 2 tasks represented the participant's response about extent of model's intention of action with each interjection. Last question was “when you ask to imitate the action, which action do you imitate?” using forced two choice question, and a choice of action with “ei” was coded as a point. The sum of the 2 tasks was recorded as the participant's imitation score.

A two-way within-participant ANOVA with interjection and prosodic control revealed the main effect of interjection and interaction of interjection and paralinguistic control were significant (interjection:
The simple effect test for the interaction revealed that the actions with "ei" were regarded as more intentional ones in both controlled and intact conditions (controlled condition: $F(1,29) = 8.83, p = .006$; intact condition: $F(1,29) = 20.01, p < .001$). Furthermore, it also revealed that the actions with intact "ach" utterances were regarded as less intentional action than the ones with controlled "ach" utterance ($F(1,29) = 4.33, p = .046$). To estimate imitation score, a one sample t-test compared with chance level (1 point) revealed that actions with "ei" utterance were significantly chosen in both controlled and intact conditions (controlled conditions: $t(29) = 5.46, p < .001, d = .997$; intact conditions: $t(29) = 5.46, p < .001, d = .997$, Figure 4).

**Discussion**

The current study revealed that native Japanese speakers of adults evaluated interjections “ach”, “ech”, or “oh” as less intentional than “ei”, “yah”, or “toh”. Also, Japanese preschooler considered “ei” as intentional cue. Furthermore, even if paralinguistic information of interjection was lost, they identified intentional and accidental acts.

There are 3 possibilities to explain information of interjections which have adults and children differently interpret the actions in term of intention. First, paralinguistic information, such as pitch, intonation, length and intensity, might work on intention inference [10]. The current study showed that this possibility is partly accurate. However, to imitate the action, Japanese speakers selected the imitate-worthy action in controlled interjection condition same as in intact interjection condition. Also, even children imitated the action with "ei" when paralinguistic information of interjection was controlled.

Second, sound symbol of interjection might be affect the intention inference and this is called as bouba/kiki phenomenon [11]. Such like the phenomenon,
some vowels included in interjections might be correlated with others’ intention.

Finally, semantic meaning of interjection might be accessible. In the current study, we used surprising cries and powerful shouts as interjections, which are often used in Japanese. If interjections themselves have the specific linguistic meanings, comprehension of the meanings might enable us to infer others’ mental states [7].

There are two ways to reveal these possibilities. One is cross-linguistic research. We should examine how non-Japanese speakers infer the others’ intentions in the tasks of the current study and how Japanese speakers infer the intentions in non-Japanese interjection tasks. The other is to reveal how infants infer the intentions when they are demonstrated interjections. If infants in pre-verbal period understand the meanings of interjection, comprehension of interjections might not have linguistic processes. These points are future prospects.

In conclusion, the current study suggested that Japanese interjections are available to infer presence or absence of intention for Japanese speakers of adults and children. To judge whether others do intentionally, we might not have to think the meanings of long sentences. We might be able to understand their mental states by listening their short utterance, interjections. And, paralinguistic information of interjection might not always have necessity. Considering the short non-syntactic utterances communicated speakers’ intentions, the current study might contribute to revealing the origin of mankind’s communication tools.

References