Inferring Verb Meanings from Syntactic Frames by Japanese Two-Year-Old Children: An Experimental Approach Using an IPL Paradigm with a Dialogue Phase

TESSEI KOBAYASHI*
NTT Communication Science Labs

TAKAAKI SUZUKI
Kyoto Sangyo University

1. Syntactic Bootstrapping

Syntactic bootstrapping is the proposal that children make use of syntactic information to learn verb meanings (Gleitman 1990). Learning verb meanings is a complex and challenging task for children. Unlike nouns, verbs refer to ever-changing events, and they often refer to only particular aspects of the events, such as actions, manners, and states. At the onset of verb learning, children need help to focus on the particular aspects that the verb describes, and syntactic bootstrapping helps children in this regard. For example, the meaning of the verb *gorp* in the sentence *The duck is gorping the bunny* is inferable to some extent. Because this sentence consists of an NP, a V, and an NP, the verb is considered transitive, and its meaning is likely to denote a certain kind of causation. Although it is impossible to determine the precise meaning of the verb, syntactic frames act as a kind of "zoom

* We thank Kaori Yuruki and Tomoko Kawaguchi for their assistance with data collection and coding, and Miki Terasaka for drawing illustrations used in this paper.

Japanese/Korean Linguistics 22.
Edited by Mikio Giriko, Kyoko Kanzaki, Naonori Nagaya and Akiko Takemura.

lens" in directing the listener's attention to certain aspects of the event (Fisher, Gleitman and Gleitman 1991; Gleitman 1990).

A seminal study on syntactic bootstrapping in young children was carried out by Naigles (1990), who used an intermodal preferential looking (IPL) paradigm to demonstrate that English-speaking 25-month-old children are able to use sentence frames to infer a causative/non-causative distinction. Her experiment consisted of two phases. In the training phase, children saw a combination of two different actions performed by two animal entities, one causative (a duck forced a rabbit into a bending position), and the other non-causative (a duck and a rabbit flexed their arms simultaneously). The children in the transitive condition heard a novel verb in the transitive frame (*Look, the duck is gorping the bunny!*), and those in the intransitive condition heard a novel verb in the intransitive frame (*Look, the duck and the bunny are gorping!*). In the following test phase, the causative and non-causative action scenes were presented separately side by side while the children were asked which action was denoted by the novel verb (*Which is gorping?*). Naigles found that the children looked preferentially at the causative scene for the transitive frame (NP-V-NP) and preferentially at the non-causative scene for the intransitive frame (NP-V), providing evidence that English-speaking 2-year-olds can use syntactic frames to infer verb meanings in terms of a causative/non-causative distinction. The original finding has been replicated in English using different methods (e.g., Fisher 1996) and testing different frames (e.g., Naigles 1996).

2. Current Issues

To better understand the nature of syntactic bootstrapping, recent investigations have addressed certain important issues. One is whether syntactic bootstrapping occurs in argument-drop languages. Previous studies on syntactic bootstrapping have been largely limited to English, in which syntactic frames are highly informative about a verb's transitivity because argument NPs must be overtly expressed. On the other hand, in an argument-drop language like Japanese, syntactic frames may not be a robust cue to transitivity. As the following examples show, argument NPs are frequently dropped as long as their information is recoverable from the context or shared by the speaker and the hearer. The transitive verb *homeru* 'praise' can be used in the following ways.

(1) Kyooko ga Saori o hometa.
Kyoko NOM Saori ACC praised
'Kyoko praised Saori.'

(2) Kyooko ga hometa.
Kyoko NOM praised

(3) Saori o hometa.
Saori ACC praised

(4) Hometa.
praised

Both a subject marked with nominative *ga* and a direct object marked with accusative *o* appear in (1), only a subject appears in (2), only a direct object appears in (3), and both the subject and direct object are omitted from (4), all of which are grammatical in Japanese. These patterns suggest that the syntactic frames used for English-type languages, such as NP-V-NP and NP-V, may not work well for identifying the transitivity of a verb. Thus, children learning argument-drop languages may not be able to infer verb meanings from syntactic frames (e.g., Rispoli 1995; Suzuki 2002).

To investigate this issue, Matsuo et al. (2012) examined Japanese-speaking 28-month-old children by adapting Naigle's method to Japanese. Using the same visual stimuli and procedure, this study tested three sentence patterns: the intransitive frame (*Ahiru-san to usagi-san ga neketteru yo* 'The duck and the rabbit are X-ing'), the transitive frame (*Ahiru-san ga usagi-san o neketteru yo* 'The duck is X-ing the rabbit'), and the transitive frame without case markers (*Ahiru-san usagi-san neketteru yo* 'The duck is X-ing the rabbit'). The results showed that Japanese-speaking 2-year-olds associated the sentences provided in the transitive frame with causative events, but the other two types of sentences were not associated with any particular events. On the basis of these results, Matsuo et al. proposed that Japanese 2-year-olds utilize the combination of the number of overt arguments and case markers to infer the meaning of a new verb.

The data from other argument-drop languages are also informative. Göksun, Küntay and Naigles (2008) and Lee and Naigles (2008) examined Turkish-speaking 2- to 5-year-olds and Mandarin-speaking 2- and 3-year-olds, respectively. In an act-out comprehension task, the children were given intransitive verbs in a transitive frame and transitive verbs in an intransitive frame, and then were asked to act-out the scenes by using puppets and props. The results showed that in both languages the children followed the syntactic frames for the interpretation of the sentences, and ignored the verb's original transitivity. These findings indicate early sensitivity to syntactic frames in children learning argument-drop languages, although it is impossible to make a direct comparison with the original IPL studies in English (Naigles 1990) because the Turkish and Mandarin studies used fa-

miliar verbs in their act-out tasks, and their participants included relatively older children.

Another important issue in recent research on syntactic bootstrapping is concerned with the availability of visual cues. Although traditional studies on syntactic bootstrapping have shown action events visually in the training phase when providing target sentences (e.g., Naigles 1990), such a procedure is unnecessary. The basic idea of syntactic bootstrapping is that syntactic information itself is informative about verb meanings (Gleitman 1990; Landau and Gleitman 1985), and so children do not need visual information in order to guess the gross meanings of a new verb. Arunachalam and Waxman (2010), and Yuan and Fisher (2009) have succeeded in observing children's syntactic bootstrapping in English by providing syntactic information alone in the absence of visual cues. In their "dialogue method," children did not see an action event but instead saw a two-woman dialogue scene in the training phase. In the dialogue, two women used a novel verb embedded in a transitive or intransitive frame. Then, causative and non-causative action scenes were simultaneously presented in the test phase. The results of both studies showed that English-speaking 27- to 28-month-old toddlers looked longer or pointed more often at the causative action in the transitive frame, providing evidence that children can infer verb meanings from syntactic information alone in the absence of visual cues.

3. Aims of the Present Study

In the present study, we investigated whether syntactic bootstrapping occurs in children learning Japanese by using a dialogue method in an IPL paradigm. A critical issue here is to determine whether the children can infer verb meanings from syntactic information alone in the absence of visual cues. The syntactic information includes morphological or morpho-syntactic properties such as case markers, and the position and the number of argument NPs. In this study, we use the term syntactic bootstrapping to refer to the utilization of these aspects of the language. Our participants were given two-woman dialogue-clips including a novel verb embedded in either transitive or intransitive frames and asked which of the events shown in the two adjacent scenes (causative scene vs. non-causative scene) indicated the meaning of the novel verb. We predicted that if the children infer the meaning of the novel verb from syntactic information alone during the training phase, those hearing the transitive frame should look longer at the causative scene and those hearing the intransitive frame should look longer at the non-causative scene in the test phase.

4. Experiment

4.1. Participants

Thirty-two Japanese-speaking children (16 boys and 16 girls) with a mean age of 27.5 months (range: 27.1 to 28.7 months) completed this study. Four additional children were tested but excluded from the present analysis due to fussiness. All the children grew up in a monolingual environment where they learned Japanese as their native language. They were recruited from the Kyoto area. Parents completed the Japanese version of the MacArthur-Bates Communicative Development Inventory (CDI): Words and Grammar (Watamaki and Ogura 2004). The mean productive vocabulary was 364.3 words (range: 72–621).

4.2. Apparatus and Stimuli

The children sat on their mothers' laps in a dimly lit testing room (250 × 250 cm) facing a 40-in. monitor about 150 cm away. Visual stimuli were presented on the monitor using Microsoft PowerPoint installed on a Macintosh G4 computer. Audio stimuli were presented through a speaker located under the monitor behind black curtains. A digital video camera was placed behind a small opening in the black curtains, and about 3 cm below the monitor. The camera was connected to a 20-in. monitor and a digital video recorder to record the direction of the children's eye fixation and to allow an experimenter in an adjacent room to observe the children's behaviors.

Visual stimuli were created for the dialogue scene in which two women were engaged in conversation with natural gestures, and for the causative and non-causative scenes in which a man and a woman were engaged in novel actions. The audio stimuli for the action scenes were produced and recorded by a female native Japanese speaker in a soundproof booth, and then synchronized with the visual stimuli using movie-editing software.

4.3. Design and Procedure

Sixteen children were randomly assigned to a Transitive condition and sixteen to an Intransitive condition. Within each condition, they were randomly assigned one of two novel verbs (*nemaru* or *wageru*). A session consisted of a dialogue phase, a baseline phase, and a test phase. All the children completed two consecutive sessions.

4.3.1. Dialogue Phase

In the dialogue phase, the children saw a color video clip of two women engaged in conversation (Figure 1). The dialogue video appeared in the center of the monitor (size: 26.1 cm × 46.3 cm). While the children in the Transitive condition heard the novel verb embedded in transitive sentence frames (5), those in the Intransitive condition heard the novel verb embed-

ded in conjoined-subject intransitive sentence frames (6). Each trial consisted of a 17-second dialogue including the novel verb in four sentences (Table 1). The children saw the same video clip three times in succession. At the end of each trial, the scene was blank for 3 seconds before the next trial began.

(5) Tomo-kun ga Yuu-chan o nemat-teiru/waget-teiru yo.
Tomo-Mr. NOM Yu-Ms. ACC verb-PROG PCL
'Tomo is X-ing Yu.'

(6) Tomo-kun to Yuu-chan ga nemat-teiru/waget-teiru yo.
Tomo-Mr. and Yu-Ms. NOM verb-PROG PCL
'Tomo and Yu are X-ing.'

Table 1. Two-Woman Dialogues Including Novel Verbs in the Transitive Condition and the Intransitive Condition

Transitive Condition (N = 16)	Intransitive Condition (N = 16)
A: *Nee nee.* *Tomo-kun ga Yuu-chan o nematteiru yo.* 'Hey. Tomo is X-ing Yu.'	A: *Nee nee.* *Tomo-kun to Yuu-chan ga nematteiru yo.* 'Hey. Tomo and Yu are X-ing.'
B: *E? Dare ga dare o nematteiru no?* 'Huh? Who is X-ing whom?'	B: *E? Dare to dare ga nematteiru no?* 'Huh? Who are X-ing?'
A: *Tomo-kun ga Yuu-chan o nemattan yo.* 'Tomo is X-ing Yu.'	A: *Tomo-kun to Yuu-chan ga nemattan yo.* 'Tomo and Yu are X-ing.'
B: *Honma?* *Tomo-kun ga Yuu-chan o nemattan ya.* *Sugoi ne.* 'Really? Tomo is X-ing Yu. Good.'	B: *Honma?* *Tomo-kun to Yuu-chan ga nemattan ya.* *Sugoi ne.* 'Really? Tomo and Yu are X-ing. Good.'
A: *Un, sugoi ne.* 'Yeah, I think so.'	A: *Un, sugoi ne.* 'Yeah, I think so.'

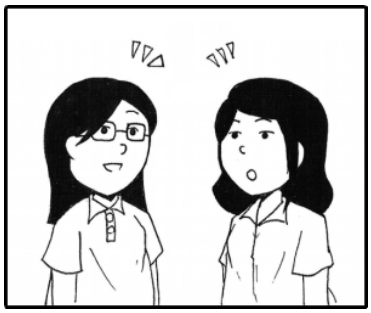

Figure 1. Illustration of Movie Clip Used in Dialogue Phase

4.3.2. Baseline Phase

Following the dialogue phase, the children were given two baseline trials. While the children in both conditions were watching two side-by-side video clips, they heard an arousal phrase such as *Mite-mite!* ('Look!'). One video clip showed a causative scene in which a man spun a woman around on a chair. The other showed a synchronous non-causative scene in which the same man and woman each waved a hand in circles (Figure 2). The size of each clip was 39.2 cm wide by 22.1 cm high, and the two clips were shown with a space of 10 cm between them. We chose this action pair because Arunachalam and Waxman (2010) used the same action pair to provide clear evidence for inferring verb meanings from syntactic frames alone.

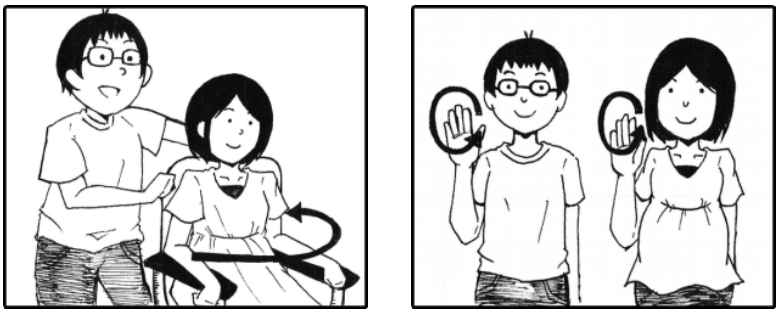

Figure 2. Movie Clip of Causative Action (Left) and Non-causative Action (Right) Used for Baseline Phase and Test Phase

4.3.3. Test Phase

The children were given two test trials in which they watched two side-by-side clips while hearing a sentence including the novel verb without syntactic frames (7).

(7) Nemat-/waget-teiru no dotchi
VERB-PROG NML which
'Which is X-ing?'

We predicted that if the children were able to infer the meaning of the novel verb that they heard during the dialogue phase, the children in the Transitive condition would look longer at the causative scene than at the non-causative scene in the test phase. In contrast, the children in the Intransitive condition would look longer at the non-causative scene than at the causative scene.

Each trial in the baseline and the test trial lasted 12 seconds. At the end of each trial, the scene became blank for 1.5 seconds before the next trial began. The left-right position of each video clip was randomized in a predetermined order.

4.4. Data Analysis

A coder who was not informed of the experimental design and condition performed an offline frame-by-frame coding (30 frames/second) to calculate how long the children looked at each scene (left and right) during the baseline and the test trials from a video recording. To assess data reliability, a different coder did offline coding of a random selection of about 20% of the primary sample. The inter-coder correlations were high (mean = .97). We calculated the proportion of time spent looking at the causative scene in each trial.

5. Results

The time spent looking at the causative scene in the test phase was analyzed with a mixed-design ANOVA with condition (Transitive vs. Intransitive) and word type (*nemaru* vs. *wageru*) as between-subjects factors, and test phase (baseline vs. test) and trial block (first vs. second) as within-subjects factors. The analysis revealed a significant main effect of test phase, F (1, 28) = 14.978, p = .001, η_p^2 = .349, and a significant interaction between test phase and condition, F (1, 28) = 24.788, p < .001, η_p^2 = .469. No other main effects or interactions were significant.

Further analysis with Bonferroni correction indicated that the children in the Transitive condition looked significantly longer at the causative scene in the test trials (M = 66.1%, SD 11.2) than in the baseline (M = 56.4%, SD 10.6), F (1, 28) = 9.864, p = .004, η_p^2 = .261. In contrast, those in the Intransitive condition looked significantly shorter at the causative scene in the test trials (M = 45.5%, SD 14.8) than in the baseline trials (M = 57.5%, SD 13.7), F (1, 28) = 15.201, p = .001, η_p^2 = .352, indicating that they looked preferentially at the non-causative scene in the test trials than in the baseline trials. These results are shown in Figure 3.

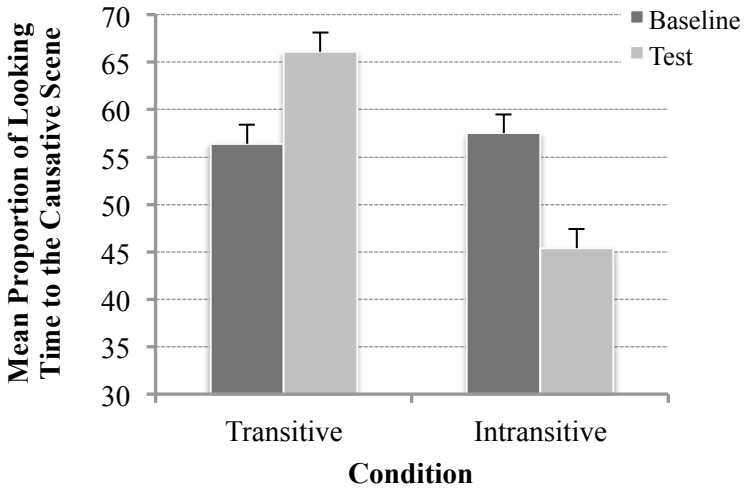

Figure 3. Mean Proportion of Looking Time to Causative Scene in Baseline Phase and Test Phase in Transitive Condition and Intransitive Condition (Error bars show standard errors.)

6. Discussion

Using an IPL paradigm that adopts a dialogue phase similar to that reported by Yuan and Fisher (2009), the present study demonstrates that the Japanese-speaking 27-month-old children hearing transitive frames during the dialogue phase looked preferentially at a causative scene over a non-causative scene, while those hearing the intransitive frames looked preferentially at a non-causative scene over a causative scene. These results reveal syntactic bootstrapping in Japanese. While previous studies using act-out tasks had reported children's sensitivity to syntactic frames in argument-drop languages such as Turkish and Mandarin (Göksun et al. 2008; Lee and Naigles 2008), the present study, using the IPL paradigm, provides additional support for Matsuo et al. (2012), which suggested that syntactic bootstrapping works in children learning Japanese as a first language.

Moreover, by using the dialogue method, we have demonstrated that Japanese-speaking children can infer verb meanings from syntactic information alone in the absence of visual cues. This finding, which parallels English studies (Arunachalam and Waxman 2010; Yuan and Fisher 2009), supports Gleitman's original idea that syntactic information is independently informative in early verb learning (Gleitman 1990), and extends this proposal to argument-drop languages. Our results indicate that Japanese-speaking children can associate causative meaning with the transitive frame and non-causative meaning with the intransitive frame in Japanese, which is

certainly consistent with the results of Naigle's original experimental study in English.

Interestingly, however, these observations are not necessarily consistent with the results of a number of previous studies. English-speaking children often failed to show their preference for a non-causative scene for intransitive frames with the traditional procedure (Hirsh-Pasek, Golinkoff and Naigles 1996; Naigles and Kako 1993) as well as with the dialogue method (Arunachalam and Waxman 2010; Yuan and Fisher 2009). Note that this is also true for the Japanese study reported by Matsuo et al. (2012). Some researchers claim that children do not associate an intransitive frame with a non-causative scene because intransitive verbs do not necessarily denote non-causative events. For example, *play* and *fight*, as in *John and Bill played/fought together*, can be a reciprocal action involving a causative meaning by the two participants (e.g., Arunachalam and Waxman 2010; Noble, Rowland and Pine 2011). However, this line of reasoning cannot explain why only some studies observed children's preference for a non-causative scene with an intransitive frame (Kidd, Bavin and Rhodes 2001; Naigles 1990). We must wait for future studies to resolve this issue, but it is important to note that this fact in itself does not undermine the possibility of syntactic bootstrapping, because all the previous studies mentioned above observed that children figured out the distinction between transitive and intransitive frames for causative/non-causative events.

7. Concluding Remarks

In summary, we suggest that Japanese 27-month-olds can make inferences about the meanings of novel verbs from syntactic information alone in the absence of visual cues about referents. This is the first evidence to show that syntactic bootstrapping works in young children learning an argument-drop language under highly controlled experimental conditions.

A further question arises as to what exactly the syntactic cues are for Japanese-speaking children. As with English, the number and the position of argument NPs may be crucial, but Matsuo et al.'s study provides evidence against this possibility. Their experimental sentences included a sentence pattern involving two NPs without case markers, and they found that this type of the sentences was not associated with either a causative or non-causative event. This leaves us with the idea that the existence of case markers is essential. To ascertain whether case markers alone can be robust cues, we need to look at single-argument sentences, namely, sentences involving only one argument with a case marker and a verb (e.g., NP NOM V / NP ACC V) (e.g., Suzuki 2007). This remains a topic for further research. Alternatively, it may be true that both the number of argument NPs and case

markers provide the essential information required for syntactic bootstrapping in Japanese.

References

Arunachalam, S. and Waxman, S. R. 2010. Meaning from Syntax: Evidence from 2-year-olds. *Cognition* 114:442–446.

Fisher, C. 1996. Structural Limits on Verb Mapping: The Role of Analogy in Children's Interpretations of Sentences. *Cognitive Psychology* 31:41–81.

Fisher, C., Gleitman, H. and Gleitman, L. 1991. On the Semantic Content of Subcategorization Frames. *Cognitive Psychology* 23:331–392.

Gleitman, L. 1990. The Structural Sources of Verb Meanings. *Language Acquisition* 1:3–55.

Göksun, T., Kuntay, A. C. and Naigles, L. R. 2008. Turkish Children Use Morphosyntactic Bootstrapping in Interpreting Verb Meaning. *Journal of Child Language* 35:291–323.

Hirsh-Pasek, K., Golinkoff, R. and Naigles, L. 1996. Young Children's Use of Syntactic Frames to Derive Meaning. *The Origins of Grammar*, ed. K. Hirsh-Pasek and R. Golinkoff, 123–158. Cambridge: MIT Press.

Kidd, E., Bavin, E. L. and Rhodes, B. 2001. Two-year-olds' Knowledge of Verbs and Argument Structures. *Research on Child Language Acquisition: Proceedings of the 8th Conference of the International Association for the Study of Child Language*, ed. M. Almgren, A. Barrena, M.-J. Ezeiz-abarrena, I. Idiazabal, and B. MacWhinney, 1368–1382. Somerville: Cascadilla Press.

Landau, B. and Gleitman, L. 1985. *Language and Experience: Evidence from the Blind Child*. Cambridge: Harvard University Press.

Lee, J. N. and Naigles, L. R. 2008. Mandarin Learners Use Syntactic Bootstrapping in Verb Acquisition. *Cognition* 106:1028–1037.

Matsuo, A., Kita, S., Shinya, Y., Wood, G. C. and Naigles, L. 2012. Japanese Two-year-olds Use Morphosyntax to Learn Novel Verb Meanings. *Journal of Child Language* 39:637–663.

Naigles, L. R. 1990. Children Use Syntax to Learn Verb Meanings. *Journal of Child Language* 17:357–374.

Naigles, L. R. 1996. The Use of Multiple Frames in Verb Learning via Syntactic Bootstrapping. *Cognition* 58:221–251.

Naigles, L. R. and Kako, E. T. 1993. First Contact in Verb Acquisition: Defining a Role for Syntax. *Cognitive Development* 64:1665–1687.

Noble, C. H., Rowland, C. F. and Pine, J. M. 2011. Comprehension of Argument Structure and Semantic Roles: Evidence from English-Learning Children and the Forced-Choice Pointing Paradigm. *Cognitive Science* 35:963–982.

Rispoli, M. 1995. Missing Arguments and the Acquisition of Predicate Meanings. *Beyond Names for Things: Young Children's Acquisition of Verbs*, ed. M. Tomasello and W. E. Merriman, 331–352. Hillsdale: Lawrence Erlbaum.

Suzuki, T. 2002. Bootstrapping from Case-marking Morphology in Japanese. *Studies in Language Sciences 2*, ed. Y. Shirai, H. Kobayashi, S. Miyata, K. Nakamura, T. Ogura, and H. Sirai, 157–174. Tokyo: Kuroshio.

Suzuki, T. 2007. Tan'itsu-kōbun no rikai kara saguru yōji no kakujosi rikai (The Development of Japanese Case Markers Observed through Children's Comprehension of Single-Argument Sentences). *Gengo Kenkyū* 132: 55–76.

Watamaki, T. and Ogura, T. 2004. *The Japanese MacArthur-Bates Communicative Development Inventory: Words and Grammar*. Kyoto: Kyoto International Social Welfare Exchange Center.

Yuan, S. and Fisher, C. 2009. "Really?" She Blicked the Baby?: Two-year-olds Learn Combinational Facts about Verbs by Listening. *Psychological Science* 20: 619–626.