Temporal Contingency Augments Attention to a Referent in a Word Learning Task

Sho Tsuji, Reiko Mazuka, and Daniel Swingley

1. Introduction

The interactive nature of social engagement and the resulting *contingency* between a child's actions and the interaction partner's reaction are considered to support early language acquisition. Contingent responsiveness has been demonstrated to lead to better learning outcomes in various measures of language, for instance number of vocalizations (Goldstein, King, & West, 2003) and phonetic learning (Elsabbagh et al., 2013).

Studies pitting learning from non-contingent video against learning from live, contingent settings illustrate the key role contingency can play in boosting learning from a teacher. For instance, 30-36 months-old toddlers learned novel verbs equally well from a live interaction partner present in the same room or introduced via video chat. If, however, toddlers were presented with the video recording of another toddler's video chat, they did not show evidence of learning (Roseberry, Hirsh-Pasek, & Golinkoff, 2014). Similar results have been obtained in the context of noun learning (Myers, LeWitt, Gallo, & Maselli, 2016), object retrieval (Troseth, Saylor, & Archer, 2006), and imitation tasks (Nielsen, Simcock, & Jenkins, 2008). The contingent interactions in these studies were designed as being socially meaningful, containing features like saying the child's name, referring to a toy that was present in the room, or giving positive feedback. In this context, the fact that toddlers learn in the interactive screen conditions can be interpreted such that toddlers need to perceive a situation (or an interaction partner or a teaching act) as socially meaningful in order to learn.

While these socially meaningful contingencies were critical factors manipulated in the above studies, other studies demonstrate that they are one of several ways that a combination of social cues can make the on-screen situation sufficiently socially meaningful to support learning. The role of social support per se, and not contingency, is shown when children learn words better (from non-contingent screens) if their mothers appear on screen rather than an unknown experimenter (Krcmar, 2010), or if they have observed a reciprocal social interaction on the screen beforehand (O'Doherty, Troseth, Shimpi, Goldenberg, Akhtar, & Saylor, 2011).

---

* Sho Tsuji, Laboratoire de Sciences Cognitives et Psycholinguistique, Dept d'Etudes Cognitives, ENS, PSL University, EHESS, CNRS, tsujish@gmail.com. Reiko Mazuka, RIKEN Center for Brain Sciences. Daniel Swingley, University of Pennsylvania.

© 2019 Sho Tsuji, Reiko Mazuka, and Daniel Swingley. *Proceedings of the 43rd Boston University Conference on Language Development*, ed. Megan M. Brown and Brady Dailey, 693-704. Somerville, MA: Cascadilla Press.

Just as social contexts without clear temporal contingencies appear to facilitate learning, temporal contingency without an ordinary social partner can facilitate gaze following. For instance, twelve-month-old infants are more ready to follow the gaze of a toy animal that previously exhibited contingent responsiveness by beeping and blinking upon infants' gaze than that of a toy animal that beeped and blinked independently of infant behavior (Johnson, Slaughter, Collins, Tyan, & Carey, 1996). Similar behavior was recently observed in eight-month-olds, who followed the turning direction of a 3D object on screen only if that object had previously shown movements contingent on infants' gaze (Deligianni, Senju, Gergely, & Csibra, 2011). Twelve-months-olds have also been shown to attribute goal-directedness to an amodal object that they had previously observed to respond to an experimenter by contingent beeping (Shimizu & Johnson, 2004), but not to its non-contingent counterpart.

One interpretation of these findings has been that contingency, among other cues that are often co-present in social situations, serves as an ostensive cue that signals the communicative intention of an interaction partner (Csibra, Gergely, & Pisupati, 2010). It is assumed that infants and toddlers can readily and automatically use these cues in order to infer communicative intention, and to consequently orient towards the source of this cue. This mechanism is suggested to elicit referential expectation in the infant, facilitating the learning of information that is provided directly after the ostensive cueing. Evidence for expectation of a communication partner's referential act after ostensive cueing comes, for instance, from studies where infants expected to find an object in the location an experimenter had looked at (Csibra & Volein, 2008), or where toddlers were able to find hidden objects based on the experimenter's gaze or pointing cues (Behne, Carpenter, & Tomasello, 2005), after the use of ostensive cues. If the gaze-following elicited by temporally contingent communication partners expresses infants' referential expectations, this expectation might contribute to better word learning after such expression of contingency, even in the absence of human-like communication partners and a rich set of social cues.

To sum up, one group of findings demonstrates how socially enriched, meaningful situations lead to a word learning advantage in toddlers. Although contingency plays a critical role in some of these studies, it is unclear to what extent it would have contributed to this outcome in the absence of the surrounding rich and meaningful remaining social cues. Another group of findings shows that temporal contingency, even if exhibited by a non-human entity in the absence of a rich set of social cues, leads to gaze following behavior. If, as suggested in the context of such studies, temporal contingency triggers referential expectation, by extension, it might also facilitate subsequent *learning*. Alternatively, the gaze following elicited by temporal contingency in previous studies might express increased attention to the object oriented towards without necessarily generating referential expectation.

In the present study, we test these hypotheses by assessing novel word learning by a contingently reacting, non-human-like entity in the absence of a richer set of social cues. We presented toddlers with two temporally contingent "teachers", who were otherwise devoid of social characteristics. Both were selfpropelled 3D-modeled objects without any human features (see Fig. 1). Each toddler saw one block with a contingent teacher, and another with a matched, non-contingent one. They were provided evidence for each teacher's contingent nature during two sets of demonstration trials, one demonstrating the teacher's contingent responsiveness (or absence thereof) towards a third person, and the other its contingent responsiveness (or absence thereof) towards the toddler. Both types of evidence have been suggested to contribute to toddlers' perception of the interaction partner (see Beier & Carey, 2014, for an overview). The teacher then oriented repeatedly towards one of two objects, with one of them being named. Toddlers' recognition of this novel word-object association was then tested.

Our central question was whether toddlers would learn novel word-object associations from these entities, and whether this would differ as a function of their contingent nature. Since gaze-following was a necessary prerequisite for disambiguating the referent of the naming instances, we also asked whether toddlers were able to gaze-follow the entities, and how that would differ as a function of contingency.

2. Methods

Our project and the analyses are preregistered on the OSF (https://osf.io/r7zy4/?view_only=beda20b2c75542ac893e86dd1967deee). If not otherwise noted, we follow the procedures and analyses as preregistered.

2.1. Participants

40 normally developing Japanese-learning 14-17 month-old toddlers (14 14-month-olds, 11 15-month-olds, 7 16-month-olds, and 8 17-months olds, median age = 474 days (range 440-542 days), 14 female) were included in the final sample. Another 34 toddlers were excluded due to not completing a single block (13), completing only one block (12), bad tracking due to excessive movements (2), trial-based exclusion as specified further below (6), or equipment error (1). The study was approved by the institutional review board, and caregivers signed an informed consent form beforehand. They received a monetary compensation of 1,000 Yen in exchange for participation.

2.2. Stimuli

Toddlers were exposed to two blocks. In each block they saw photographs of two novel objects and heard their labels. The four novel objects (Fig. 1) were real inanimate objects, unfamiliar to children and distinct in color and shape. The labels were based on the phonotactically legal Japanese non-words *kippo, monsha, dejo, sappu* [kippo, monɕa, de:ʑo, sappu]. They were constructed as disyllables with heavy-light syllable weight, a structure frequently occurring in words used in Japanese infant-directed speech (Mazuka, Kondo, & Hayashi, 2007). Frequency of occurrence of the constituting syllables as well as of the whole string were matched. The original recordings were made by a female native speaker of Japanese. *Kippo* and *monsha* were always paired in one block and *dejo* and *sappu* in the other block, such that each pair corresponded to one teacher's voice. We strove to make the two non-words within a pair as distinct as possible by using distinct vowels and consonants as well as different pitch patterns. We then modified these recordings as follows in order to achieve a voice quality to match up with our two non-human teachers using Praat (Boersma, & Weenink, 2017) and Audacity (Audacity Team. 2018). We first replaced the first vowel portion of each non-word by selecting 8 pitch pulses of that vowel, fading in the first two pulses and partially fading out the last two pules (40%) to create a rippled effect. We then copied this new sequence as many times as necessary to match the length of the original vowel, and replaced it. We then used the Vocoder Praat add-on (Corretge, 2012) to modify the vocal tract properties of the whole non-words with a formant shift of 1.99. Finally we modified the mean pitch of the first pair of non-words to 500 Hertz, and the mean pitch of the second pair to 400 Hertz, to give the two teachers distinct voices.

In addition, the same speaker also recorded a short, unspecific exclamation, which was used during the greeting and familiarization phases. Again, we replaced the vowel portion by a modified, rippled passage, and modified the vocal tract properties. Finally, we created two versions of the sequence, one corresponding to each teacher, with a mean pitch of 500 Hertz and 400 Hertz. We also inverted the pitch for the 400 Hertz sound. These sounds are henceforth referred to as beeps.

In each block, toddlers saw one 3D entity that functioned as the teacher of the novel word-object association. These entities were also distinct in color and shape (Fig.1). They were created using the free online 3D design tool Tinkercad (https://www.tinkercad.com/). One entity (green/round) was always paired with two distinct objects. One of them was always labeled *kippo* in the naming phase, and the other, only named during the test phase, was labeled *monsha*. The other entity (purple/triangle) was paired with the other two objects. One of them was was labeled *dejo* and was named during the naming phase; the other's label was *sappu*. Which group of stimuli was presented in the contingent block and which was presented in the non-contingent block was counterbalanced across participants.

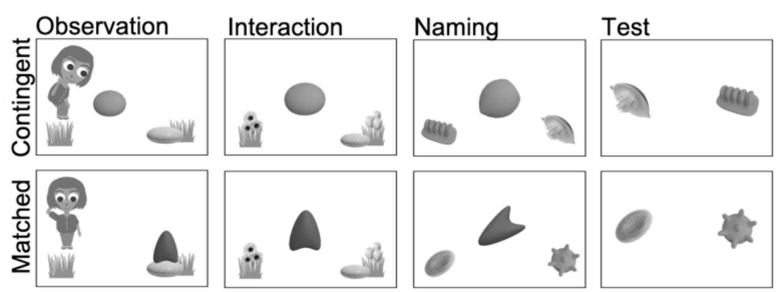

Figure 1. Illustration of experimental phases for each condition, with details provided in the text. Order of conditions and matching of entity/objects to condition were counterbalanced across toddlers.

2.3. Procedure

Toddlers were seated on a caregiver's lap in a sound-attenuated room facing the screen of a Tobii XL eye-tracker. The experimenter was hidden behind a wall in the same room. The caregiver wore headphones with masking music. Infants' gaze was calibrated with Tobii Studio's child-friendly 5-point calibration. Their gaze was recorded with a sampling rate of 120 Hz, and the experiment was administered using E-Prime 2.0.

The experiment consisted of two blocks, one of which contained all trials of the contingent condition, and the other all trials of the matched condition. Block order was counterbalanced across infants. Within each block, toddlers were exposed to one greeting trial, two familiarization trials, eight naming trials, and seven test trials, described next. Each trial was preceded by a fixation cross and started manually once the toddlers fixated the screen.

The *observation trial* functioned to demonstrate either that the entity was communicative (in the contingent block) or not (in the matched block) with regard to a third person. In the contingent block, a girl animation walked into the screen from one side, followed by the entity from the other side. They then engaged in an interactive sequence, where a smile and handwave by the girl was reciprocated by the entity's jumping up and down and emitting the beep sound. This sequence lasted around 30 seconds. In the matched block, the two agents walked in as well, but instead of engaging in an interaction, the girl scratched her head, and the entity jumped up and down on a stone while emitting the sounds. Thus, the number of movements and sounds was matched between conditions, but only one of them demonstrated the entity's contingent reactivity on part of the entity.

The two *interaction trials*, each 15 s long, demonstrated that the entity was also contingently reacting to the toddler (contingent block) or not (matched block). The entity was placed in the center of the screen, with two flower beds to the right and left bottom corner of the screen. In the contingent block, if the toddler gazed at the entity, it would jump up and down and emit the beeping sound, and if the toddler gazed to the right or left the entity would orientation-follow the direction of the toddler's gaze. In the matched block, the entity's movements were yoked based on the movements elicited by one of the previous toddlers' contingent blocks. For the first two toddlers, this was done based on pilot toddlers' data. Thus, the total amount of movements was controlled between blocks.

During the *naming trials*, the entity was again placed in the center of the screen, and the two associated novel objects were placed on the left and right bottom of the screen (side counterbalanced). The entity oriented towards one of the objects, six times towards the target object and two times towards the distractor object. Each time it oriented to target, it first extended towards it once, and its label was then presented twice after a second extension. When it oriented to the distracter it extended only once, with no labeling following.

Finally, during *test trials*, the target and distractor objects were presented side by on the screen (side counterbalanced), and one of them was named. The target was named in four trials, and the distracter in the remaining three.

2.4. Data cleaning and analysis

To clean our data, we first discarded all gaze data that were marked with with a gaze validity code higher than 1 as recommended by the manufacturer (Tobii Technology, Inc, 2016). We then excluded data of individual toddlers that did not accumulate sufficient looking time to screen by trial type, as follows. For the naming trials, we excluded toddlers if their total sum looking time during all naming trials was less than one third, or if the they had less than one third of trials in which they looked more than one third of the time of trial duration. For the test phase, the same criteria applied, but we additionally excluded toddlers who had not reached the above criteria during the familiarization and naming trials, since attention during these was considered crucial for interpreting looking behavior in the test phase. Note that our preregistration initially set this number to 50% instead of 33%, but since this excluded more than 25% of toddlers, we lowered this criterion, following a preregistered decision rule.

We analyzed the data from the naming phase to assess whether toddlers succeeded in gaze-following the entity. In order to see whether gaze-following differed as a function of naming, we separated naming trials into two separate time-windows, one from 400 ms after the entity was fully pointing for the first time until naming started, and one from 400 ms after naming onset until end of trials. We aggregated the total looking times to the 400x400 pixel squares containing the target and distractor images, and analyzed the data with a linear mixed effects model using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2018). Our predictor variables were Block (contingent, matched) and Phase (pointing, naming), and the analysis formula was lmer(Looks ~ Block*Phase + (1 | TrialId) + (Block | Subject)). Block and phase were effect-coded, such that the intercept term was indicative of overall performance against chance, with p-values estimated using normal approximation. The effect of predictor variables was assessed using model comparison.

We then analyzed the test trials to assess word learning. We analyzed the aggregated looking time data in a time-window of a 400-2400 ms after onset of naming; the area of interest were the 500x500 pixel squares containing the target and distractor images. Our predictors were Block (contingent, matched) and Word Type (learned, novel), again effect-coded, and the formula was lmer(Looks ~ Condition * Word type + (Condition |SID)+ (1|Trial)). We additionally conducted a non-preregistered growth curve analysis (GCA) modeled after Mirman (2014) on the same time-window, since visual inspection of the data suggested that aggregating over the whole time window might not have captured all differences present in the dataset. GCA accounts for the dynamic nature of gaze data by not only assessing overall differences in looking times but additionally differences in the shape and latency of the gaze curve. The time course of the word recognition effect was captured with third-order orthogonal polynomials and with fixed effects of condition on all time terms, as well as random effects of participant and trial on all time terms. Data were grouped into 50ms bins, and the analysis formula was lmer(Looks ~ Block * Word Type*(t1+t2+t3) + (1 + t1+t2+t3 | TrialId) +(BlockType+ot1+ot2+ot3| Subject).

2.5. Results

Toddlers gaze-followed the entity significantly overall ($b = 0.044$, se = 0.021, $t = 2.14$, $p = 0.032$), with no difference between blocks [$X^2(1) = 0$, $p = .99$], phase [$X^2(1) = 1.59$, $p = 0.207$], or interaction between the two [$X^2(1) = 0.083$, $p = 0.773$]. Toddlers thus were able to gaze-follow the abstract entities' orientation, passing the prerequisite for learning the novel word-object association. There was, however, no difference in gaze-following between the contingent and matched blocks.

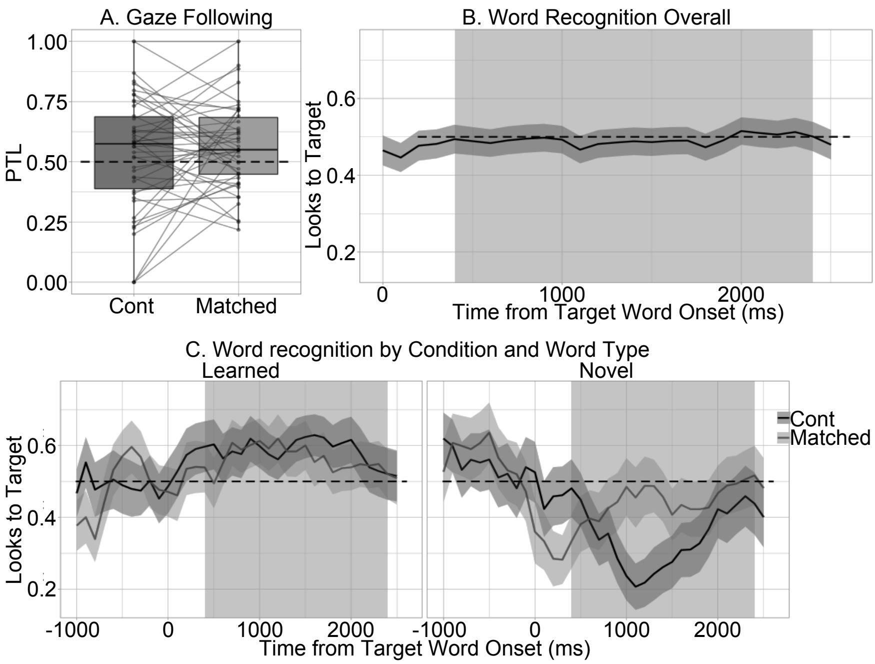

Figure 2. A: Proportion of looks to side the entity oriented towards in naming trials. Points represent individual toddlers' proportion of looks; lines connect points within a toddler. B: Gaze trajectories during test phase, overall. X axis represents time from target word onset, and y axis looks to target (chance is at 0.5). Shaded rectangles represents time window of analysis (400-2400 ms after target word onset), and shaded areas around gaze trajectories represent standard errors of the mean. C: Gaze trajectories during test phase, by Block and Word Type. Cont: Contingent Block; Matched: Matched Block; PTL: Proportion of Target Looks.

With regard to the test phase, the analysis over the whole time window revealed no significant overall word recognition (*b* = 0.002, *se* = 0.027, *t* = 0.086, *p* = 0.931). Model comparison showed no main effect of Block [$X^2(1) = 0.189$, *p* = 0.664], but a main effect of Word Type [$X^2(1) = 10.69$, *p* = 0.001], and no interaction between the two factors [$X^2(1) = 0.470$, *p* = 0.493]. We followed up on the main effect of word type by constructing separate models for learned and novel words, and found significant word recognition for learned words (*b* = 0.071, *se* = 0.036, *t* = 1.985 *p* = 0.047), suggesting that toddlers were able to recognize the words that had been named before. For novel words, we found a significant effect in the opposite direction (*b* = -0.067, *se* = 0.028, *t* = -2.365, p = 0.018). This shows that toddlers were looking at the previously labeled object no matter which label they heard. To verify whether toddlers generally preferred looking at the labeled object regardless of the naming act, or whether they preferred looking at it after hearing any label, we did a post-hoc analysis comparing the mean proportion of looks to the previously labeled object in the 1000 ms time-window before naming onset (*m* = 0.44, *sd* = 0.25) to the mean proportion of looks to this object in the 400-2400 ms time-window after naming (*m* = 0.59, *sd* = 0.23). We found that toddlers significantly increased their proportion of looks to the previously labeled image after hearing either label [$t(37) = 2.97$, *p* = .005], suggesting that this shift in looking behavior was driven by hearing any of the two labels.

A post-hoc growth-curve analysis to better understand this effect showed, in addition to confirm the significant effect of Word Type [$X^2(1) = 192.11$, $p < .001$] but not of Block [$X^2(1) = 0.000$, *p* = < .990], a significant interaction between Block and WordType [$X^2(1) = 17.204$, *p* = < .001] as well as significant effects on the quadratic (*p* = .004) and cubic (*p* = .033) time terms on this interaction effect. We followed up on these analyses by separate models per Word Type. For the learned words, we did not find any significant effect involving Block, while for the novel words, we found significant effects of the *linear* (p = .006) and *quadratic* (p < .001) time terms on Block. These additional results suggest that toddlers' diverging looking behavior to learned versus novel words was affected by block type: While toddlers recognized learned words regardless of block types, in case of novel words, they looked more towards the learned object in the contingent block than in the non-contingent block.

3. Discussion

The present study compared toddlers' gaze following and learning of novel word-object associations from contingently reacting or not contingently reacting non-human like entities in the absence of a rich set of accompanying social cues. Toddlers gaze-followed both types of teachers, but did not learn the words. Exposure to a temporally contingent teacher, however, increased their later attention to the named object. Our findings propose a specific role for communication partners' temporal contingency of in the word learning process, guiding attention to particular elements in complex environments.

In line with previous studies, toddlers gaze-followed the entities' orientation (Deligianni et al., 2011; Johnson et al., 1996). This result is nonetheless not trivial, since our entities had fewer features hinting towards animacy than previous ones, who either had fur and at least one eye-hole (Johnson et al.) or were made up of several distinct shapes with a clear front (Deligianni et al.). The entities' contingency did, however, not influence toddlers' gaze-following behavior. Inspection of the interaction trials suggests that this was not likely due to an unsuccessful contingency manipulation; for instance, toddlers' first look away from the entity was significantly later in the contingent condition (*m* = 1,459 ms, *sd* = 599) compared to the matched condition (*m* = 1,275 ms, sd = 590) [$t(66) = 2.13$, $p = 0.037$]. We discuss alternative interpretations of this lack of difference further below.

Our second outcome of interest was learning of the novel word-object associations as a function of teacher's contingency. We found that toddlers looked more at the previously labeled object, regardless of whether they heard the learned label or a novel label. In other words, they looked at the same object no matter which word they heard. A first conclusion of the present results is that toddlers do not succeed in learning novel word-object associations from a temporally contingent, but otherwise non-social entity. These results are in line with accounts suggesting that contingency accounts for the learning advantage found in previous studies on learning from live or videotaped situation only in combination with a rich social context (e.g. Roseberry et al., 2014).

While toddlers thus did not learn words differently as a function of the teacher's contingency, our post-hoc analyses show that they instead showed increased attention to the labeled object especially after learning from a contingent teacher. Note that both objects had been present on screen for the same amount of time throughout the experiment, and the increased attention paid to this object in the test phase can thus not be explained with differences in exposure. Instead, entities had oriented towards the labeled object more frequently, as well as labeled it, and this difference is likely to drive these differences in looking behavior. The fact that toddlers preferred to look at the previously labeled object especially in the contingent block might suggest that the entity's temporal contingency enhanced toddlers' attention to its focus. Considering that paying attention to the part of the world a communication partner is referring to is a prerequisite for learning, these results thus suggest a specific role of temporal contingency in the learning process, namely of directing attention.

This pattern of results emerged despite the fact that toddlers did not show increased gaze-following in the contingent block. Other studies find a similar dissociation between infants' gaze following and their subsequent learning as a function of the situation's social value. Wu & Kirkham (2010) presented infants with two different displays, in each of which two distinct locations were co-occurring with a distinct sound. In addition to that, one of the locations was cued. In one condition the cue was social, namely the gaze direction of a human face that had previously ostensively cued the infant by smiling, directly gazing and greeting the infant with a phrase. In the other condition it was non-social, meaning a flashing square surrounding the object in the target location. Infants gaze-followed both types of cues equally during the exposure phase, suggesting they were equally successful in guiding attention. In the subsequent test phase, they heard one of the two distinct sounds while the display stayed empty, and their anticipatory gaze was monitored. Those infants who had been socially cued looked mostly at the previously cued location, while those who had been cued by the flashing square did not show a significant difference in their looking behavior to the two locations in which the target image had appeared. Thus, despite comparable gaze following behavior, only infants in the social cueing condition learned the association between location and sound (see also Wu, Tummeltshammer, Gliga, & Kirkham, 2014).

Together with the present results, these findings suggest that infants, despite similarly directing their gaze towards a target, might learn differently depending on context. In particular, only in situations perceived as social would they learn about the referential target of the teacher. While this interpretation aligns with our findings, we nevertheless want to mention another, perhaps less theoretically appealing possibility. Since we modified the recordings of the non-words to make them sound less human, toddlers might not have perceived them as speech, and not found it relevant to distinguish between the learned and novel labels in the test phase. If so, they might simply have associated *any* sound with the learned object, and hence oriented towards it no matter which non-word they heard. While we consider this possibility unlikely, because the sound files were distinct in vowel/consonant quality and pitch pattern, we are nevertheless investigating this possibility in a follow-up study using the unmodified voice of a female speaker.

In sum, the present study demonstrated that a temporally contingent, but otherwise non-human and non-social teacher increases attention to the named object, but not the learning of novel word-object associations, but. The role of temporal contingency during the process of learning from a social communication partner might thus be to guide infants' attention toward relevant elements in a complex environment; a prerequisite, but not a sufficient condition, to augment language learning.

References

Audacity Team (2018). *Audacity(R): Free Audio Editor and Recorder* [Computer application]. Version 2.2.1.0 retrieved December 6th 2017 from https://audacityteam.org/

Bates, Douglas, Maechler, Martin, Bolker, Ben, & Walker, Steve (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01.

Behne, Tany, Carpenter, Malinda, & Tomasello, Michael (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*(6), 492-499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

Beier, Jonathan S., & Carey, Susan (2014). Contingency is not enough: Social context guides third-party attributions of intentional agency. *Developmental Psychology, 50*(3), 889. https://doi.org/10.1037/a0034171

Boersma, Paul, & Weenink, David (2017). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.33, retrieved September 26th 2017 from http://www.praat.org/

Corretge, Ramon (2012). *Praat Vocal Toolkit*. http://www.praatvocaltoolkit.com

Csibra, Gergely, & Gergely, Györgi (2009). Natural Pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. http://doi.org/10.1016/j.tics.2009.01.005

Csibra, Gergely, & Volein, Ágnes (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26*(1), 1-11. https://doi.org/10.1348/026151007X185987

Deligianni, Fani, Senju, Atsushi, Gergely, György, & Csibra, Gergely (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology, 47*(6), 1499–503. http://doi.org/10.1037/a0025659

Elsabbagh, Mayada, Hohenberger, Annette, Campos, Ruth, Van Herwegen, Jo, Serres, Josette, de Schonen, Scania, … Karmiloff-Smith, Annette (2013). Narrowing perceptual sensitivity to the native language in infancy: Exogenous influences on developmental timing. *Behavioral Sciences, 3*(1), 120–132. http://doi.org/10.3390/bs3010120

Goldstein, Michael H., King, Andrew P., & West, Meredith J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America, 100*(13), 8030–5. http://doi.org/10.1073/pnas.1332441100

Johnson, Susan, Slaughter, Virginia, Collins, Kimberly, Tyan, Jean, & Carey, Susan (1996). Whose gaze will infants follow? Features that elicit gaze-following in twelve-month-olds. *Infant Behavior & Development, 19*, 527. http://doi.org/10.1016/S0163-6383(96)90581-0

Krcmar, Marina (2014). Can infants and toddlers learn words from repeat exposure to an infant directed DVD? *Journal of Broadcasting & Electronic Media, 58*(2), 196–214. http://doi.org/10.1080/08838151.2014.906429

Mazuka, Reiko, Kondo, Tadahisa, & Hayashi, Akiko (2007). Japanese mothers' use of specialized vocabulary in infant-directed speech: Infant-directed vocabulary in Japanese. In N. Masataka (ed.), *Origin of Language* (pp.39-58). Tokyo, Japan: Springer.

Mirman, Daniel (2014). *Growth curve analysis and visualization using R.* Boca Raton, FL: Chapman and Hall/CRC Press.

Myers, Lauren J., LeWitt, Rachel B., Gallo, Renee E., & Maselli, Nicole M. (2016). Baby FaceTime: Can toddlers learn from online video chat? *Developmental Science, 20*(4). https://doi.org/10.1111/desc.12430

Nielsen, Mark, Simcock, Gabrielle, & Jenkins, Linda (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science, 11*(5), 722–731. http://doi.org/10.1111/j.1467-7687.2008.00722.x

O'Doherty, Katherine, Troseth, Georgene L., Shimpi, Priya M., Goldenberg, Elizabeth, Akhtar, Nameera, & Saylor, Megan M. (2011). Third-party social interaction and word learning from video. *Child Development, 82*(3), 902-915. https://doi.org/10.1111/j.1467-8624.2011.01579.x

R Core Team (2018). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roseberry, Sarah, Hirsh-Pasek, Kathy, & Golinkoff, Roberta M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development, 85*(3), 956–70. http://doi.org/10.1111/cdev.12166

Shimizu, Y. Alpha, & Johnson, Susan C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science, 7*(4), 425–430. http://doi.org/10.1111/j.1467-7687.2004.00362.x

Tobii Technology, Inc. (2016). *Tobii Studio User's Manual* (version 3.4.5). Retrieved from https://www.tobiipro.com/

Troseth, Georgene L., Saylor, Megan M., & Archer, Allison H. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77*(3), 786-799. http://doi.org/10.1111/j.1467-8624.2006.00903.x

Wu, Rachel, & Kirkham, Natasha Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology, 107*(2), 118-136. http://doi.org/10.1016/j.jecp.2010.04.014

Wu, Rachel, Tummeltshammer, Kristen S., Gliga, Teodora, & Kirkham, Natasha Z. (2014). Ostensive signals support learning from novel attention cues during infancy. *Frontiers in Psychology, 5*, 251. https://doi.org/10.3389/fpsyg.2014.00251

Proceedings of the 43rd annual Boston University Conference on Language Development

edited by Megan M. Brown
and Brady Dailey

Cascadilla Press Somerville, MA 2019

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-096-8 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com