Let's talk action: Infant-directed speech facilitates infants' action learning

Melanie S. Schreiner * [1, 2, 3, 4], Johanna E. van Schaik * + [5, 6], Jelena Sucevic [7], Sabine Hunnius [5], & Marlene Meyer [5, 8]

[1] Psychology of Language Research Group, University of Göttingen, Germany

[2] Leibniz ScienceCampus Primate Cognition, Göttingen, Germany

[3] Clinic for Cognitive Neurology, University of Leipzig, Germany

[4] Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

[5] Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, the Netherlands

[6] Educational Studies, Vrije University, the Netherlands

[7] Department of Experimental Psychology, University of Oxford, UK

[8] Department of Psychology, University of Chicago, USA

* Should be considered joint first author

+ Corresponding author

Address corresponding author: Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands.

Email corresponding author: j.e.van.schaik@vu.nl

Conflict of Interest

The authors declare no conflicts of interest with respect to the authorship or the publication of this article.

Let's talk action: Infant-directed speech facilitates infants' action learning

Abstract

Parents modulate their speech and their actions during infant-directed interactions, and these modulations facilitate infants' language and action learning, respectively. But do these behaviors and their benefits cross these modality boundaries? We investigated mothers' infant-directed speech and actions while they demonstrated the action-effects of four novel objects to their 14-month-old infants. Mothers (*N* = 35) spent the majority of the time either speaking or demonstrating the to-be-learned actions to their infant while hardly talking and acting at the same time. Moreover, mothers' infant-directed speech predicted infants' action learning success beyond the effect of infant-directed actions. Thus, mothers' speech modulations during naturalistic interactions do more than support infants' language learning; they also facilitate infants' action learning, presumably by directing and maintaining infants' attention towards the to-be learned actions.

Keywords: infant-directed speech; infant-directed actions; action learning

Decades of research in the language domain have investigated parents' speech modulations (Ferguson, 1964), and more recently research in the action domain has examined parents' action modulations during infant interactions (Brand, Baldwin, & Ashburn, 2002). Although in daily life parent-infant interactions are typically multimodal, including speech and action, little is known about how these parental modulations interact to affect infants' learning across modalities. Given that infant-directed speech enhances infants' attention to language input (Schreiner, Altvater-Mackensen, & Mani, 2016; Soderstrom, 2007), the present study investigates whether this attentional enhancement can cross over to benefit learning in the action modality. This study thus attempts to bridge the gap between these two lines of infant-directed behavior research by investigating whether mothers' infant-directed speech plays a role in their action demonstrations and, importantly, infants' action learning.

Infant-directed speech

Adults across most cultures tend to exaggerate their speech during natural interactions with infants as compared to conversations with adults (Ferguson, 1964; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2001). In particular, adults modify their prosody in infant-directed speech (hereafter, IDS) by using higher pitch, and a wider pitch range than in adult-directed speech (hereafter, ADS; Fernald et al., 1989; Soderstrom, 2007). There is ample evidence that IDS, relative to ADS, facilitates infants' word segmentation (Thiessen, Hill, & Saffran, 2005), and word learning (Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011), and that the quality of IDS predicts infants' later vocabulary outcomes (Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013).

Nevertheless, the underlying mechanisms of the benefits of IDS on early language learning are still disputed. While earlier work suggested hyperarticulation to be a possible mechanism of IDS (Kuhl et al., 1997), recent research speaks against this assumption, showing that IDS is actually less clearly articulated than ADS (Martin et al., 2015). A different account proposed that the function of IDS is to provide attentional enhancement, that is, it attracts and maintains infants' attention to language input, thereby, facilitating language learning (Schreiner et al., 2016; Soderstrom, 2007). Accordingly, numerous studies have demonstrated that infants prefer IDS over ADS (Cooper & Aslin, 1990; The ManyBabies Consortium, 2020), and that they also select their communicative partners dependent on the speech register they employ (Schachner & Hannon, 2011). Taken together, these findings suggest that within IDS, parents enhance the prosodic characteristics of their speech. This enhancement seems to direct and maintain infants' attention to the relevant speech input in order to boost learning.

Infant-directed action

Caregivers also adjust their action demonstrations for infants compared to adults (Brand et al., 2002). Within these infant-directed actions (hereafter, IDAs), parents tend to perform demonstrations closer to their child, make actions larger, adjust the velocity of their movements, and show the effects of actions for longer (Brand et al., 2002; Rohlfing, Fritsch, Wrede, & Jungmann, 2006; van Schaik, Meyer, van Ham, & Hunnius, 2020). Demonstrations performed in an IDA-manner have been shown to facilitate infants' object exploration and imitation compared to adult-directed demonstrations (Koterba & Iverson, 2009; Williamson & Brand, 2014).

In summary, the existing body of research indicates that parents, on the one hand, modulate their speech, and, on the other hand, modulate their actions, while interacting with their infants. Whereas it has been clearly shown that these modulations facilitate infants' learning within the respective modalities, very few studies investigated the interaction between these domains and potential cross-modal effects.

A multimodal environment: IDS and IDA

Though most studies of parents' infant-directed behaviors are unimodal, typical parent-infant interactions are not. Do infant-directed behaviors within one modality, such as IDS, occur within the context of another modality, like action demonstrations, and does this facilitate learning? Thus far, to our knowledge, no research has been done on the relation between the prosodic characteristics of IDS and action-learning, but the timing of speech during action demonstrations has been examined. Hirsh-Pasek and Golinkoff (1996) proposed that speech timing might help infants to structure observed streams of action, and evidence suggests that infants can indeed use speech to parse action sequences (Brand & Tapscott, 2007). Moreover, parents have been found to naturally use speech to structure their sequential IDAs; in two infant-directed cup-stacking studies, parents enveloped sub-goals of the action sequence with social signals including speech (Nagai & Rohlfing, 2009; Rohlfing et al., 2006). These studies highlight a potential role of speech as structuring, but not overlapping with, action streams, and are in line with the overshadowing hypothesis. This hypothesis posits that upon simultaneous presentation, auditory input tends to grab infants' attention thereby delaying processing (and hence affecting learning) in the visual domain (Robinson & Sloutsky, 2004, 2007). Extending these findings on the basis of IDS research, we propose that the prosody-driven attention enhancing qualities of IDS could likewise also affect action learning. Through exploring the effect of prosodic modulations within IDS on infants' action-learning, we might be able to better explain variations in action learning during naturalistic action demonstrations (e.g., van Schaik et al., 2020) and discover a possibly broader attentional enhancement function of the prosodic modulations of IDS.

Overview of the current study

In our study, we examined whether and to what extent mothers use not only IDAs but also IDS during action demonstrations to their infants and whether mothers' prosodic modulations, over and above action modulations, benefit infants' action learning. To investigate this, we quantified mothers' prosodic and kinematic modulations during naturalistic action demonstrations as well as their 14-month-old infants' action learning. Mothers were asked to demonstrate the actions of four novel objects to their infants and were free to exchange the objects and speak to their infants. The dataset used for the present investigation stems from a previous study investigating the effects of IDA on infants' action learning (van Schaik et al., 2020).

Our first research question was whether mothers use IDS during action demonstrations, and, if so, how much of the time mothers talk, act, or do both simultaneously. Analogous to the attention-grabbing effects of IDS in language learning, we propose that IDS might also be beneficial for drawing attention to the interaction (cf., Schreiner et al., 2016; Soderstrom, 2007) and therefore hypothesize that mothers make use of IDS during action demonstrations. It is an open question whether mothers might either speak or act rather than doing both simultaneously, as one might expect that simultaneously speaking and acting on the objects might impede action learning, following the overshadowing hypothesis (Robinson & Sloutsky, 2007).

Our second research question investigated whether IDS is related to infants' action learning. More specifically, do the quintessential prosodic characteristics of IDS contribute to infants' learning outcomes in action-learning contexts? Given that IDS enhances infants' attention to the relevant speech stimuli (Schreiner et al., 2016), we expected that mothers' IDS influences infants' action learning, with enhanced prosodic characteristics being related to better action learning.

Method

The present analyses utilize data from a previous study (van Schaik et al., 2020) which compared parents' infant- and adult-directed actions using optical motion tracking. In particular, parents were asked to demonstrate the different action-effects of four novel objects to their 14-months-old infants and to two adults. Given the goal of the present investigation was to examine the role of IDS during infants' action learning, we now additionally assessed mothers' IDS during the mother-infant interactions.

Participants

The data of 35 mothers and their 14-month-old Dutch-learning infants was included in this study. On average, infants were 14.3 months old (range: 13.5-15.2 months; 16 girls). The sample size was determined by the previously collected dataset. Post-hoc power analyses were performed using GPower (Faul, Erdfelder, Buchner, & Lang, 2009). Given the R^2 of .28 of the final regression model, the sample size of 35, and an alpha of .05, the resulting power is 0.84.

Five parent-infant dyads from the original data set (van Schaik et al., 2020), were excluded because of insufficient usable audio data (*n*=2) or participation of the father instead of the mother (*n*=3). Only mothers were included in the analyses as males show different prosodic characteristics (i.e., lower pitch height) than females (Fernald et al., 1989). With only three fathers, a separate analysis could not be conducted for the fathers. This research was approved by the Ethics Committee of the Faculty of Social Sciences of Radboud University ("Cognitive research with babies and young children"; ECG2012-1301-006, -006a, -006a2) and was carried out in accordance with the Declaration of Helsinki.

Materials

Objects. Four novel cylindrical objects were used in this study (see Figure 1). The visual appearance of the objects was identical except for their color. Despite looking similar, each object could be operated with a distinct action to produce a unique auditory or visual action-effect: the grey object rattled when shaken, the orange object could be twisted around its radial axis to produce a “kggrr” sound, when the top of the red object was pressed a light turned on, and the two magnetic halves of the yellow object could be pulled apart. The objects were hidden from the infant’s view in the beginning of each demonstration and presented one at a time in a counterbalanced order.

Recordings. Mothers’ movements were recorded with a Qualisys Oqus 5+ optical motion tracking system (www.qualisys.com). The three-dimensional movements of two reflective markers worn on mothers' index fingers were tracked using seven motion cameras and a video camera. Interactions were also recorded using two corner cameras and a ceiling microphone mounted directly above the mother (Noldus Information Technology, Media Recorder, Version 2.5).

Procedure

Mothers were first familiarized with the objects and their specific effects. Subsequently, mothers were asked to demonstrate how the objects worked to an adult partner, then their infant, and then another adult partner. Thereby, mothers were sufficiently familiar with the objects before demonstrating them to their infants.

The demonstration took place with mother and infant seated across from each other at a table. The objects were placed behind a small screen in the right-hand corner on the mothers' side of the table so that mothers could easily access them while also being out of the infants' reach and sight. Mothers were instructed to demonstrate one object at a time following a pre-specified order (counterbalanced across mothers) and to demonstrate the object at least once before passing it to their infant. Mothers were then free to interact naturally with their infant, including speaking to their infant and exchanging the objects back-and-forth as often as they liked. Mothers could move on to the next object at their own pace. The demonstration session ended after the mother had demonstrated and exchanged all four objects. The mean length of the demonstration session was 191 seconds (*SD* = 92).

Measures

Maternal utterances were coded using PRAAT (Boersma & Weenink, 2019), from the mothers' first touch of the first object until her return of the last object. Mothers' action demonstrations were coded using ELAN (2018; Lausberg & Sloetjes, 2009) while they were holding each object. From this, two types of measures were calculated: (A) the percentage of time mothers spent talking, acting or both and (B) the prosodic characteristics of IDS. The IDA characteristics and infants' learning data were obtained from the previous study (van Schaik et al., 2020).

Distribution of time spent on IDS and IDA. The utterances coded in PRAAT were imported into the ELAN files that contained the action coding. This allowed calculation of the percentages of time mothers were speaking, demonstrating actions, or doing both simultaneously. Since mothers could speak both when they were and were not in possession of the objects but could logically only act on the objects while in possession, we distinguished between time speaking when in possession and when not in possession of the object. We then calculated the time mothers spent speaking while not in possession, speaking (not acting) while in possession, acting (not speaking) while in possession, and speaking and acting at the same time while in possession, as a percentage of the summed time they spent speaking and acting. In other words, the percentages indicate the time mothers spent speaking and/or acting relative to the total amount of time mothers were speaking and acting. We calculated these relative percentages instead of, for example, the percentages speaking and acting out of the total interaction time, because we were specifically interested in how mothers divided their infant-directed behavior time into speaking, acting, or both.

IDS. All utterances were analyzed for four pitch characteristics: minimum, maximum, and mean pitch as well as pitch range using PRAAT. Descriptive outcomes of all IDS measures are in Table 1, together with previously reported pitch characteristics of IDS found for German and British English for comparison (Fernald et al., 1989). Variables were converted to z-scores to ensure consistency in the analyses (see IDA calculations below), resulting in one score per IDS measure per mother.

IDA. Four measures were obtained for each action demonstration and were averaged across repetitions of that object: the 3D distance covered, velocity, proximity of mothers' actions, and duration of the action effects (Table 1). Since the actions for the four objects differed with respect to absolute kinematics values, we calculated z-scores per object before averaging across objects to arrive at one score per IDA measure per mother.

Infant learning. Infants' object-directed actions when they received the objects from their mother were coded. Per object, the number of times the infant successfully operated the object to produce the effect was divided by the infant's number of attempts to operate the object. These ratio scores were averaged across the four objects, resulting in one learning ratio per infant.

Data analyses

To address the first research question, pertaining to the distribution of time mothers spent on IDAs and IDS, the percentage of time mothers used either one and both was examined. With respect to the second research question on the contributions of IDA and IDS to action learning, we first employed a principal component analysis (PCA, see Bair et al., 2006) to reduce the number of variables that represent IDS and IDA. Two PCAs were conducted separately, one for the four measures of IDA and one for the four measures of IDS. Next, to predict infants' learning, we first entered the extracted IDA component into a hierarchical linear regression. In the next step, to assess the possible added benefit of IDS on infants' learning, we entered the extracted IDS components. Analyses were performed in IBM SPSS.

Results

Distribution of time spent on IDS and IDA

The calculations showed that, of the total amount of time mothers spent on IDS and/or IDA, they performed both simultaneously only 6.2% of the time on average (see Figure 2). Mothers spent on average 45.5% of the time on IDS while not in possession of the object. While in possession of the objects, mothers spent 19% on IDS and 29.3% on IDA; specifically, 25 of the 35 mothers used more IDA than IDS, whereas seven mothers used more IDS than

IDA, and three mothers devoted equal amounts of time to IDS and IDA (see Supplemental Material, Figure S 1).

Additionally, we explored the possibility that the combination of IDS and IDA was influenced by the type of effect an action produced (i.e., auditory vs. visual). We calculated overlap of IDS and IDA separately for the objects with an auditory and visual effect. Specifically, we calculated the average overlap when mothers were in possession of an object, with the grey and the orange object representing overlap during auditory-effect object demonstrations and the red and yellow object representing overlap during visual-effect object demonstrations[1]. There was a significant difference in the simultaneous use of IDS and IDAs for the objects with an auditory effect relative to the objects with a visual effect, $t(34) = -3.20$, $p = .003$, $d = 0.54$. In particular, mothers simultaneously spoke and acted on average 9.6% of the time when demonstrating auditory-effect objects, while for objects with visual effects, their IDS and IDAs overlapped on average 14.8% of the time (see Figure 3).

IDA, IDS, and infant learning

In order to assess the impact of IDA and particularly of IDS on infants' action learning success, we first conducted PCAs to reduce the number of variables for the IDA and IDS measures. The PCA on the IDA measures resulted in one principal component, and the PCA on IDS measures resulted in two[2]. These explained 47.0% (IDA) and 99.6% (IDS) of the variance. Component 1 of IDA captures features of larger 3D-distance coverage, higher velocity, longer effect durations, and smaller proximity (demonstrations further away from the infant; see Figure 4). Component 1 of IDS reflects extreme values in pitch (high minimum, maximum and mean pitch) with a small pitch range, and component 2 of IDS captures

---

[1] Note that this calculation is different from the overall distribution calculation where speech was also assessed when mothers were not in possession of an object.

[2] The cutoff point of an eigenvalue larger than 1 was used to determine the number of principal components to be extracted (Kaiser, 1960).

exaggeration of maternal speech input highlighting the range in pitch (with high maximum, low minimum pitch, and a wide pitch range; see Figure 5).

In order to estimate the contribution of IDA and IDS to infants' learning, we conducted a stepwise hierarchical linear regression analysis with infants' learning success as dependent variable. Initial examination of Pearson correlations did not reveal significant correlations between the IDA and IDS components (see Table 2). In the first step of the regression, the extracted IDA component was entered to test whether IDA predicts action learning. In the second step, both extracted components of IDS were entered to evaluate whether a model including both IDA and IDS better predicts infants' learning success. For the first step, a significant regression model was found, $F(1, 33) = 4.49$, $p = .042$ with an R^2 of .12. The IDA component 1 was a significant predictor of infants' learning success. The second step of adding the extracted IDS components significantly improved the model, $F_{change}(2, 31) = 3.30$, $p = .049$ with an R^2_{change} of .16. In particular, IDS component 2, but not 1, was a significant predictor of infants' action learning (see Table 3). This final model explained 28% of the variance with IDA contributing at a level of marginal significance and IDS contributing significantly to the model. In summary, this analysis revealed that both IDA and IDS predict infants' learning. Importantly, IDS had a unique contribution to learning, above the effects of IDA.

Discussion

We investigated the presence of IDS during mothers' demonstrations of different novel objects to their 14-month-old infants and whether mothers' IDS contributed to infants' action learning. Despite being engaged in an action demonstration context, mothers spent the greatest percentage of time speaking to their infants without demonstrating an action, on average 64.5%. Mothers spent 29.3% of the time on IDA and these two behaviors, IDA and IDS, overlapped for only 6.2%. This pattern indicates that mothers were typically engaging in only one infant-directed behavior at a time, and, even though the goal of the interaction was to teach infants actions, mothers spent a considerable amount of their time speaking. Furthermore, mothers' IDA significantly predicted infants' action learning, and this model significantly improved when mothers' IDS was added. In other words, mothers' IDS significantly predicted infants' action learning beyond the effect of IDA. This finding is especially interesting considering that infants were learning new actions rather than language.

The presence of mothers' IDS during naturalistic action demonstrations and the predictive power of IDS for infants' action learning are in line with two hypothesized benefits discussed in the literature on IDS, IDA, and multimodal learning cues. First, it can be speculated that the benefit of mothers' IDS is reflective of the attention-enhancing role of IDS. Behavioral studies have shown that infants prefer IDS as compared to ADS (The ManyBabies Consortium, 2020) and neurophysiological studies have demonstrated increased brain activation in response to IDS relative to ADS (Zangl & Mills, 2007), indicating that IDS may attract and regulate infants' attention and arousal. Interestingly, the findings of the present study suggest that these attention effects might be cross-modal, as they enhance learning in the action domain. Additionally, as indicated by the factor loadings of IDS component 2, it is wider pitch range, with low minimum pitch and high maximum pitch rather than higher pitch overall, that significantly predicted infants' learning. This suggests that variation in pitch rather than generally higher pitch values in mothers' speech might be important for attracting infants' attention.

Second, both the limited overlap in mothers' speech and action demonstrations and the benefit of IDS for action learning, suggest that parents might be sensitive to the benefits and drawbacks of using different infant-directed behaviors across modalities. These findings extend those of previous IDA studies in which parents used speech to segment sub-goals in action sequences (Nagai & Rohlfing, 2009; Rohlfing et al., 2006). Here, in the context of single actions, speech was prominent in mothers' interactions, but similarly did not overlap with mothers' actions very often. Additional analyses revealed that this was particularly the case for the auditory-effect toys, where speaking at the same time would make the sound effects more difficult to hear, and therefore less salient.

An unexpected result of the present study was that there was no evidence for a correlation between IDA and IDS. Authors have paralleled IDS characteristics and the effects on language learning with those of IDA and its effects on action learning (Brand et al., 2002). However, this exploration of mothers' naturalistic behaviors did not show evidence for the assumption that mothers who modulate their behavior more in one modality also do so in the other. Although interpretation of a null result should be taken with caution, one explanation for this outcome could be due to the level of analysis, grouped over entire interaction sessions, which may not be fine-grained enough, or because the action-focused nature of the task skewed parents' natural behaviors. Future research is needed to investigate this matter further, for example by examining possible contributors to both IDS and IDA such as parental emotional expressivity. Emotional expressivity, which has been found to strongly manifest in IDS (Benders, 2013; Trainor, Austin, & Desjardins, 2000) and is also considered characteristic of IDA (Brand et al., 2002), might underlie both infant-directed behaviors.

This study provides new evidence for the contribution of language to action learning. Existing research has primarily investigated the opposite direction of influence, namely how movement can contribute to infants' language learning (e.g., Pence, Golinkoff, Brand, & Hirsh-Pasek, 2005). For example, providing consistent visual cues has shown to facilitate preverbal infants' discrimination of phonetic categories (Yeung & Werker, 2009). Further, in object name-learning tasks, mothers have been shown to synchronize the naming of target words with moving the target objects (Gogate, Bahrick, & Watson, 2000), and the synchrony of mothers' movement and speech seems to facilitate infants' learning of object-name pairs (Matayaho & Gogate, 2008). Taken together with the current findings, the commonality of benefiting learning suggests that parental modulations should be seen as a larger, multimodal behavioral repertoire. In this manner, future research might be better able to pinpoint the likely common learning mechanisms.

Furthermore, at the intersection of, but distinct from language and action, are gestures. In contrast to object-directed actions as in IDAs, gestures are non-object-directed movements with representational goals and are typically used in communicative settings (Wakefield, Novack, & Goldin-Meadow, 2018). Research on gestures indicates that gestures are beneficial for learning and generalization in contexts like word learning (Wakefield, Hall, James, & Goldin-Meadow, 2018), solving mathematical problems (Congdon et al., 2017), and also action learning (Novack, Goldin-Meadow, & Woodward, 2015). Though the interpretation of movements as gestures and learning from gestures has mostly been shown in older children, young children could start being able to learn from them by the age of 2 years (Novack et al., 2015; Novack, Filippi, Goldin-Meadow, & Woodward, 2018; Wakefield, Novack, & Goldin-Meadow, 2018). While in the present study maternal object-directed action demonstrations could not co-occur with maternal gestures, gesture and speech can co-occur and thereby enhance learning (Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018). Given the present cross-modal relation between IDS and action learning and the role of gesture in cross-modal learning and generalization, future work should investigate the existence and role of infant- (or toddler-) directed gestures in action learning and its relation to the other infant-directed behaviors of IDA and IDS.

In conclusion, this study suggests that the way mothers speak relates to how well infants learn new actions. While teaching new actions to their infants, mothers spent a considerable amount of time talking to their infants and the way mothers modulated their speech significantly predicted infants' action learning beyond the effect of IDA. We posit that IDS drew attention to the action demonstrations, while IDA maintained infants' attention and highlighted aspects important for reproducing the action. These findings thus provide a starting ground for future work into how parents are expertly able to simplify this multimodal world into coordinated tidbits of information for their learning infants.

References

Bair, E., Hastie, T., Paul, D., & Tibshirani, R. (2006). Prediction by supervised principal components. *Journal of the American Statistical Association, 101*(473), 119-137.

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, *36*(4), 847-862. doi: 10.1016/j.infbeh.2013.09.001

Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.50, retrieved 31 March 2019 from http://www.praat.org/

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, *5*(1), 72-83. doi 10.1111/1467-7687.00211

Brand, R. J., & Tapscott, S. (2007). Acoustic packaging of action sequences by infants. *Infancy*, *11*(3), 321-332. doi: 10.1111/j.1532-7078.2007.tb00230.x

Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O'Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction*, *50*, 65-74.

Cooper, R. P. & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584-1595.

ELAN (Version 5.2) [Computer software]. (2018, April 04). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. doi: 10.3758/BRM.41.4.1149

Ferguson, C. A. (1964). Baby talk in six languages. *American anthropologist*, *66*(6), 103-114. doi: 10.1525/aa.1964.66.suppl_3.02a00060

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477-501. doi: 10.1017/S0305000900010679

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878-894. doi: 10.1111/1467-8624.00197

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The Origins of grammar: Evidence from comprehension. Cambridge, MA: MIT Press.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151.

Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2001). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, *24*(4), 372-392. doi: 10.1016/S0163-6383(02)00086-3

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development*, *32*(4), 437-444. doi: 10.1016/j.infbeh.2009.07.003

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684-686. doi: 10.1126/science.277.5326.684

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers, 41*(3), 841-849. doi:10.3758/BRM.41.3.591.

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, *7*(3), 185-201. doi: 10.1080/15475441.2011.579839

Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers speak less clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological Science*, *26*(3), 341-347.

Matatyaho, D. J., & Gogate, L. J. (2008). Type of maternal object motion during synchronous naming predicts preverbal infants' learning of word–object relations. *Infancy*, *13*(2), 172-184. doi: 10.1080/15250000701795655

Nagai, Y., & Rohlfing, K. J. (2009). Computational analysis of motionese toward scaffolding robot action learning. *IEEE Transactions on Autonomous Mental Development*, *1*(1), 44-54. doi: 10.1109/TAMD.2009.2021090

Novack, M. A., Filippi, C. A., Goldin-Meadow, S., & Woodward, A. L. (2018). Actions speak louder than gestures when you are 2 years old. *Developmental Psychology*, *54*(10), 1809. doi: 10.1037/dev0000553

Novack, M. A., Goldin-Meadow, S., & Woodward, A. L. (2015). Learning from gesture: How early does it happen? *Cognition*, *142*, 138–147. doi:10.1016/j.cognition.2015.05.018

Pence, K., Golinkoff, R. M., Brand, R. J., & Hirsh-Pasek, K. (2005). When actions can't speak for themselves: How might infant-directed speech and infant-directed action influence verb learning. *From orthography to pedagogy: Essays in honor of Richard L. Venezky*, 63-79.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*(5), 1387-1401.

Robinson, C. W., & Sloutsky, V. M. (2007). Visual processing speed: Effects of auditory input on visual processing. *Developmental science*, *10*(6), 734-740. doi: 10.1111/j.1467-7687.2007.00627.x

Rohlfing, K. J., Fritsch, J., Wrede, B., & Jungmann, T. (2006). How can multimodal cues from child-directed interaction reduce learning complexity in robots? *Advanced Robotics*, *20*(10), 1183-1199. doi: 10.1163/156855306778522532

Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Developmental Psychology*, *47*(1), 19.

Schreiner, M. S., Altvater-Mackensen, N., & Mani, N. (2016). Early speech segmentation in naturalistic environments: Limited effects of speech register. *Infancy, 21*, 625-647. doi: 10.1111/infa.12133

Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, *40*(3), 672-686. doi: 10.1017/S0305000912000141

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501-532. doi: 10.1016/j.dr.2007.06.002

The ManyBabies Consortium (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science, 3*(1), 24–52. doi: 10.1177/2515245919900809

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*(1), 53-71.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, *11*(3), 188-195. doi: 10.1111/1467-9280.00240

van Schaik, J. E., Meyer, M., van Ham, C. R., & Hunnius, S. (2020). Motion tracking of parents' infant-versus adult-directed actions reveals general and action-specific modulations. *Developmental Science*, *23*(1), e12869. doi: 10.1111/desc.12869

Wakefield, E. M., Hall, C., James, K. H., & Goldin-Meadow, S. (2018). Gesture for generalization: gesture facilitates flexible learning of words for actions on objects. Developmental Science, 21(5), e12656. doi: 10.1111/desc.12656

Wakefield, E. M., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*, *21*(6), e12664. doi:10.1111/desc.12664

Wakefield, E. M., Novack, M. A., & Goldin-Meadow, S. (2018). Unpacking the ontogeny of gesture understanding: How movement becomes meaningful across development. *Child Development*, *89*(3), e245-e260.

Williamson, R. A., & Brand, R. J. (2014). Child-directed action promotes 2-year-olds' imitation. *Journal of Experimental Child Psychology*, *118*, 119-126. doi: 10.1016/j.jecp.2013.08.005

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition*, *113*(2), 234-243. doi: 10.1016/j.cognition.2009.08.010

Zangl, R., & Mills, D.L. (2007). Brain activity to infant versus adult directed speech in 6- and 13-month olds. *Infancy, 11*, 31-62.

Acknowledgements

We would like to thank Rebecca Brand for her valuable comments on the results of our study, and our student assistants Hilde Krajenbrink and Camila van Ham for their help in coding the motion data of the study, as well as our lab manager Angela Khadar.

Figures and Tables

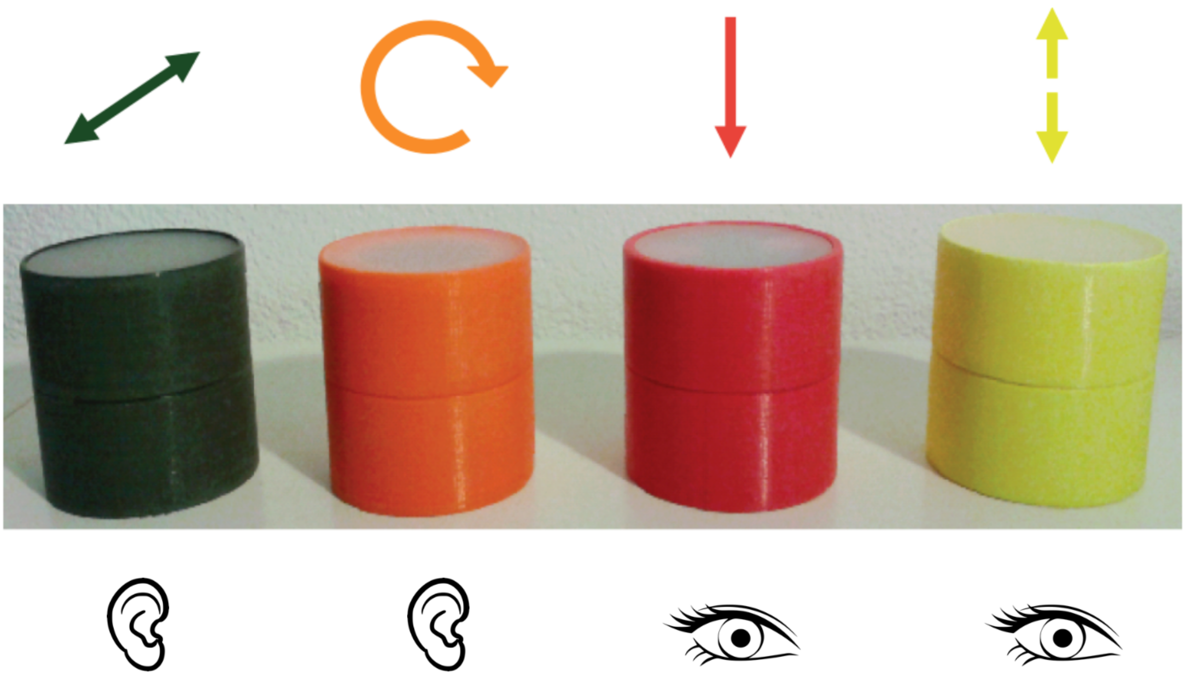

Figure 1. The four novel objects with auditory or visual action effects used for the action demonstrations. Adapted from “Motion tracking of parents’ infant- versus adult-directed actions reveals general and action-specific modulations” by J.E. van Schaik, M. Meyer, C.R. van Ham, S. Hunnius, 2020, *Developmental Science, 23*(1), e12869, p4. CC BY-NC. Published by John Wiley & Sons Ltd.

Table 1

Mean (with standard deviation) and range of IDA and IDS variables for the current Dutch sample. Note that for the analyses averaged z-scores were used. Means of IDS characteristics are also provided for German and British English (from Fernald et al., 1989).

Variable	**Measure**	**Mean (*SD*) Dutch**	**Min-Max**	**Mean German**	**Mean English**
				(Fernald et al., 1989)	
Infant-directed Action	velocity (mm/s)	311.91 (85.75)	132.27-535.18		
	3D-distance (mm)	455.10 (193.97)	191.38-956.29		
	proximity (mm)	494.36 (83.27)	346.74-645.51		
	effect duration (s)	1.80 (.49)	1.03-3.02		
Infant-directed Speech	maximum pitch (Hz)	318.69 (40.99)	213.01-384.64	367	382
	minimum pitch (Hz)	255.19 (29.50)	177.58-322.16	178	198
	mean pitch (Hz)	285.84 (30.47)	194.63-334.01	241	262
	pitch range (semitones/s)	3.96 (2.34)	0.19-8.38		

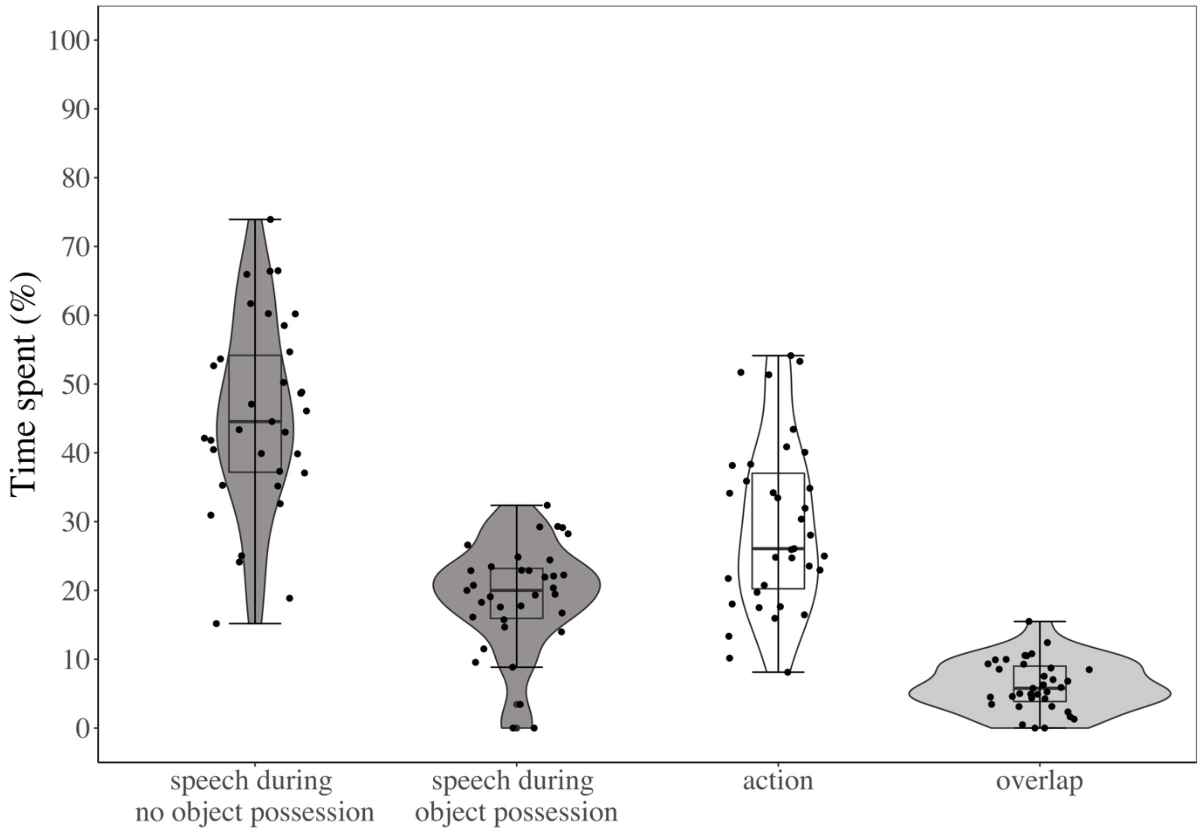

Figure 2. The percentages of time mothers spoke while not in possession of the object, spoke while in possession of the object, acted while in possession of the object, or spoke and acted at the same time, out of the total time mothers spent speaking and/or acting (the width of the violin plot represents the density of the data; dots represent individual data points and are horizontally jittered to ensure visibility).

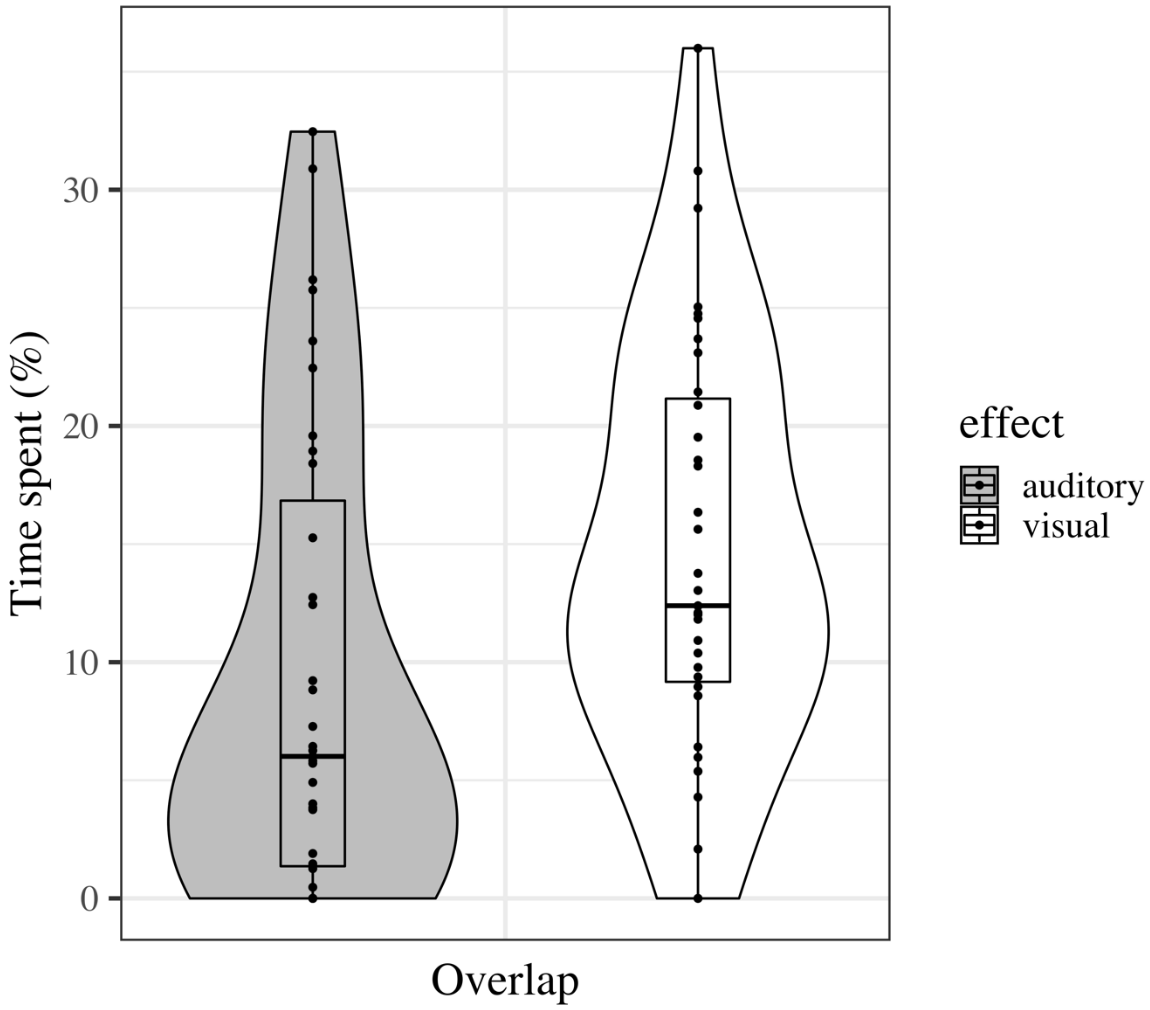

Figure 3. Percentages of time mothers used speech and action in combination during the demonstration of objects with an auditory or a visual effect (the width of the violin plot represents the density of the data; dots represent individual data points).

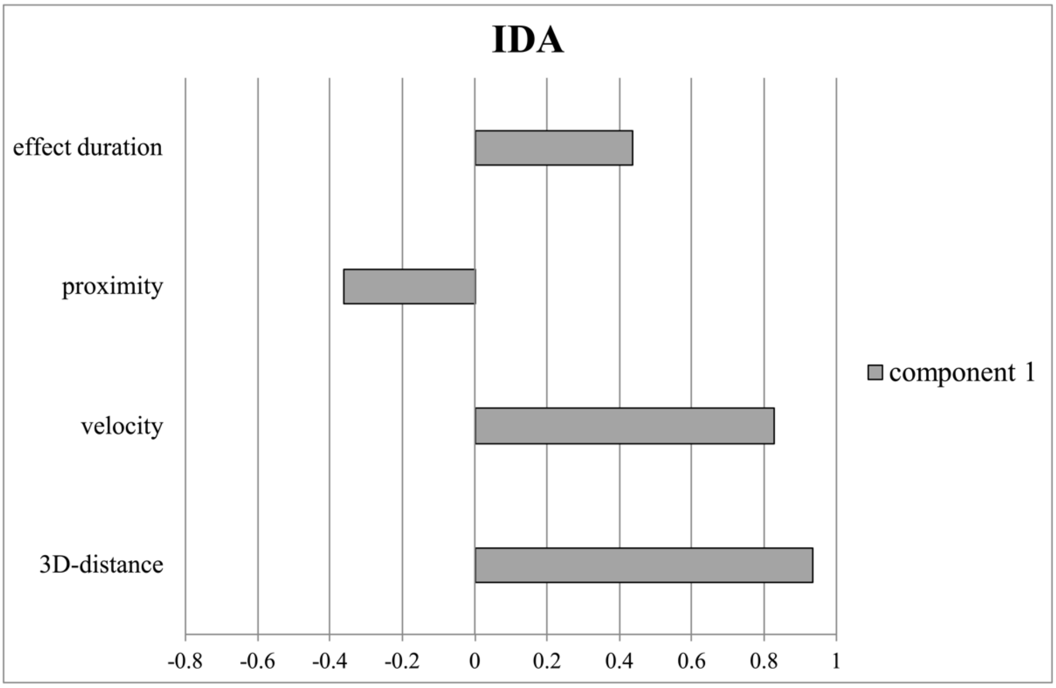

Figure 4. Factor loadings of the extracted IDA component 1.

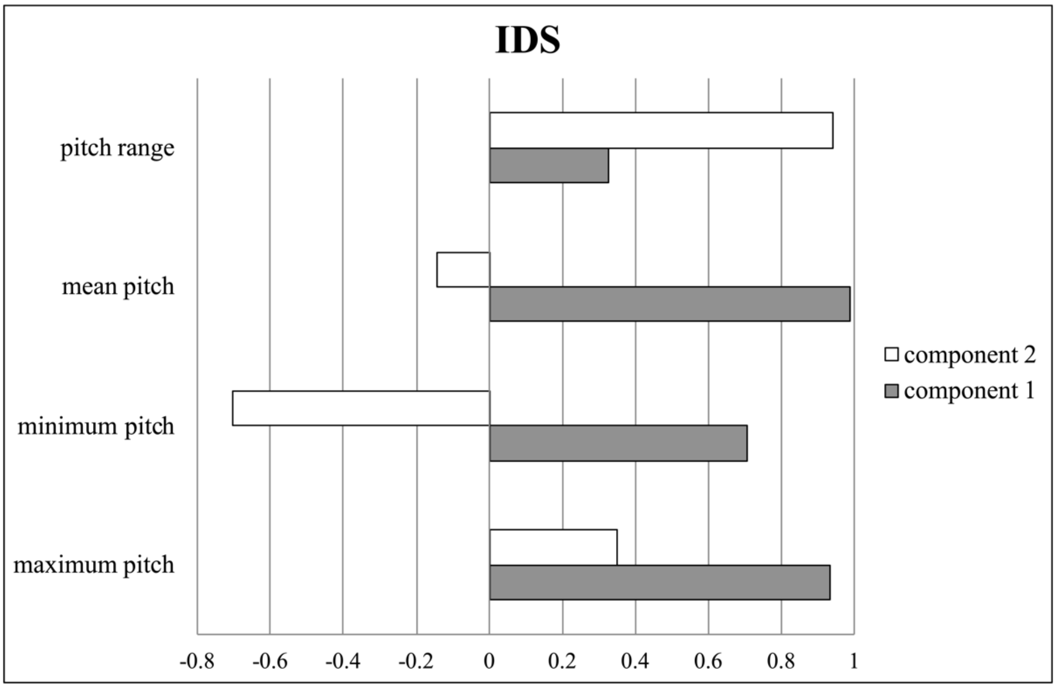

Figure 5. Factor loadings of the extracted IDS components 1 and 2.

Table 2

Correlations of infants' action learning success, IDA component 1, IDS component 1, and IDS component 2.

	Learning success	IDA component 1	IDS component 1	IDS component 2
Learning success	1.00	.35*	-.03	.45**
IDA component 1		1.00	.12	.19
IDS component 1			1.00	.00
IDS component 2				1.00

* $p < 0.05$; ** $p < 0.01$

Table 3

Linear model of predictors of infants' action learning success.

	Model 1			Model 2		
	b	*SE b*	*β*	*B*	*SE b*	*β*
IDA component 1	0.06	0.03	.35*	0.05	0.03	$.28^{+}$
IDS component 1				-0.01	0.03	-.07
IDS component 2				0.07	0.03	.40*
R^2		0.12			0.28	
***F* for change in R^2**		4.49*			3.33*	

$^{+}$ $p < 0.1$; * $p < 0.05$

Supplemental Material

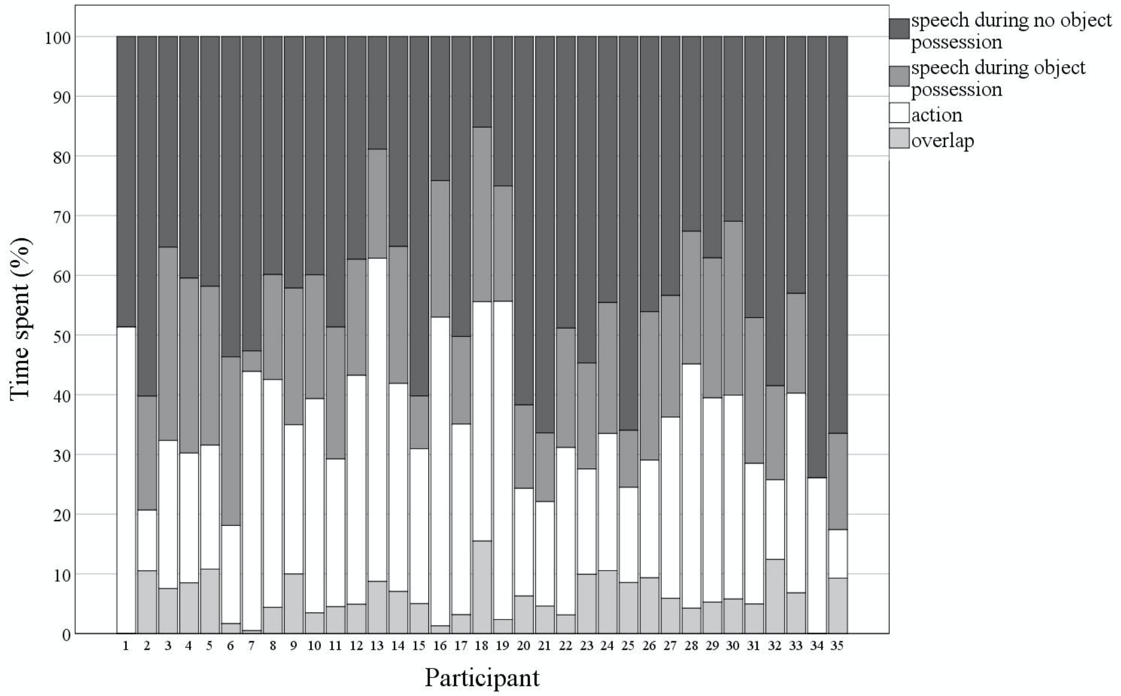

Figure S 1. Percentages of use of speech, action, and overlap of speech and action for each mother.