This paper was presented at the 6th Gesture and Speech in Interaction Conference that was held at Paderborn University, Germany from September 11-13, 2019.

To cite this paper:

Grimminger, A. (2019). The timing of pointing-speech combinations in typically developing and language-delayed toddlers. In: Grimminger, A. (Ed.): *Proceedings of the 6th Gesture and Speech in Interaction – GESPIN 6* (pp. 44-49). Paderborn: Universitaetsbibliothek Paderborn. doi:10.17619/UNIPB/1-807

The timing of pointing-speech combinations in typically developing and language-delayed toddlers

Angela Grimminger

Paderborn University, Germany
angela.grimminger@uni-paderborn.de

Abstract

Research on the development of the gesture–speech integrated system suggests that the temporal alignment becomes closer with progression in linguistic skills. In this study, the multimodal communicative combinations of pointing gestures with speech (vocalizations and first words) in two groups of 18-month-old children with different developmental trajectories in their linguistic development were analyzed: a group of typically developed children and a group of children delayed in language acquisition—as attested retrospectively by a standardized test. Using the reliable paradigm of the decorated room to elicit pointing behavior in children, the analyses focussed on the timing between the two modalities and the temporal distances between gesture and speech onsets. Similar patterns of gesture–speech integration were found for both groups.

1. Introduction

For adult speakers, there is wide consensus that gesture and speech form an integrated communicative system (Kendon, 1980; Kita & Özyürek, 2003; McNeill, 1992). This view is based on the observation that gesture and speech are temporally and semantically synchronized (McNeill, 1992). Synchronization on the semantic level means that both modalities refer to the same idea, either by expressing similar information in gesture and speech (i.e., with one modality being redundant with or complementing aspects to the content of the other, e.g., Iverson & Goldin-Meadow, 2005), or by gesture and speech expressing information that supplement one another. Semantic synchronization has been used to show how children first express themselves multimodally and are increasingly able to combine information within one modality (Iverson & Goldin-Meadow, 2005).

Beside the semantic level, synchronization on the temporal level means that the most prominent part of the gesture, i.e. the stroke, co-occurs with the most prominent part of a speech unit. Following Kendon (1980), a gesture's execution can be described in different phases: the "preparation phase", in which the hand leaves its rest position, the "stroke", and the "retraction", in which the hand or arm return to the rest position. For adult speakers, it has been reported that the onset of a gesture precedes the onset of speech (Bergmann, Aksu, & Kopp, 2011), while the gestural stroke is temporally closely aligned with the onset of the spoken part of an utterance (with a mean temporal distance of about 128 ms between the stroke onset and the onset of speech; Bergmann et al., 2011).

For children's developing communication system, it is of question how and when speech and gesture become integrated. This question is addressed by analyzing the age at which gesture and speech are synchronized on the semantic and temporal level respectively (e.g., Butcher & Goldin-Meadow, 2000; Esteve-Gibert & Prieto, 2014; Murillo, Ortega, Otones, Rujas, & Casla, 2018). In this paper, I will focus on the temporal level.

Studies on precursors for the integrated gesture–speech system investigated the co-development of hand movements and movements of the mouth or vocalizations, and found both modalities to be temporally integrated from very early on (e.g., Ejiri & Masataka, 2001; Iverson & Thelen, 1999; Masataka, 2003). However, with increasing linguistic capabilities this temporal relation becomes closer, as Iverson and Thelen (1999) report for children between 16 and 18 months of age compared to younger children. Similar results were obtained for communicative gestures, of which deictic gestures are the most frequently used type observed in young children (Bates, 1976; Capone & McGregor, 2004). Butcher and Goldin-Meadow (2000) reported that children temporally synchronize verbal utterances and their communicative gestures, mainly deictic gestures at this age,

not until they started to produce their gestures together with meaningful words, as opposed to gestures produced with speech sounds. Note that in this study, temporal integration was operationalized as a complete overlap of the verbal part with the gestural stroke. Esteve-Gibert and Prieto (2014) also included partial overlaps in their analyses of a longitudinal study and were able to shed light onto developmental changes in the temporal coordination of the earliest gesture–speech combinations in a fine-grained way using different measures. This way, the authors revealed support for a closer temporal relation between communicative gestures and speech with increasing linguistic abilities: At 11 months of age, when infants are at the babbling stage, they already combine about 40% of their (mainly deictic) gestures with verbal utterances; critically, once infants produce their first words, the majority of gestures are produced together with speech. Further, and similar to the temporal coordination reported for adult speakers, in this study, the infants' gesture onset preceded the onset of the verbal utterance. Interestingly, the temporal distance between those two measures appeared to be more adult-like in the single-word period compared to the babbling period. Analyses of the temporal distance between the gestural stroke and the onset of the spoken part of the utterances revealed very small differences between the two measures, thus showing that gestures and words were almost simultaneously produced. In sum, the two studies mentioned above show that children start to temporally integrate speech and gestures at an early age, and this temporal relation becomes closer aligned with progress in linguistic skills (see also Murillo et al., 2018).

Many studies have shown a strong and positive relation of infants' use of deictic gestures and their subsequent lexical and syntactic development (e.g., Beuker, Rommelse, Donders, & Buitelaar, 2013; Colonnesi, Stams, Koster, & Noom, 2010; Rowe & Goldin-Meadow, 2009; Rowe, Özçalışkan, & Goldin-Meadow, 2008), not only in children whose language develops typically but also in clinical populations (Brady, Marquis, Fleming, & McLean, 2004; Lüke, Grimminger, Rohlfing, Liszkowski, & Ritterfeld, 2017; Özçalışkan, Adamson, & Dimitrova, 2016; Stolt et al., 2014). Beyond these results, recently the focus seems to have moved from considering the frequency of pointing alone to considering pointing–speech combinations, which might be an even better predictor of lexical skills (Igualada, Bosch, & Prieto, 2015; Murillo & Belinchón, 2012; Wu & Gros-Louis, 2014) and advances in syntactic development (e.g., Fasolo & D'Odorico, 2012; Rowe & Goldin-Meadow, 2009).

Up to date, however, studies addressing the use of pointing–speech combinations in children who show different rates in their early language development are missing. Given the results that the frequency of pointing–speech combinations predicts later linguistic skills, and that the temporal integration becomes closer with increasing linguistic skills, for this study, I hypothesize that the timing of pointing–speech combinations of LD children will differ from that of children with whose linguistic skills are lower. More specifically, I assume that the temporal distance between the onset of the gesture stroke and the onset of the spoken part of the utterance is greater in children with LD. A second hypothesis is that the combinations of pointing gestures with words are temporally closer aligned than pointing gestures with vocalizations. These two hypotheses are examined in children at the age of 18 months, at which period typically developing children start to produce two-word-utterances in speech (cf. Klann-Delius, 2016).

2. Methods

2.1. Participants

14 German-learning children were drawn from a larger sample of 34 families participating in a longitudinal study between 12 and 30 months of age (Grimminger, 2017). Within the whole sample, eight children were identified as being language delayed (LD) at 24 months of age (see below). One of the children with LD was excluded from the analyses here because the parents did not give consent for further analyses. The final sample thus consisted of seven children with LD (2 girls, 5 boys). Seven children with typical language development (TD) were matched for gender.

2.2. Setting and procedure

To elicit spontaneous pointing and verbal utterances, the infants and one of their caregivers (85 % mothers) were observed in a semi-naturalistic setting within a laboratory room that was selectively decorated with 16 interesting objects, pictures (Liszkowski & Tomasello, 2011), and events (e.g., sudden onset of a water fountain). Because we started the longitudinal study when the children were

12 months old, at 18 months of age, they had been in this decorated room with a caregiver several times before. At every session, caregivers were instructed to engage with their children while carrying them for 6 minutes and looking at the objects presented in the room without touching any of them. The data were videorecorded using four cameras from different angles of the room.

2.3. Assessment of language development

To assess the children's language development at 24 months of age, a German standardized language test was administered (*Sprachentwicklungstest für zweijährige Kinder – SETK-2)* [test of language acquisition for two-year-old children] (Grimm, 2000). In accordance with other authors (Heilmann, Ellis Weismer, Evans, & Hollar, 2005; Sachse & von Suchodoletz, 2008), a two-year-old child was defined as being language delayed if she or he scored 1½ standard deviations below the mean (i.e., T-score of ≤ 35) in at least one of the four subtests of the SETK-2 and one standard deviation below the mean in at least one additional subtest (i.e., T-score of < 40).

2.4. Coding

All verbal utterances and gestures of the children were coded using ELAN (Sloetjes & Wittenburg, 2008). Pointing gestures, defined as the extension of the index finger or the whole hand towards an object or location, were the majority of gestures used, and thus, other gesture types will be omitted for the analyses here. For each pointing gesture that was accompanied by a verbal utterance the gesture onset, i.e. the beginning of the preparation phase, and the stroke was coded in order to analyze the timing between the onset of a verbal utterance and the onset of a gesture and its stroke. The gestural stroke was coded as the interval in which the arm and/or index finger were maximally extended (see Esteve-Gibert & Prieto, 2014). The pointing–speech combinations were assigned to one of the following categories: (a) pointing+vocalization; (b) pointing+protoword (German "da"): these combinations were coded as an extra category (see also Liszkowski & Tomasello, 2011), because they mark a transition phase to semantically more complex forms of pointing–speech combinations that contain a word (Clark, 1978); (c) pointing+word. Pointing+two-word utterances were observed only on few occasions, and were therefore excluded from the statistical analyses of temporal distance.

Pointing gestures and verbal utterances that did not at least partially overlap were not considered. If a pointing gesture was accompanied by more than one vocalization, the vocalization closest to the gestural stroke was considered. As an additional measure of temporal coordination, it was coded if the gestural stroke overlapped with the verbal part, either fully, partially or not.

3. Results

First, in most cases of pointing–speech combinations (371 cases in total), the gesture onset preceded the onset of the verbal part (91.4%), and children in both language development groups, TD versus LD, were as likely to do so: TD children, M=92.5% (SD=5.6), children with LD, M=84.8% (SD=16.5), $p > .05$ (Mann-Whitney). Second, in the majority of cases of pointing–speech combinations, the onset of the pointing gesture's stroke followed the onset of speech (73.8%). Again, no differences between both language development groups were found, TD children, M=70.1% (SD=2.3), children with LD, M=77.4% (SD=14.7), $p > .05$ (Mann-Whitney). These results are consistent with previous research and show that at 18 months of age, TD children and children with LD are comparable in their overall pattern of how pointing gestures and speech are aligned. However, some of the children with LD mainly used either vocalizations together with their pointing gestures or protowords, whereas TD children used all forms of combinations. This observation is reflected in the results showing significant group differences in the number of pointing+1-word utterance, Z= -3.27, p= .001, and pointing+two-word utterances, Z= -2.25, p= .05, but not in the number of pointing+vocalizations and pointing+protowords (both $p > .05$, Fig. 1).

While above, the onsets of the two modalities were compared to each other, in the following analyses their temporal distance is considered. Following the finding that temporal integration becomes closer with increasing linguistic skills, it was hypothesized that group differences can be found in the temporal distance between gesture and speech. Thus, the mean temporal distances in

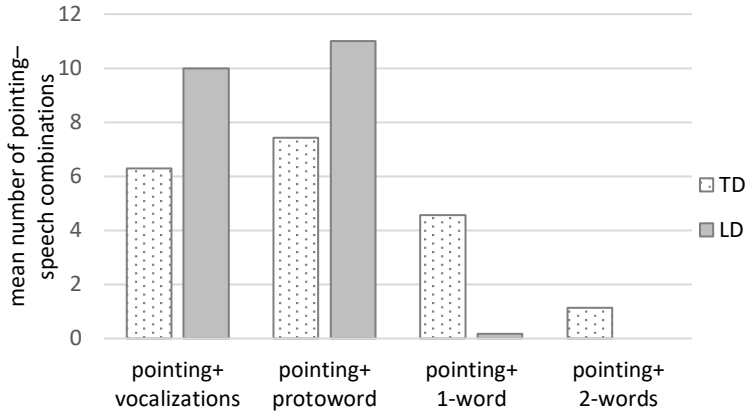

Figure 1. Comparison of the two groups in their use of different types of pointing–speech combinations.

milliseconds between (1) gesture onset and speech onset, and (2) between speech onset and gestural stroke onset were calculated (see Esteve-Gibert & Prieto, 2014). For each participant, the mean temporal distances in milliseconds were separately calculated for each category of pointing–speech combinations, and these variables were compared across the two language development groups. Only those cases were included in which the gesture onset preceded the onset of the verbal part. Contrary to the assumption, the groups did not differ significantly, neither in the temporal distance between pointing onset and speech onset, separately for combinations with vocalizations, proto-words, or one-word utterances (each $p > .05$), nor in the temporal distance between speech onset and stroke onset, again separately for each type of combination (each $p > .05$, Table 1 for the descriptive statistics).

Table 1
Descriptive statistics of temporal distances in millseconds

	TD		LD	
	M	*SD*	*M*	*SD*
gesture onset – speech onset				
vocalization	673,68	220,77	948,62	647,03
protowords	568,68	203,89	487,31	188,10
1-word	455,90	290,72	530,00	.
speech onset – stroke onset				
vocalization	225,51	134,67	253,45	94,92
protowords	266,58	107,13	260,95	92,42
1-word	291,10	163,83	434,00	---

A further hypothesis was that the combinations of pointing gestures with words are temporally closer aligned than pointing gestures with vocalizations. Because no group differences were found, the sample was analyzed as a whole to address this hypothesis. As shown above, the children with LD barely used pointing+1-word combinations, and a Friedman's test was conducted to compare the temporal distances of pointing onset with vocalizations, protowords, and one-word utterances, respectively. This analyses did not yield significant differences, $X_2(2) = .67$, $p > .05$. Because four of the children with LD did not produce any pointing gestures with words, this analysis was performed with $N = 8$. To include more children, we additionally compared the temporal distances of pointing onsets with vocalizations and pointing onsets with protowords using a Wilcoxon test. No differences were found, $Z = -1.65$, $p > .05$, $N = 11$. To compare the distance between the onset of vocalizations and protowords, respectively, and the onset of the gesture stroke, we applied a Wilcoxon test. Again, no differences were found, $Z = -1.48$, $p > .05$, $N = 12$. We also did not find significant differences between the language development groups when comparing the percentage of pointing–speech combinations in which the gesture stroke did not overlap with the spoken part.

4. Discussion

For adult speakers, gesture and speech form an integrated system. Research on gesture–speech integration from a developmental perspective focusses on how and when this integration is achieved, and has impressively shown that infants quite early start to temporally align their manual and vocal activities (e.g., Ejiri & Masataka, 2001; Iverson & Thelen, 1999; Masataka, 2003). However, with increasing linguistic skills, this integrated system becomes more adult-like (Esteve-Gibert & Prieto, 2014; Murillo et al., 2018). Therefore, in this study, I investigated the temporal integration of children's early vocalizations, protowords and words with their gestures in two groups of children that retrospectively differed in their language developmental paths to receive insights into the development of the system. I hypothesized that the timing of pointing–speech combinations of children with LD will differ from that of TD children who I assumed to show a closer temporal alignment of the two modalities. In addition, it was hypothesized that the combinations of pointing gestures with words are temporally closer aligned than pointing gestures with vocalizations. No group differences were found in how children synchronize their pointing gestures with vocal behaviors. Instead, the results for both groups of children are consistent with previous research in adults and infants: The gesture onset preceded the onset of the verbal part, and the onset of the pointing gesture's stroke followed the onset of speech (Bergmann et al., 2011; Esteve-Gibert & Prieto, 2014). This result is intriguing as it suggests a similar processing of early communicative behavior for both, the TD children and children with LD. The only difference that was found pertains to the number of pointing gestures with words being higher in TD children than in children with LD (see Figure 1). Thus, while the TD children at 18 months of age might have a more diverse repertoire of their multimodal communicative means, by using pointing gestures together with different kinds of verbal utterances, the children with LD make greater use of pointing gestures with vocalizations or protowords still. This finding accords with previous findings suggesting that while toward the end of the second year TD children increasingly used words, the pointing gestures of children with LD are accompanied by reduced expressive language (Lüke et al., 2017).

However, some methodological limitations of this study should be stressed. First, the sample size in each group is very small which is due to fact that all children were recruited at 12 months of age for this longitudinal study, and language delay could only be assessed as early as 24 months of age. The results here therefore need further verification. Second, by averaging the temporal distances of gesture onset or stroke onset and speech onset across each subject, the data of the stability in individual patterns were likely stripped: It is possible that the multimodal utterances of children with LD are less consistently synchronized when considering all pointing attempts; vice versa, TD children appear to be rather stable in the way they integrate their vocal behaviors with pointing gestures. To confirm such observations, more fine-grained analyses and methods that are sensitive to patterns on the individual level are necessary. Secondly, even though Lüke et al. (2017) report about early communicative attempts consisting of two types of pointing gestures, namely the index-finger pointing and hand pointing, we excluded pointing gestures performed with the whole-hand from the analyses here, because pointing with the whole hand was shown to be rather negatively related with later language skills. Further investigations taking any gestural form into account might thus look at the way integration with verbal behavior is achieved in whole-hand pointing.

Concerning the debate whether children with LD use their gestures to compensate for their language deficits, our results can be interpreted against the compensation effect, because we found similar integration of pointing with verbal behaviors in both group of children. In addition, as can be viewed from Figure 1, it is rather the group of TD children who is using more pointing–word combinations. With respect to the other forms of speech, no group differences were found. However, it is possible that focusing our analyses on the data point at which children were 18 months old is problematic, since the compensation effect might come into play when children use words rather than protowords and vocalizations. Thus, it can rather be that while the gesture–speech system is developing a similar way in both groups, with increasing linguistic skills, it might start to serve different functions resulting in increasing compensation. This possibility links to a discussion about early communicative attempts being of similar or different nature than conventional use of language (Dore, 1975).

Acknowledgments

This work was supported by the German Research Foundation (DFG; RO 2443/3-1). I thank Juliane Rode for help with annotating the data.

References

Bates, E. (1976). *Language and context: The acquisition of pragmatics* (Vol. 13). New York: Academic Press.

Bergmann, K., Aksu, V., & Kopp, S. (2011). The relation of speech and gestures: Temporal synchrony follows semantic synchrony. *Proceedings of the 2nd Workshop on Gesture and Speech in Interaction (GeSpIn 2011)*.

Beuker, K. T., Rommelse, N. N. J., Donders, R., & Buitelaar, Jan. K. (2013). Development of early communication skills in the first two years of life. *Infant Behavior and Development*, *36*, 71–83.

Brady, N. C., Marquis, J., Fleming, K., & McLean, L. (2004). Prelinguistic predictors of language growth in children with developmental disabilities. *Journal of Speech, Language, and Hearing Research*, *47*, 663.

Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one-to two-word speech: When hand and mouth come together. *Language and Gesture*, *28*, 235–257.

Capone, N. C., & McGregor, K. K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech, Language, and Hearing Research*, *47*, 173–186.

Clark, E. V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In J. S. Bruner (Ed.), *Human growth and development* (pp. 85–120). Oxford: Clarendon Press/Oxford University Press.

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*, 352–366.

Ejiri, K., & Masataka, N. (2001). Co-occurences of preverbal vocal behavior and motor action in early infancy. *Developmental Science*, *4*, 40–48.

Esteve-Gibert, N., & Prieto, P. (2014). Infants temporally coordinate gesture-speech combinations before they produce their first words. *Speech Communication*, *57*, 301–316.

Fasolo, M., & D'Odorico, L. (2012). Gesture-plus-word combinations, transitional forms, and language development. *Gesture*, *12*, 1–15.

Grimm, H. (2000). *Sprachentwicklungstest für zweijährige Kinder (SETK-2)*. Göttingen: Hogrefe.

Grimminger, A. (2017). *Gestische und sprachliche Kommunikation von 12 – 16 Monate alten Kindern und ihren Bezugspersonen in wiederkehrenden, semi-naturalistischen Interaktionen und individuelle Unterschiede in der späteren Sprachentwicklung*. Paderborn: Universitätsbibliothek.

Heilmann, J., Ellis Weismer, S., Evans, J. L., & Hollar, C. (2005). Utility of the MacArthur—Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology*, *14*, 40–51.

Igualada, A., Bosch, L., & Prieto, P. (2015). Language development at 18 months is related to multimodal communicative strategies at 12 months. *Infant Behavior and Development*, *39*, 42–52.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychol. Sci.*, *16*, 367–371.

Iverson, J. M., & Thelen, E. (1999). Hand, mouth and brain. The dynamic emergence of speech and gesture. *Journal of Consciousness Studies*, *6*, 19–40.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The Relationship of Verbal and Nonverbal Communication* (pp. 207–227). Mouton: The Hague.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.

Klann-Delius, G. (2016). *Spracherwerb* (3rd ed.). Stuttgart: J.B. Metzler.

Lüke, C., Grimminger, A., Rohlfing, K. J., Liszkowski, U., & Ritterfeld, U. (2017). In infants' hands: Identification of preverbal infants at risk for primary language delay. *Child Development*, *88*, 484–492.

Masataka, N. (2003). From index-finger extension to index-finger pointing: Ontogenesis of pointing in preverbal infants. In S. Kita (Ed.), *Pointing: Where Language, Culture, and Cognition Meet* (pp. 69–84). Mahwah: Lawrence Erlbaum.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, USA: Chicago University Press.

Murillo, E., & Belinchón, M. (2012). Gestural-vocal coordination: Longitudinal changes and predictive value on early lexical development. *Gesture*, *12*, 16–39.

Murillo, E., Ortega, C., Otones, A., Rujas, I., & Casla, M. (2018). Changes in the Synchrony of Multimodal Communication in Early Language Development. *Journal of Speech, Language, and Hearing Research*, 1–11.

Özçalışkan, Ş., Adamson, L. B., & Dimitrova, N. (2016). Early deictic but not other gestures predict later vocabulary in both typical development and autism. *Autism*, *20*, 754–763.

Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, *12*, 182–187.

Rowe, M. L., Özçalışkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, *28*, 182–199.

Sachse, S., & von Suchodoletz, W. (2008). Early Identification of Language Delay by Direct Language Assessment or Parent Report? *Journal of Developmental & Behavioral Pediatrics*, *29*, 34–41.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by Category: ELAN and ISO DCR. *LREC Proceedings*, 816–820. Marrakesch, Marokko.

Stolt, S., Mäkilä, A.-M., Matomäki, J., Lehtonen, L., Lapinleimu, H., & Haataja, L. (2014). The development and predictive value of gestures in very-low-birth-weight children: A longitudinal study. *International Journal of Speech-Language Pathology*, *16*, 121–131.

Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, *34*, 72–90.