How hand movements and speech tip the balance in cognitive development

de Jonge-Hoekstra, Lisette

DOI:
10.33612/diss.172252039

Publication date:
2021

6

General Discussion

General Discussion

In this dissertation, my aim was to understand how cognitive development is related to how children move their hands and how they speak during cognitive tasks, over time and at multiple (time) scales. I therefore examined how children's hand movements, gestures, and speech relate to each other and to the physical and social environment, when children engage in science and technology tasks. Previous studies have shown that children's hand movements lead cognitive development, over speech. For example, children use their hands to explore (e.g. Adolph & Franchak, 2017), and they show understanding in gestures before they can put this into words (e.g. Cook & Goldin-Meadow, 2006; Goldin-Meadow, 2001). My research questions and hypotheses, as well as the methodologies and data-analyses, were based on the theoretical perspectives of *complex dynamical systems*, *coordination dynamics*, and *affordances*, because I expected these perspectives to bring unique insights about hand movements' leading role in cognitive development. A detailed explanation of the theoretical perspectives can be found in the General Introduction. In this dissertation's studies I investigated *hand movements*, *gestures*, or a *combination* of the two[1]. Depending on whether a study emphasized hand movements or gestures, I will choose one of the terms in the study descriptions below.

Summary of studies

Study 1 - Asymmetric dynamic attunement of speech and gestures in the construction of children's understanding

In the first study (Chapter 2), I investigated whether the leading role of hand movements and gestures (hereafter: gestures) in children's cognitive development is also evident *within* tasks, as opposed to *between* tasks, which previous studies have shown. Using a multiple case study design (N = 12), I researched how children performed a science and technology task together with an adult. I coded children's gestures and speech, and assigned levels of understanding to both their gestures and speech. Furthermore, Chromatic and Anisotropic Cross Recurrence Quantification Analysis (Cross RQA) was used to investigate the coupling between levels of understanding of gestures and speech, in terms of which modality is leading in time, and which modality attracts the other more strongly (see Figure 1). In addition, I examined whether these within-task-measures of coupling were related to more static measures of cognitive development.

[1] Gestures are hand movements which are typically tightly coordinated with speech, while hand movements are a broader category and are not necessarily coordinated with speech. The boundary for when a hand movements becomes a gesture is fuzzy, however. I will also touch upon this later in the General Discussion.

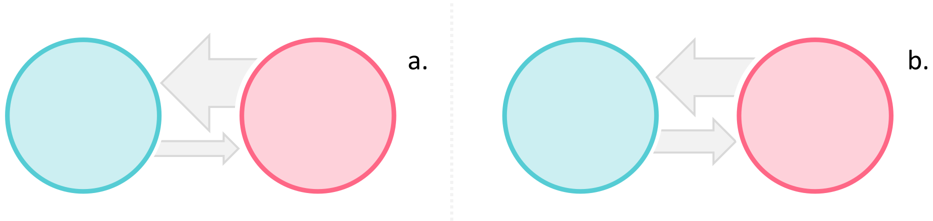

Figure 1. Schematic visualization of coupling between levels of understanding of gestures (blue) and speech (pink). Panel a shows that understanding in speech attracts understanding in gestures much more than vice versa, as displayed by the large difference in arrow size. This pattern was typical of children with a high score on a standardized language test. Panel b also shows that understanding in speech attracts understanding in gestures more than vice versa, but it is more balanced than the coupling relation between gestures and speech in panel a. This last pattern was typical for children with a high score on math or past science and technology tasks.

The results showed that Kindergartners' (n = 5) understanding in gestures was leading understanding in speech by 18 s, on average. For first-graders (n = 7) understanding in gestures and speech was more synchronized, whereby speech was slightly leading (M = 0.71 s). I thus only found a leading role of gestures in time for younger children. Furthermore, for all children speech attracted gestures more than vice versa. Interestingly, this asymmetry in speech attracting gestures was more pronounced for children who scored higher on a standardized language test[2]. For children who scored higher on a standardized math test[2], or on past science and technology tasks, the asymmetry was less pronounced, and speech and gestures were thus more balanced. This last finding could be taken to suggest that, for these children, speech constrains gestures relatively less, and that children benefit from this during math or science and technology tasks.

Besides the empirical findings, in this first study I proposed a theoretical perspective on the coupling between gestures and speech. I proposed that gestures and speech were two coupled, yet separate synergies. In other words, both gesturing and speaking involves the functional organization of many components (e.g. muscles, bones, neurons) at many scales (e.g. cells, muscles, brain) throughout the neuromuscular system. Critically, there is overlap in the components which are involved in the simultaneously occurring synergies of gestures and speech (Figure 2a). For instance, our lungs, muscles in our back, and the neural structure of Broca's area are involved in both gesturing and speaking. Based on this proposal, I suggested that gesture-speech mismatches are due to a competition and subsequent decoupling between the synergies of gestures and speech during difficult tasks (Figure 2b). This idea was empirically challenged in the second study.

[2] We used the Kindergarten CITO as our standardized language and math test. This test is not used anymore in primary schools in the Netherlands, because its predictive value of later language and math scores is limited.

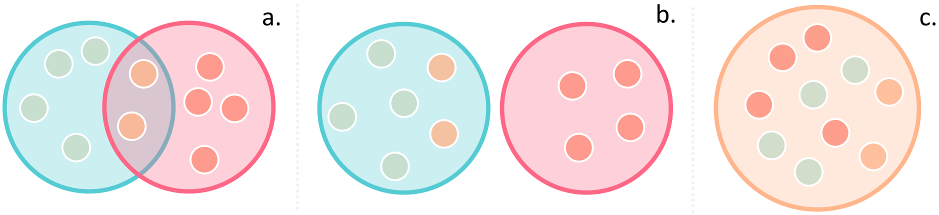

Figure 2. Schematic visualization of synergies of gestures and speech. The small dots represent components at many scales. The large circles represent synergies. Panel a displays the coupled synergies of gestures (blue circle) and speech (pink circle), whereby there is an overlap between the two synergies with respect to the involved components. Panel b displays the decoupled synergies of gestures and speech during gesture-speech mismatches, due to competition between the two synergies during difficult tasks (as proposed in Study 1/Chapter 2). Panel c displays how the components involved in the synergies of gestures and speech have organized into one gesture-speech-synergy (orange circle), due to difficult tasks (as proposed in Study 2/Chapter 3).

Study 2 - Easier said than done? Task difficulty's influence on temporal alignment, semantic similarity, and complexity matching between gestures and speech

In the second study (Chapter 3), I investigated whether task difficulty influences gesture-speech synchronization. Different from the other three studies in this dissertation, I investigated bachelor students instead of children. We chose to investigate bachelor students instead of children, because bachelor students would be able to persevere for 1100 trials, while this would be nearly impossible for children.

In the experiment, task difficulty was manipulated by means of the task layout (also see Figure 4, Chapter 3, p. 74). In the easy condition participants had to match targets of the same color by pointing to locations and saying the word of the locations in a *regular* pattern. In the difficult condition participants also had to match targets of the same color, but they had to point to locations and say the word of the locations, which were presented in a *random* pattern. I analyzed three forms of gesture-speech synchronization: Temporal alignment, semantic similarity, and complexity matching. Temporal alignment pertained to the time between pointing to a location and saying the word of a location. Semantic similarity referred to whether the location to which a participant pointed matched with the location that the participant said in words. Important to note is that gesture-speech mismatches can be seen as semantic *dissimilarity* between gestures and speech. Complexity matching can be thought of as a similarity in the multiscale (i.e. including a whole range of shorter and longer timescales) organization of gestures and speech. I used Multi Fractal Detrended Fluctuation Analysis (MFDFA), which informs about systems' multiscale organization, to investigate complexity matching between gestures and speech.

I found that task difficulty indeed influences students' gesture-speech synchronization. I found less temporal alignment, less semantic similarity (thus more gesture-speech mismatches), and more complexity matching between gestures and speech in the difficult than in the easy condition. This last finding suggests that the synergies of gestures and speech *do not* compete and subsequently decouple during difficult tasks, as we suggested in the first study. Instead, it seems like the components involved in the synergies of gestures and speech organized into a combined gesture-speech-synergy (Figure 2c). Interestingly, this gesture-speech-synergy during difficult tasks seems to be characterized by less temporal alignment and less semantic similarity than the coupled synergies of gestures and speech during easy tasks (Figure 2a).

A second interesting finding from the second study was that participants never *pointed* to the incorrect location, whereas they frequently *said* the incorrect location. This means that all the instances of semantic dissimilarity were due to incorrectly saying the location. I drew a parallel between this finding of gestures always being correct while speech was not, and gestures' leading role in cognitive development during gesture-speech mismatches. To explain both phenomena, I proposed that gestures and hand movements, regardless of age, are strongly coupled to spatiotemporal properties of the environment, while speech is not strongly coupled to these properties. I further investigated this proposal in the third study.

Study 3 - Movers and shakers of cognition: Hand movements, speech, task properties, and variability

Previous studies have shown that a transition from "old" to "new" understanding is a reorganization of the system which goes together with an increase in variability[3] (Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009). As a brief reminder, in the second study I proposed that hand movements' leading role in cognitive development is related to hand movements' typically stronger coupling with spatiotemporal task properties. If this is true, one would expect a difference in these participants' hand movements' variability upon a change in relevant spatiotemporal task properties (i.e. length of the balance scale, or relative difference in mass between the two weights), but not in speech. In the third study (Chapter 4), I researched whether a change in the spatiotemporal task properties of a balance scale task differentially affected the variability of speech and hand movements (including gestures; hereafter: hand movements) during children's explanations. Children between 4 and 7 years were investigated, which is the age range in which children typically transition from only regarding the mass of the

[3] I explained this increase in variability during transitions and reorganization using a LEGO-metaphor (Chapter 4, p. 99): One can only build a new structure from an old structure if one breaks the old structure (increase the variability) and uses the bricks to build a new structure.

weights to solve balance scale problems to including the distance of the weights from the fulcrum in their explanation (see Figure 3).

Children were asked to predict and explain about balance scale problems to an experiment leader. They hereby participated in one of two experiments (also see Figure 1, Chapter 4, p. 100). In the first experiment, children worked with a long balance scale in the first half of the task, and with a short balance scale in the second half (Long-Short condition), or vice versa (Short-Long condition). The difference in length of the balance scale was related to the relevant dimension of distance from the fulcrum for solving balance scale problems. In the second experiment, children worked with a large difference in mass of the weights in the first half of the task, and with a small difference in mass in the second half (Large-Small condition), or vice versa (Small-Large condition). The difference in mass of the weights was related to the relevant dimension of weight. Children's hand movements and speech were separately coded, in broad behavioral categories, and with a frequency of 2 Hz (2 data points per second). On these time series of hand movements and speech, I calculated two measures of variability. The first measure was *Diversity*, which was calculated by means of the Shannon entropy of the frequency distribution of the behavioral categories and their duration. Diversity indicates a system's adaptability (e.g. Adolph et al., 2015), and is related to the range of behaviors across the task. The second measure was *Complexity*, which was derived by applying categorical RQA, and calculating the block entropy (Leonardi, 2018). Complexity indicates a system's flexibility, and is related to the temporal structure of behavior across the task (Adolph et al., 2015).

Across the two experiments, and thereby four conditions, I found a difference in Complexity after a change in relevant spatiotemporal task properties, but not in Diversity, for *both* hand movements and speech in all but the Long-Short condition. Furthermore, only in the Large-Small condition the change in Complexity (and Diversity) of hand movements and speech was related. In other words, changes in relevant spatiotemporal task properties thus often affected the flexibility (but not adaptability) of both hand movements and speech. The differences in flexibility of hand movements and speech *within* children were mostly unrelated, however.

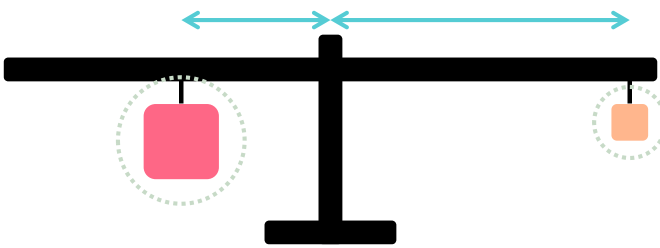

Figure 3. Example of a balance scale task. The dotted lines are drawn around the weights. The arrows indicate the distance of the weights from the fulcrum.

These findings suggest that changing the relevant spatiotemporal properties of a task thus influences the flexibility of children's hand movements and speech, but it is yet unclear whether and how such changes benefit cognitive change in children. Furthermore, I suggested that changes in spatiotemporal task properties might not even have to be related to the *relevant* dimensions of the task in order to elicit an increase in variability. In line with this, Stephen, Dixon, et al. (2009) found that *random* changes in spatiotemporal task properties increased the variability of participants' hand movements, and thereby raised the likelihood of discovering a new cognitive strategy. In addition, children not only needed to attune their hand movements (and speech) to spatiotemporal task properties, but also to the experiment leader to which they explained their answers. These social interactions shape children's hand movements and speech as well, which might explain why I found changes in both hand movements' and speech's flexibility. I further addressed how children attune their multimodal behavior to each other during social interactions in the fourth and last study.

Study 4 - Get it together: How do children interact multimodally and at multiple time scales when they solve dyadic balance scale problems?

In the final study, I investigated interpersonal coordination between children's speech, hand movements, and head movements, when children solve balance scale problems in pairs. I collected data of 25 dyads ($N_{children}$ = 50), between 6 and 10 years of age, in the Connecticut Science Center. For each balance scale problem, children were asked to first individually discuss about the outcome of the balance scale problem, by pressing a button on a game controller. If they disagreed or if they both gave the wrong prediction, the children needed to discuss what they individually predicted and why. After these discussion moments, children were asked to give a second, individual prediction about the particular balance scale problem. I measured children's hand movements and head movements during discussion moments, and recorded their speech. To investigate interpersonal coordination, I applied cross wavelet analysis on the time series of intensity of speech, hand movements, and head movements during discussion moments. Cross wavelet analysis allowed me to look for similarities in the periodicities (i.e. frequency of the waves [∿] which make up communication) of each modality between children in a dyad over time (see Figure 1, Chapter 5, p. 130, and Figure 4, Chapter 5, p. 138, for a more elaborate explanation), and thereby investigate which multimodal coordination patterns occurred at which timescales when children solve dyadic balance scale problems. Furthermore, I examined whether these measures of interpersonal coordination predict dyadic task performance after discussing.

I found that children together coordinated their speech mostly at a (fast) timescale from 2-8 Hz (i.e. one ∿ per 0.5 to 0.125 s), while they coordinated their hand movements and head movements mostly at a (slow) timescale from 0.25-2 Hz (i.e. one ∿ per 4 to 0.5 s). These findings are comparable to results about interpersonal coordination between adults, albeit less pronounced. Furthermore, I found that children's speech at a timescale from 0.5-8 Hz (i.e. one ~ per 2 to 0.125 s) is mostly in an in-phase (i.e. 0-90°) relation, which means that one of the children typically followed the other child's talking, with a delay between 0 and 1 s (depending on the timescale as well as the exact phase relation). There were no clear phase relations between the other modalities at either of the timescales, however. Lastly, except for an anti-phase relation between children's speech being predictive of *not* adopting the other child's prediction, there were no relations between the measures of interpersonal coordination and indices of dyadic task performance. In line with our findings, previous studies found interpersonal coordination between people who do something together, but findings are mixed about whether there is a relation between measures of interpersonal coordination and task performance. I therefore suggested that more *systematic*[4] and *qualitative* research is needed to learn how interpersonal coordination contributes to how well people work together in general, and how children learn from each other during cognitive tasks in particular.

In the following sections I will discuss the findings summarized above from several angles. Firstly, I will deal with how to understand the role of children's hand movements and speech in cognitive development from the perspective of complex dynamical systems, coordination dynamics, and affordances, and what the studies, methods, and analyses, in this dissertation contribute to this understanding. Secondly, based on this as well as on the studies' limitations, I will lay out important directions for future and follow-up research. Finally, I will discuss which practical implications, for instance for education, follow from this dissertation.

Coupling between hand movements and speech during science and technology tasks

The findings of Study 1 and 3 (Chapter 2 and 4) point into one direction: When children are asked by someone else to explain about science and technology tasks, their hand movements and/or gestures and speech are clearly coordinated and coupled. This finding resonates with many previous studies (e.g. Alibali et al., 2014; Church & Goldin-Meadow, 1986; Goldin-Meadow, 2001; Goldin-Meadow, 2017; Kita, 2000). With regard to the specifics of this coupling, in Study 1 (Chapter 2) I found that understanding in speech attracts understanding in gestures more than vice versa (i.e. *asymmetric* coupling). This result possibly reflects that speech often

[4] At the moment, different studies have applied diverse measures of interpersonal coordination to a diverse range of interpersonal tasks. This makes it difficult to compare the outcomes of the studies. If we would systematically apply the same measure of interpersonal coordination to different tasks, and/or apply diverse measures of interpersonal coordination to one task, it would be possible to disentangle the influence of task specifics and measurement specifics on the research findings.

recruits hand movements when someone is speaking (Iverson & Thelen, 1999), while many of our hand movements, such as manipulating objects, happen without speech. Similar asymmetric coupling has been found between movements of someone's dominant and non-dominant hand (de Poel et al., 2007). Furthermore, in Study 3 (Chapter 4) I found that changes in spatiotemporal task properties influence both hand movements' and speech's variability. This might also indicate a bidirectional coupling between hand movements and speech, whereby changes in hand movements recruit changes in speech as well. In line with this, Pouw, Harrison et al. (2019) found entrainment between forceful, vertical hand movements, and the fundamental frequency and amplitude envelope of speech.

The above findings are supplemented and extended by the results of Study 2 (Chapter 3), in which students (instead of children) performed a simple cognitive task. In an easy, repetitive task, I found phase synchronization between gestures and speech, in line with the entrainment found by Pouw, Harrison et al. (2019). Furthermore, a difficult task with random variation even induced synergetic coupling between gestures and speech, which entails synchronization on multiple timescales. One way to interpret these findings is that they reflect two possible coordination patterns between gestures and speech. In the first coordination pattern, gestures and speech are like two phase-locked oscillators. This coordination pattern is characterized by close temporal alignment and high semantic similarity (Wagner et al., 2014). I hypothesize that this coordination pattern is the "default" mode of gesture-speech coordination, in line with how weakly coupled systems ubiquitously have been shown to synchronize (Rosenblum et al., 2001). This coordination pattern could be seen as a strong and stable attractor. In the second coordination pattern gestures and speech no longer behave as two oscillators, but instead form one large functional organization with different behavioral characteristics, such as low temporal alignment, low semantic similarity, but high accuracy of gestures. This second coordination pattern is a much weaker and less stable attractor than the first. Only situations and tasks with specific characteristics are able to elicit the second coordination pattern. I will further address the relation between the coupling between hand movements and speech and tasks with specific characteristics in the next section.

Hand movement-speech coupling is embedded and nested

Children's hand movements and speech, and the coupling between them, are embedded within the characteristics of the environment. Embedded means that the environment opens up and constrains possibilities for action (i.e. affordances; Adolph, 2019; Adolph & Hoch, 2019; Gibson, 1966), such as hand movements and speech. For example, in Study 2 (Chapter 3), students engaged in a task with a specific physical structure (targets in a regular or random pattern) and also a very specific instruction (match targets of the same color by means of pointing and saying their location). The differences in physical structure between the two conditions resulted in clear differences in the coupling between students' hand movements and speech.

Also with regard to the science and technology tasks that I used in Study 1, 3, and 4 (Chapter 2, 4 and 5), children's hand movements and speech are embedded within tasks with very specific characteristics. These tasks require children to verbally, and preferably correctly, answer someone else's questions, and are thus characterized by a particular social structure. Anything a child does during these tasks is directed towards someone else. In addition, the verbal answering causes most of the child's hand movements to happen together with speech. In line with the social structure of such tasks, in Study 4 (Chapter 5), I found that children coordinate their speech, hand movements, and also head movements with the person with whom they are doing the task. Furthermore, the task material itself has a specific physical structure. For example, in a balance scale task the physical task characteristics are laid out so that only the distance from the fulcrum and the weights can change. Moreover, the social and physical structure of the task are also intertwined, as behaving according to the social structure (correctly answering the questions) requires children to attune to the relevant physical structure of the task material (also see E. J. Gibson & Pick, 2000; Adolph & Kretch, 2015).

Besides being embedded in science and technology tasks with particular social and physical characteristics, children's hand movements and speech, and the coupling between the two, are nested within cognitive understanding (see Figure 4 for a schematic visualization; also see General Introduction, p. 19-22; Adolph & Hoch, 2019; Kloos & Van Orden, 2009; Newen et al., 2018). Cognitive understanding can be seen as a coordination pattern which is functional within a specific task (for a more in-depth explanation, please see the General Introduction, p. 19-22). In the case of a science and technology task, this coordination pattern involves things like perceiving and attuning to particular physical properties of the environment, and communicating with someone else. Being nested means that cognitive understanding constrains and enables children's hand movements, speech, and hand movement-speech coupling, while children's hand movements, speech, and hand movement-speech coupling also constrain and enable cognitive understanding. Furthermore, the above definition of cognitive understanding as a functional coordination pattern within a specific task implies that cognitive understanding is also embedded within the characteristics of the environment, such as the science and technology tasks in this dissertation (Study 1, 3, and 4/Chapter 2, 4, and 5). The embeddedness and nestedness of children's (coupling of) hand movements and speech and cognitive understanding has consequences for hand movements' role in cognitive development.

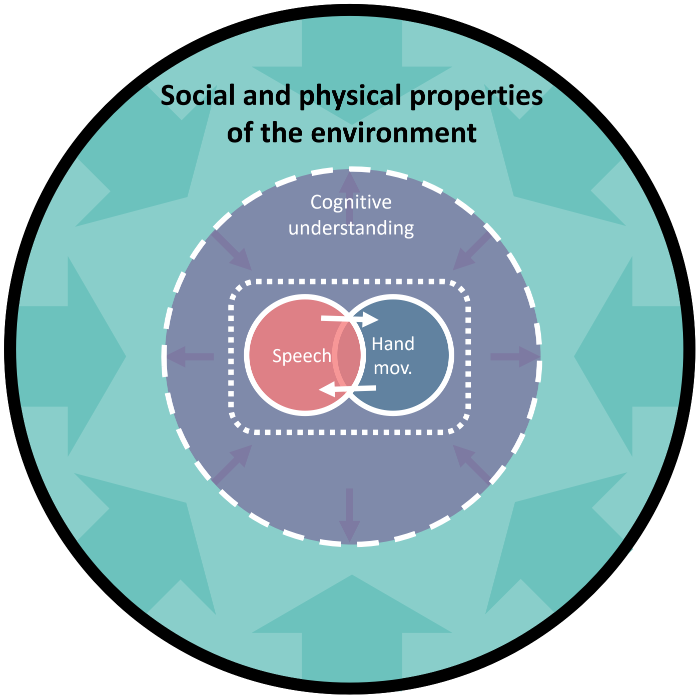

Figure 4. Speech, hand movements, and hand movement-speech-coupling are nested within cognitive understanding. This means that cognitive understanding constrains and enables children's hand movements, speech, and hand movement-speech coupling, while children's hand movements, speech, and hand movement-speech coupling also constrain and enable cognitive understanding (the bidirectional purple arrows). Furthermore, both cognitive understanding and speech, hand movements, and hand movement-speech-coupling are embedded in the social and physical properties of the environment (the uni-directional blue-green arrows).

Hand movements' (leading?) role in cognitive development

If children's hand movements and speech, and the coupling between them, as well as cognitive understanding are embedded within the characteristics of the environment, then hand movements' (leading) role in cognitive development is also dependent on, and embedded in, the task. For example, manual exploration enables someone to learn about the physical structure (e.g. shape, size, surface) of objects, among other things. People use their hands for this exploration, because hands are usually the parts of the body which are simply most suited for functionally exploring objects. Using speech (i.e. sound waves) to explore things like the shape, size or surface of an object would provide you with little to no useful information about the physical structure of the object. I therefore propose that it is strange to say that hand movements are leading exploration of objects over speech, while speech was never able to provide any useful information about the physical structure of objects in the first place[5].

With regard to gestures in specific, researchers often have used tasks similar to the science and technology tasks that I used in this dissertation (Study 1, 3, and 4/Chapter 2, 4, and 5) to investigate gestures' role in cognitive development. As described in the previous section, these tasks have very specific social and physical characteristics, whereby children need to perceive and attune to the relevant physical, spatial properties of the task in order to correctly answer someone's questions. Similar to manual exploration, children's hands are the prime candidates to explore the physical, spatial properties of the task material, while this is simply impossible for speech (also see Study 3 and the Discussion of Study 2). In other words, these tasks are devised to elicit hand movements. Different from manual exploration, such tasks also require children to talk, in order to answer the questions. As described before, speaking often recruits hand movements. In line with this, I could speculate that (deictic and representational) gestures reflect hand movements' simultaneous coordination with the rhythm of speech (coordination dynamics) and the typical possibilities of action of hands (affordances) (also see Pouw et al., 2018; Wagner et al., 2014). The above has implications for gestures' leading role in cognitive development, in the form of gesture-speech mismatches.

Why do gesture-speech mismatches occur?

During gesture-speech mismatches, children's hands correspond to the relevant physical properties of the task material for answering the questions, while their speech does not. Analogous to manual exploration, it is strange to say that gestures are leading cognitive development about tasks, which have specific physical and spatial properties, while speech is just not suited to fully correspond to these properties. Interestingly, also structural changes in eye movements, which are capable of moving according to the spatial structure of the task as well, have been shown to precede changes in verbal explanations (e.g. Dixon et al., 2012; Grant & Spivey, 2003). In that sense, instead of asking whether gestures are leading cognitive development, the following question about gesture-speech mismatches might be more useful: Why do children not speak the right words to indicate the relevant physical structure of the task, while their hands (or eyes) have started to attune and correspond to this physical structure?

[5] This does **not** mean that speech is never important for manual exploration. Instead, many of children's exploratory actions will emerge in the context of social interactions, in which children speak with others. Speech and exploration will thus often elicit and constrain each other, which again emphasizes the tight coupling between speech and hand movement on many levels, their embeddedness in the physical and social environment, and their nestedness within cognitive development.

One important ingredient for answering this question is that gesture-speech mismatches are known to happen around cognitive insights (e.g. Church & Goldin-Meadow, 1986; Perry et al., 1992). As cognitive understanding is a functional coordination pattern, a cognitive insight is the emergence of a new coordination pattern. In line with this, a cognitive insight shows characteristics of a self-organized transition of the system (e.g. Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009). Such a transition is typically depicted as moving from a stable coordination pattern, through a variable transition phase, to another (more or less) stable coordination pattern (also see Study 3/Chapter 4). The variable transition phase means that both children's hand movements and speech will be variable too around cognitive insights.

Although hand movements are variable during such a transition phase, children can still (relatively easily) move and shape their hands according to the physical properties of the environment, as these physical properties are directly specified by the environment. However, there is relatively little in the environment that directly corresponds to children's speech, or the structure of sounds which corresponds to particular words (also see Study 2/Chapter 3). Therefore the variability in speech, in the form of searching for the right words, is more apparent than the variability of children's hand movements, even though hand movements also become more variable around a cognitive insight (Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009). The findings in Study 2 (Chapter 3), in which the difficult task led to a merged gesture-speech synergy, with low temporal alignment, low semantic similarity, but high gesture-accuracy, might also reflect such a variable phase.

Within science and technology tasks, in extreme cases the variability in speech might even cause children to reiterate a structure of sounds that they are familiar with and are able to put into words, and which one might interpret as "old" understanding[6]. When children's hands movements simultaneously attune to and correspond to the relevant physical properties of the task, one might interpret this combination as a gesture-speech mismatch. In addition, the coding systems which have led to the discovery of gesture-speech mismatches might exaggerate subtle differences between the content of gestures and speech (also see Koschmann, 2017), leading to even more coded gesture-speech mismatches.

In conclusion, hand movements' role in cognitive development is embedded in and dependent on the specifics of the task. When tasks have a physical, spatial structure, children's hand movements are often the most direct and suitable body parts to explore and learn about this physical structure. When a task also requires children to speak and answer questions, hand

[6] In a sense, this might be similar to the perseveration in the A-not-B error (see General Introduction, p. 21-22 for an elaborate exploration), whereby children continue to search at location A while the toy has been hidden at location B.

movements are recruited by speech and become gestures, which are attuned to the rhythm of speech while also corresponding to the physical properties of the environment. During cognitive insights, children's hand movements and speech become variable, whereby hand movements can be structured according to the directly specified physical properties of the task, while there is nothing in the environment that directly specifies speech. While this might seem as if understanding in hand movements is ahead of speech, and is therefore coded as such, both hand movements and speech reflect the variability of the system during cognitive insights. In other words, neither hand movements nor speech are driving or leading cognitive development over the other. Instead, cognitive development and cognitive insights are fuzzy processes, which arise from the complex and nuanced interactions between the system as a whole, the coupling between a child's hand movements and speech, and the social and physical properties of the environment (also see the General Introduction, p. 11-14; 19-22).

Capturing the nuance of cognitive development

Cognitive development and cognitive insights are thus fuzzy processes, which emerge from what children do, whereby their actions are nested in an environment with specific characteristics. Capturing cognitive development in all its typical nuance, fuzziness, and nestedness requires specific methods, which are grounded in the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances - which I did. Most of these methods had never been applied to study children's (and students') hand movements, gestures, and speech during (science and technology) tasks.

To be specific, Study 1 (Chapter 2) was the first study to apply Cross RQA to investigate the coupling between gestures and speech. Moreover, it was the first study ever to apply anisotropic Cross RQA to examine the directionality of coupling between two systems. Applying anisotropic Cross RQA showed that speech typically attracts gestures more than vice versa, which would not have been found using other methods. Furthermore, Study 2 (Chapter 3) was the first study to simultaneously investigate gesture-speech synchronization in terms of temporal alignment, semantic similarity, and complexity matching, whereby complexity matching between gestures and speech had never been examined before. This led to the counterintuitive finding of more complexity matching between gestures and speech in the difficult than in the easy condition. To my knowledge, Study 3 (Chapter 4) was the first study which *empirically* applied a new categorical RQA measure for system's variability, namely block entropy. Using this RQA measure, I found changes in variability which were not picked up by another measure for variability that did not include temporal order. Lastly, Study 4 (Chapter 5) was the first study to simultaneously research interpersonal coordination between speech, hand movements, and head movements during collaborative problem solving of dyads of children. I found that the timescales of interpersonal coordination between children's speech, hand movements, and head movements are similar but less pronounced than in adults, while the degree of interpersonal coordination between children did not seem to influence task performance.

In conclusion, applying methods which are grounded in the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances has yielded new as well as counterintuitive and nuanced findings. Furthermore, the studies in this dissertation show that these methods are applicable to research and capture a whole range of change processes in (cognitive) development and particularly its relation to hand movements and speech, including the nuanced and fuzzy processes that are typical for development, which future research could benefit from.

Different perspectives on children's hand movements and speech within cognitive development

Although the studies in this dissertation yielded new understanding of children's hand movements and speech within cognitive development, a number of important areas of study are also missing. In the next sections, I will focus on the most important ones, and discuss possible venues for future research.

First, investigating children's hand movements, gestures, and speech from the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances, meant that I focused on temporal patterns in this dissertation, instead of qualitative descriptions of children's behavior. However, connecting findings on the level of temporal patterns with findings on a qualitative level will yield additional and more pronounced understanding of the processes, coordination patterns, attractors, and relations between children and their environment, that underlie cognitive development. For example, such an approach would make it possible to empirically address my proposal that gesture-speech mismatches are (merely) a reflection of hand movements and speech being variable around cognitive insights. This could entail 1) capturing the intensity of children's and experimenters' hand movements and speech over time, and investigate hand movements' and speech's variability and coupling, 2) detailed coding of the timing *and* semantic content of children's and experimenter's hand movements and speech on multiple timescales (i.e. word by word, movement by movement, utterance by utterance, different strategies), and 3) a detailed description of the tasks' physical and social structure. Next to providing important understanding about cognitive development, I think qualitative descriptions are essential for bridging between quantitative findings and educational practice.

Second, while I focused on children's hand movements, gestures and speech within experimental and controlled tasks with a duration of 10 to 20 minutes, cognitive development extends far beyond these limited settings. If we are taking the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances, and the nested timescales which are inherent to these perspectives, seriously, we need to complement experimental studies with studies in which we investigate children longitudinally, with dense and diverse sampling of datapoints, and in a range of naturalistic settings (Adolph, 2019). In light of this, technological advances, such as OpenPose (Cao et al., 2021), DeepLabCut (Mathis et al., 2018), and wearables are promising. Such automated capturing and analysis techniques and devices allow accessible, cheap, and non-intrusive measuring of many different behaviors, such as movements, speech, heart rate, location, etc. (Developmental) Psychologists should profit from this, as these techniques and devices open up new possibilities to bring the methods we use to study behavior over time out of the lab and into the real world. Furthermore, I think studying children longitudinally and in naturalistic settings, using the technological advances that I just described, is essential for research findings to be meaningful in educational practice.

Lastly, in none of the studies did I address why children would want to learn something new in the first place. Moreover, this crucial issue is often not taken into account in studies about cognitive development, as well as in most studies which are grounded in the perspectives of complex dynamical systems, coordination dynamics, and affordances. However, as a mother observing and interacting with my children, it is impossible to overlook the importance of things like joy, curiosity, laughing, and mutual reinforcement for things my children learn. From the perspective of affordances, affordances are framed as possibilities for action, which enable and constrain, but *not* determine, what children (or other animals) do (e.g. Withagen et al., 2017). This leaves room for children to be motivated, at least to some degree, to engage with either one affordance or another. In the words by Withagen et al. (2017), the child is an agent, that is, an individual who is able to affect the degree to which an affordance has an effect on them. Furthermore, Wagman et al. (2016) propose that at any given time, many, hierarchically nested, affordances are present for any given individual. This could imply that researchers, parents, or teachers, can modulate this hierarchy of nested affordances, by means of shaping the physical and social characteristics of the environment, and thereby affording individual children to learn. We need more research to establish how we could do this for individual children at any point in time.

Practical implications for education

In line with the process of cognitive development, the practical implications of this dissertation are not straightforward (as they hardly ever are). However, I have some modest things to share with people involved in educational practice.

I found that children and students (hereafter: children) use both their hands, speech, and even move their heads according to the social and physical structure of the environment when they engage in (science and technology) tasks. Furthermore, I found that children's hand movements, gestures and speech are tightly and continuously coupled. I think it is important to be aware of this continuous coupling between children's modalities, and to allow children to communicate in such multimodal ways, and not only attend to, or favor, their speech. In addition, I found differences between kindergartners and first graders with regard to coupling between gestures and speech, which suggest a potential age difference in multimodal communication. This age difference could be important to take into account while teaching. Besides being coupled to each other, children's hand movements, gestures and speech are also attuned to the physical and social environment. Emphasizing particular properties of a task, either by means of the physical or social structure that is provided, influences children's hand movements, gestures, and speech, and thereby their cognitive development. Lastly, the findings in the studies show that cognitive development is not linear or straightforward, but instead is a nuanced, complex, fuzzy process, comprised of and nested within many other processes. This implies that children both need space to explore and try things out, using the parts of their body most suited for the job, as well as functional constraints as provided by the social and physical structure of the task.

I genuinely hope that, one day, I will have the chance to discuss these rather general practical implications with teachers. I would love to devise a plan *together* for how to translate these implications into an applied research program, deemed useful by teachers themselves (and yes, this is an invitation!)