How hand movements and speech tip the balance in cognitive development

de Jonge-Hoekstra, Lisette

DOI:
10.33612/diss.172252039

Publication date:
2021

1

General Introduction

General Introduction

When we educate children in primary education, we aim to induce development and teach them something new. While this development in children spans many, equally important, developmental areas, we tend to emphasize *cognitive development*. Inducing cognitive development requires educators and children to communicate. In other words, educators and children need to speak, gesture, and move together for cognitive development to happen (e.g. Novack & Goldin-Meadow, 2015; Pennings et al., 2018; van de Pol et al., 2010; Van der Steen et al., 2012). Moreover, children's hand movements in general, and gestures in specific, have been found to lead cognitive development, over speech (e.g. Adolph et al., 2015; Adolph & Franchak, 2017; Adolph & Kretch, 2015; Alibali & Goldin-Meadow, 1993a; Church & Goldin-Meadow, 1986; Fischer & Bidell, 2006; Goldin-Meadow et al., 1992, 1993; Perry et al., 1992; Roth, 2002).

Existing explanations for gestures' leading role in cognitive development center around abstract concepts, such as implicit and explicit knowledge (Broaders et al., 2007), conflicting cognitive representations (Church & Goldin-Meadow, 1986; Goldin-Meadow et al., 1993; Perry et al., 1992), and cognitive load (Cook et al., 2012; Melinger & Kita, 2007). However, these explanations disregard that moving your hands and speaking is not (only) abstract. Instead, hand movements and speech are 1) actions which involve many physical components at many scales which interact over time, 2) physically coupled to each other, and 3) related and adapted to the physical and social environment.

These three characteristics of hand movements and speech are related to complex dynamical systems, coordination dynamics, and affordances, respectively, which are the theoretical grounds on which this dissertation is build. Previous research from these theoretical perspectives has yielded crucial understanding about diverse areas of child development and skill acquisition (e.g. Adolph et al., 2018; Gibson & Pick, 2000; Smith & Thelen, 2003; Thelen et al., 1987; van Geert, 2008). My goal in this dissertation, based on these theoretical perspectives, is 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 scales-, and how their hand movements and speech relate to each other, and to the physical and social environment.

In this General Introduction, after giving a brief overview of hand movements, speech, and cognition in development, I will introduce the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances. As it is entirely possible to write whole books with sophisticated detail and mathematical precision about either of the topics that I address in the General Introduction (e.g. E. J. Gibson & Pick, 2000; J. J. Gibson, 1966; Kelso, 1995; Thelen & Smith, 1994; Van Geert, 2008), I will focus on the core ideas of these perspectives.

Hand movement and speech in (cognitive) development

The relation between hand movements and vocalizations starts early in life. Even long before birth, human fetuses have been shown to suck their fingers at 10 to 15 weeks of gestation (e.g. de Vries et al., 1982), and to coordinate hand moving and mouth opening at 19 to 35 weeks of gestation (e.g. Myowa-Yamakoshi & Takeshita, 2006). After birth, the frequency of this hand-mouth-coordination sharply increases (Butterworth & Hopkins, 1988; Sparling et al., 1999). Throughout the first year of life, infants' hand-mouth-coordination differentiates into new patterns, such as bringing objects to their mouth to explore them orally, rhythmical manual banging and vocal babbling, and pointing gestures and saying their first word (for an overview, see Adolph & Franchak, 2017; Iverson & Thelen, 1999). In line with differentiation into new and more patterns, Abney, Warlaumont et al., (2014) found both hand movements and vocalizations of one infant to become more flexible and context-dependent over time, from 51 to 305 days of age. In particular, changes in the variability of the infant's hand movements and vocalizations were related. These early couplings between infants' hand movements and vocalizations provide the basis for more adult-like gestures and speech in communication (Iverson, 2010; Iverson & Fagan, 2004; Iverson & Thelen, 1999).

Hand movements and speech continue to develop hand in hand, also after the first year of life. The coordination between children's hand movements – and body movements in general – and speech is pivotal for prosody development, i.e. how children learn speech rhythm and intonation (Esteve-Gibert & Guellaï, 2018). With regard to semantic development, children's first pointing precedes saying their first word, whereby children who were early pointers also tend to be early speakers (for a review, see Goldin-Meadow & Alibali, 2013). Furthermore, the moment of children's first gesture + word combinations predicts the moment of their first word + word combinations (Iverson & Goldin-Meadow, 2005). While children's pointing initially is accompanied by some form of speech for only 40% of the time, gestures predominantly occur together with speech (Esteve-Gibert & Prieto, 2014) after a couple of months. This pattern, of hand movements leading and speech "catching up", also extends to cognitive development more generally.

With regard to cognitive development, children use their hands to explore and gesture about the world around them (Adolph & Franchak, 2017; Adolph & Kretch, 2015; Chapter 4 of this dissertation). Children (and adults) from all ages reach for objects that interest them, and feel and manipulate these objects using their hands, in ways they are unable to do by speaking. Within primary education, hands-on learning activities also rely on such manual exploration (Fischer & Bidell, 2006; Roth, 2002). When children talk about objects, they also gesture and thereby extend their array of manual action (Roth, 2002). Encouraging children to gesture while they reason about something they do not yet understand, such as conservation problems or mathematical equivalence problems, fosters their understanding (Broaders et al., 2007), particularly when children are instructed to shape these gestures according to relevant task properties (Brooks & Goldin-Meadow, 2015; Goldin-Meadow et al., 2009).

Moreover, children as young as 5 years old have been shown to convey their “new” understanding in gestures, while simultaneously putting their “old” understanding into words (Church & Goldin-Meadow, 1986; Pine et al., 2004). For example, in the context of a liquid conservation task, a child may still say that one glass contains more water because the level of water is higher (i.e. old understanding = only taking the water level into account), while simultaneously make a C-shape with their hand to indicate the width of the glass in gestures (i.e. new understanding = also taking the width of the glass into account). This phenomenon has been called a gesture-speech mismatch. However, we still grapple to understand how these gesture-speech mismatches fit with, and could originate from, an integrated and tightly coordinated gesture-speech system (Koschmann, 2017; Pouw et al., 2017). To better understand why hand movements in general, and gestures in specific, are leading over speech in cognitive development, we investigate them from a complex dynamical systems perspective in this dissertation.

Complex dynamical systems

Complex dynamical systems are systems that consist of multiple components, typically at multiple scales of a system, which interact and spontaneously coordinate over time by means of self-organization (e.g. Kelso, 1995; Smith & Thelen, 2003; Thelen & Smith, 2007; Van Geert, 1998; Van Geert, 2008; Van Orden et al., 2003; Van Geert, 2019). Examples of complex dynamical systems are weather systems, ant colonies, and the stock market, to name a few. Due to the interactions between components, a complex dynamical system is a whole greater than the sum of its parts. More specifically, interacting components self- organize into global patterns, whereby new patterns emerge. For example, weather systems self-organize into hurricanes, ant colonies self-organize into hyper-efficient trails to bring food into their nest, and the stock market self-organizes into sudden recessions (see Figure 1). Such global patterns, also known as attractors or collective states, are relatively stable, thus tending to resist perturbations – at least to a certain degree. During such stable states, the coupling between a system’s components is strong.

Albeit relatively stable, changes from one stable state to another can occur. This is characterized by a reorganization of a system’s components and their relations. For example, hurricanes tend to dissolve above land, ant colonies reorganize into different trails when they find new food sources, and stock markets reorganize into growth after a recession. Such a

Figure 1. Examples of self-organization in complex dynamical systems. Panel a shows a hurricane, panel b shows an ant trail, and panel c shows a recession.

system reorganization and transition is typically accompanied by an increase of variability in the system's behavior, when the coupling between components weakens. Such an increase in variability is seen as a hallmark of change.

Attractor landscape

The stability of stable states, or attractors, and the variability surrounding transitions between them can be metaphorically described using an attractor landscape (see Figure 2)[1]. In this landscape, an attractor is depicted as a well with a certain width and depth. Furthermore, there can be one or multiple wells, corresponding to the existence of an equal number of attractors. More attractors typically indicate that a system is capable of adapting to different circumstances. For example, if someone knows multiple ways to bring food to their mouth, they can adaptively use one to eat either soup or chocolate.

With regard to stability and variability of attractors, one can imagine what would happen with a ball rolling across the landscape. If a well is wide, the chance that the ball rolls in the well is relatively large, as compared to a narrow well. Analogously, some attractors are relatively stronger than others. For instance, when we were on a holiday, my daughter took off to the playground, which was to the right, about 50 times per day. However, when she needed to go to the bathroom, which was to the left, she still would take off to the right, indicating that the running-towards-the-playground-attractor was relatively strong.

Furthermore, if a well is deep, the chance that the ball will get out of the well is relatively small, as compared to a shallow well. This analogy corresponds to some attractors being more stable, or more resistant to perturbations, than others. With regard to the previous example of running

[1] It should be noted that the attractor landscape is only capable of describing a particular type of attractor, namely point attractors. Many other attractors also exist, but explaining them in detail would go beyond the scope of this General Introduction. For beautiful pictures, one can search the internet for "strange attractors".

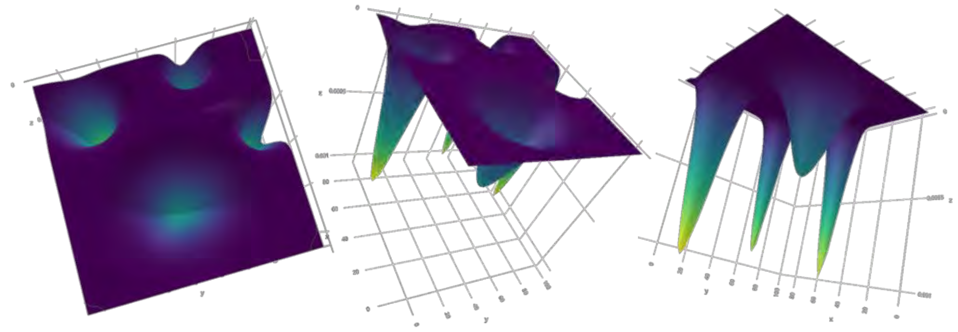

Figure 2. Example of an attractor landscape with four wells, or attractors. These wells differ in width and depth, which analogously correspond to attractors with a different strength and stability, respectively.

towards the playground, my daughter eventually took a U-turn and ran towards the bathroom. This indicates that the she did not get stuck running towards the playground, and the attractor thus was not particularly stable.

Lastly, the attractor landscape changes over time, with some wells becoming wider or deeper, while other wells appear or disappear, in correspondence with what can happen with real attractors. Again returning to the previous example, over the course of a couple of days, I noticed that the time it took my daughter to make the U-turn towards the bathroom became less and less, and eventually she immediately ran left towards the bathroom and right towards the playground. In other words, next to the running-towards-the-playground-attractor also a running-towards-the-bathroom-attractor had emerged.

People as complex dynamical systems

As implied by the examples in the previous section, people are complex dynamical systems as well, as they also consist of multiple components at multiple scales which interact over time (e.g. Kelso, 1995; Smith & Thelen, 2003; Thelen & Smith, 2007; Van Geert, 1998; Van Geert, 2008; Van Orden et al., 2003; Van Geert, 2019). For example, people consist of different types of cells, which self-organize into different structures (systems) such as bones, muscles, blood vessels, or brain parts. These structures are self-organized in larger structures such as the skeleton, muscular system, circulatory system, or central nervous system, and these larger structures themselves are all self-organized in a coherent human being. We can scale this example even further up to larger systems (e.g., people are part of a family, which is part of a community), as well as identify intermediate scales, such as the musculoskeletal system, but also the cognitive system (Thelen & Smith, 1994; Thelen & Smith, 2007), or the gesture-system and the speech-system (Iverson & Thelen, 1999; Rusiewicz & Esteve-Gibert, 2018). Just like any complex dynamical system, people's components are coupled at many different scales and self-organize into coherent wholes and stable states, which is dazzlingly complex to realize.

Furthermore, similar to other complex dynamical systems, also people show a destabilization and increase in (various forms of; see previous section) variability upon changes from one stable state to another (e.g., Adolph et al., 2015; Bassano & van Geert, 2007; Shockley et al., 2002; Van der Maas & Molenaar, 1992; Van Geert & Van Dijk, 2002). For example, participants show an increase in hand movement variability upon discovering a new cognitive strategy (Stephen et al., 2009). Moreover, critical fluctuations precede large shifts in symptom severity of patients with a mood disorder (Olthof, Hasselman, Strunk, van Rooij, et al., 2020), and destabilization of self-ratings is related to better intervention outcomes (Olthof, Hasselman, Strunk, Aas, et al., 2020). A last example is that an increase in variability of utterance length precedes structural changes in language development (Bassano & van Geert, 2007; Van Dijk & van Geert, 2011). How change exactly arises from multiple components when they are coupled and "work together", is the topic of the field of *coordination dynamics*.

Coordination dynamics

Coordination is everywhere. Examples of coordination include a couple dancing the tango, an acrobat juggling with six balls, male fireflies synchronizing their flashes to attract female fireflies, a baby learning to walk (or talk, etc.), Usain Bolt sprinting towards victory, bees running their hive together, or car drivers slowing down and accelerating together during a traffic jam. In all these examples, two or more things are *coupled*, be it physically and/or perceptually. Due to this coupling they adjust their actions to each other, and their behavior becomes coordinated. Moreover, often these coupled systems and their components behave as if they were one – a synergy (e.g. Haken, 1987; Kelso, 2013; Latash, 2008; Strogatz, 2012; Turvey, 2007; also see Warren, 2006).

A synergy is a functional grouping of systems that "work together" and self-organize in the service of a particular "goal" (Kelso, 2013; Latash, 2008; Turvey, 2007). In our previous examples, we can identify functional organizations, such as dancing, juggling, attracting female fireflies, walking, winning, running the hive, and driving somewhere while keeping the car in one piece, respectively. Within a synergy, fluctuations of one component are compensated for by fluctuations of other components, as to preserve the functional organization of the synergy. For example, if Usain Bolt steps on a stone with his foot, muscles in other parts of his body will compensate for this and he will still be able to maintain a stable running pattern, leading him to win the match. If babies step on a stone however, they still lack the ability to compensate for this fluctuation in one component, and they will probably fall. In other words, while the coordination of many components goes smoothly for Usain Bolt, this is not (yet) the case for the baby. Smooth coordination of many components is related to the problem of degrees-of-freedom in motor control.

The problem of degrees-of-freedom entails that any movement, no matter how big or small, entails the coordination of numerous diverse body components (Bernstein, 1967). For instance, uttering one syllable already involves the cooperation of more than 70 muscles (Turvey, 2007). In theory, the number of possible configurations (degrees-of-freedom) for each movement is astronomically large, and controlling each individual component that is involved in a movement would take an immense effort (hence the problem). However, instead of being individually controlled by some central command system, the components self-organize into collectives: Synergies (Haken, 1987; Kelso, 2013). Within a synergy, the degrees-of-freedom are compressed and the components are constrained to act as a functional unit. To maintain this functional unity, changes in degrees-of-freedom in one component of the synergy (e.g. Usain Bolt's foot) are compensated for by changes in degrees-of-freedom in other components of the synergy (e.g. Usain Bolt's muscles in his leg and back). In Kelso's words: "Retaining stability is, for a synergy, the retaining of functional integrity." (Kelso, 2013, p. 1541).

Behavior, however, not only shows stability, but also flexibility, and is adaptive to changing circumstances. That is, there are multiple attractors in an attractor landscape, from which the system can choose. Synergies are task specific and always ready to become something else in a moment. For example, writing this thesis involves the components of the neuromuscular system responsible for my hand movements to change between typing on a keyboard, moving the mouse, writing on paper, grabbing a coffee mug, and fidgeting my hair. This fits with degeneracy in complex dynamical systems, which means that multiple combinations of components can achieve one function, and one combination of components can achieve multiple functions (e.g. Edelman & Gally, 2001; Seifert et al., 2016; Whitacre, 2010). Furthermore, the coordination patterns between components differ for different functions. For example, typing on a keyboard involves different fingers of both my hands to engage in temporally and spatially tightly coordinated movement cycles of pressing and releasing keys (see Figure 3, left panel). Writing on paper, however, involves all the fingers of my right hand to

Figure 3. Coordinative patterns of typing on a keyboard (left panel) and hand writing (right panel).

engage in a cycling motion as well, albeit with a different spatial configuration for each finger (see Figure 3, right panel).

Following von Holst (1938), Kelso (2013) identifies three general patterns of coordination. During *absolute* coordination, components are locked in time - a pattern also known as *phase synchronization* (Pikovsky et al., 2001). The earlier example of fireflies' synchronous flashing illustrates absolute coordination. During *relative* coordination components become locked for some period of time and then unlock again, such as in the earlier example of car drivers slowing down and accelerating during a traffic jam. Lastly, components can go about independently, which is the case with *no* coordination. Furthermore, these three general patterns of coordination can also mix and coexist. In addition, more forms of coordination exist, which will be explained in Chapter 3 (Study 2). Changes between such coordination patterns are called *phase transitions*.

Phase transitions in human motor behavior have been extensively studied within rhythmic motor tasks, following the Haken-Kelso-Bunz (HKB) study paradigm (Haken et al., 1985). In bimanual coordination, people move their fingers in two distinct coordinative patterns at lower speeds: either *in phase* (or parallel; see Figure 4, left panel) or *anti phase* (or mirror; see Figure 4, right panel). However, when people increase their movement speed, they involuntarily switch from anti phase to in phase coordination at a certain threshold, while this is not the case for in phase coordination. In other words, at higher movement speeds only one coordinative pattern is possible, namely in phase coordination. Furthermore, when people lower their movement speed again, the threshold at which they switch back to anti phase coordination is lower than the threshold at which they switched to in phase coordination. This phenomenon is called hysteresis, and shows that coordinative patterns are dependent on what happened before, i.e. the history of the system. In terms of attractors, these findings have been modelled as two stable attractors at lower speeds: An in-phase and an anti-phase attractor. When the speed increases, the stability of the anti-phase attractor increasingly diminishes until it virtually disappears, and only the in-phase attractor exists.

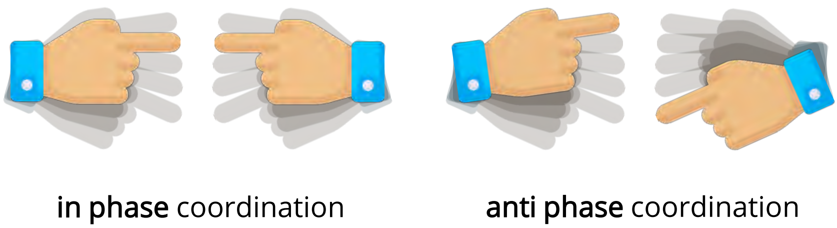

Figure 4. In phase (parallel) bimanual coordination pattern (left panel) and anti phase (mirror) bimanual coordination pattern (right panel) in the Haken-Kelso-Bunz study paradigm.

The findings from the HKB paradigm (i.e. inter-limb coordination), including transition and hysteresis phenomena, have been extended to coordination between manual and vocal actions (i.e. gesture-speech coordination; e.g. Treffner et al., 2008; Treffner & Peter, 2002) and to coordination between people (i.e. interpersonal coordination; e.g. Richardson et al., 2007; Schmidt & Richardson, 2008). Interestingly, Richardson et al. (2007) found that directly looking at each other while rocking in rocking chairs resulted in more stable interpersonal movement coordination, compared to peripherally seeing each other. In other words, *perception-action couplings* between people modulates their movements, and thus modulates the coordinative patterns that arise between them. (e.g. E. J. Gibson & Pick, 2000; J. J. Gibson, 1966; Marsh et al., 2009; Warren, 2006). Not only does this perception-action coupling allow us (and other animals) to adapt our actions to each other, but it also allows us to adapt our actions to our physical surroundings, which is captured by the concept of *affordances*.

Affordances – the match with the environment

Affordances are possibilities for action which the environment offers to the animal, thereby matching its capabilities (e.g. Adolph & Kretch, 2015; E. J. Gibson & Pick, 2000; J. J. Gibson, 1966, 1979). Given the importance of the match between animal and environment, animals (including humans) should be considered within their natural surroundings, doing the things they normally do. Moreover, animals and their surroundings are reciprocal. That is, animals adjust their actions to the environment and the environment offers possibilities for action accordingly. Furthermore, the environment provides information that specifies these action possibilities, which animals attune to and use to guide their actions.

Perception and action thus are reciprocal too, which is known as the *perception-action loop* (e.g. Adolph & Kretch, 2015; E. J. Gibson & Pick, 2000; J. J. Gibson, 1966, 1979). First, perceiving means that the animal *actively gathers information* about things and events in the environment, by means of looking, listening, feeling, tasting, and moving and manipulating the environment to optimize the information. Crucially, the information that flows through the senses is *rich*, in the sense that the energy in the form of light, sound, pressure, and chemicals is structured by objects and events in the environment in a specifying manner. Second, perception informs the animal about the actions it can perform with the objects and about what to expect in the environment (called prospection), while the animal's own movements are informative about its (changing) relation to the environment.

To give an example of perception-action reciprocity in relation to the structure of the environment: When Usain Bolt runs through an area covered with obstacles, he will move and turn his head and body to change his angle of approach towards each obstacle on his way. These changes in visual angle enable Usain Bolt to regulate perception of distance between him and the obstacles, their size, and time-to-contact, thereby enabling him to avoid them. On the other hand, a young toddler going through the same area will also turn their head and body to optimize perception, but the exact movements will be very different. These different movements are due to differences in the toddler's size, speed, strength, visual accuracy, and motor coordination, compared to Usain Bolt. Moreover, these differences in body capacities also make that the toddler will probably not run through the area and avoid obstacles, but will rather climb them, or walk around them, among other things that the toddler can and would like to do (see also next paragraph). In conclusion, affordances are animal-specific, which means that they depend on the match between an individual's bodily scales and action capabilities and the properties of the environment (e.g. Fajen et al., 2009).

Children need to learn to perceive and realize affordances (possibilities for action; Adolph & Kretch, 2015; E. J. Gibson & Pick, 2000). For example, while a couch affords sitting for an adult, it affords pulling up to stand for a baby (see Figure 5). Throughout development, children need to learn what their body can do and what the environment has to offer to make use of this. By means of exploration, such as a toddler doing different things on and around obstacles, children become increasingly better at attuning to the relevant information specifying the possibilities for action in a given situation. This is a lifelong process, whereby the match between a growing body and increasing action repertoire constantly changes, and new opportunities for actions in the environment continue to arise. Importantly, these new action opportunities, in turn, provide new things to be explored and new skills to be learned. For example, sitting requires strong core muscles to keep the torso stable. When that has been mastered, the child's hands free up, which gives them the possibility to reach for and grasp objects. With this new skill, the child can make all kinds of wonderful discoveries.

Figure 5. Affordances of a couch for an adult (sitting; left panel) and for a baby (pulling up; right panel).

Affordances and structure of the environment are also apparent in children's hand movements and speech, and social interactions. Regarding hand movements, children's hands are crucial to learning about affordances, especially for objects that require fine motor skills to handle (Adolph, 2019; Adolph & Franchak, 2017). Children use their hands to explore these objects: They feel its surface structure, size and weight. Furthermore, they pick it up and turn it in order to see it from different angles, hear the sounds the object is making, and put it in their mouth to taste it and explore its texture. In addition, they 'use' the objects to explore the surfaces around them, by bouncing on them or by using an object to change something in their surroundings (i.e. tool use; Lockman, 2000; Smitsman & Bongers, 2003). Gestures are also hand movements, whereby gestures can be thought of as moving one's hands according to the rhythmic structure of speech (e.g. Wagner et al., 2014, also see Pouw et al., 2018), as well as according to the spatial structure of objects and events in the environment.

With regard to vocalizations and speech, from early on vocalizations (e.g. crying) are very effective to elicit or stop someone else's actions within particular situations. Furthermore, babies very quickly learn that making sounds, such as cooing, captures their caregiver's attention for longer periods of time (e.g. Jaffe et al., 2001). During these interactions, caregivers actively and voluntary as well as involuntary structure children's vocalization patterns (e.g. reacting, turn taking, mimicking), and at the same time over-emphasize the relevant acoustic structure of their own speech (e.g. so-called *motherese*) (e.g. Stern et al., 1983). In other words, embedded within everyday social interactions, children learn to mutually and adaptively structure their vocal sounds on many levels with their interactions partners (e.g. Reed, 1995; van Dijk et al., 2013). Speaking thereby opens up many new possibilities for action together with other people, such as collaborating, sharing thoughts and feelings, and teaching and learning about cognitive tasks, which extend to both the past, present, and future (e.g. Smith & Gasser, 2005).

Cognitive development from the perspective of complex dynamical systems, coordination dynamics, and affordances

A recent review (Adolph & Hoch, 2019; also see Adolph, 2019; Adolph et al., 2018; Newen et al., 2018) summarizes the characteristics of motor development as embodied, embedded, enculturated and enabling. Embodied refers to the fact that the current specifics of the body determine possibilities for action, embedded implies that the environment opens up and constrains possibilities for action, enculturated indicates that motor development is shaped by social and cultural forces, and enabling means that each new skill opens up a whole new range of opportunities to learn other skills, and thereby can bring about a developmental cascade. This echoes the descriptions already given above about new possibilities for action, which continuously arise throughout development. Following previous researchers (e.g. Kloos & Van Orden, 2009; Thelen & Smith, 1994; Thelen & Smith, 2007), I am convinced that these characteristics of motor development also apply to cognitive development in general.

Based on the framework above, cognitive understanding within cognitive development is the equivalent of what a motor skill is within motor development. This entails that cognitive understanding is a functional coordination pattern too, similar to motor skills. Functional hereby means that it arises when a particular child is in a particular physical and social environment, such as when an adult asks them to explain about a particular task (see e.g. Study 1 and 3/Chapter 2 and 4). Depending on the specifics of the environment, cognitive understanding can take many forms, such as talking and gesturing, but also writing on paper, or hands-on problem solving. Similarly, also motor skills come in many different forms, such as walking, running, climbing, or swimming, depending on the environment that someone is in and the particular motor problem one is confronted with, such as moving on a horizontal surface, a slanting or vertical surface, or in the water, respectively. This suggests that any form of cognitive understanding, just like any particular form of motor skills, only exists for a specific child doing a concrete task in a specific environment.

Opponents of such a view typically emphasize that viewing cognitive understanding about a particular concept as being similar to a motor skill ignores that cognitive understanding, at least in part, is abstract, symbolic, disembodied and ungrounded. This expresses that cognitive understanding about a particular task, once it is well-developed, is supposed to happen "in someone's head", and is thereby relatively independent from the specific environment that someone is in or in which the understanding emerged. However, I would like to challenge the idea that a motor skill is any less (or more) abstract or "in someone's head" than cognitive understanding about a particular task, using the example of swimming.

Few people would disagree that swimming is a motor skill that depends just as much on the specifics of the environment as that it depends on a person's capability to adjust to that in a very typical way. This specific environment is a pool of water large enough for a person to move about in. Swimming on land is, strictly speaking, impossible, because the resistance of air is much lower than the resistance of water, while a floor, on the other hand, is much too resistant. Furthermore, flapping your arms and legs in the air in a pattern that looks like swimming will not get you anywhere and is thus not functional (unless your aim was to make other people laugh). Swimming thus only exists and can be concretely defined in the water. In addition, learning to swim entails learning to coordinate many components of your body so that you stay afloat and move forwards or backwards while being in the water. When you have learned to swim, we expect you to be able to swim whenever you are in the water. However, when you are not in the water and are thus not swimming, we do not think that you are not a skillful swimmer anymore. We typically do not ask “where your skill of swimming went”. No one considers it to be *abstract* or *in your head*, when you are not in the water.

Similar to swimming, cognitive understanding about a particular task only exists and can be concretely defined when a child is in a particular physical and social environment. For example, talking and gesturing about balance scale problems (see also Study 3 and 4, Chapter 4 and 5, respectively) only happens when a child is in a situation in which a balance scale and weights are present and an adult asks them to explain about balance scale problems. If a child would do a similar coordination pattern while playing hide and seek, this would give away their location, and would thus not be functional. Furthermore, having learned to correctly (from the perspective of the adult) explain about balance scale problems entails paying attention to, speaking, and gesturing about both mass of the weights and distance from the fulcrum whenever a child is in a situation that requires them to do so. This is thus similar to a skilled swimmer being able to swim whenever they are in the water. I therefore assert that asking “where the cognitive understanding about balance scale problems went” when a child is not in that particular situation is just as meaningful, or rather meaningless, as asking “where the skill of swimming went”.

One last counterargument, which is in favor of cognitive understanding being fundamentally different from motor skills, is that cognitive understanding about a particular task transfers to many other situations, while this is not the case for motor skills. However, this argument disregards that the ability to adaptively use a motor skill in an increasing number of diverse situations is inherent to learning a motor skill (e.g. Adolph, 2019; Adolph et al., 2018; Adolph & Hoch, 2019). With regard to the previous example of swimming, while children typically learn to swim in calm waters, such as a swimming pool, later on they will learn to swim in water with waves, or currents, such as in a sea or river. On the other hand, adverse circumstances, such as heavy clothing or stormy waters, will make swimming impossible for even the most skilled swimmers.

Moreover, cognitive understanding is known to be grounded and highly sensitive to environmental circumstances. I will illustrate this with the famous example of the A-not-B error (see Figure 6). The A-not-B error pertains to a classical Piagetian task, in which a toy is repeatedly hidden at a location A (the A-trials), where the child subsequently and correctly finds the toy. After a number of A-trials, the toy is hidden at location B. Children between 7 to 12 months old have been found to continue searching at location A, instead of location B. This has been coined as the A-not-B error (Piaget, 1954). Piaget attributed the error to the idea that children at that age have not yet developed the concept of *object permanence*. However, in a series of studies,

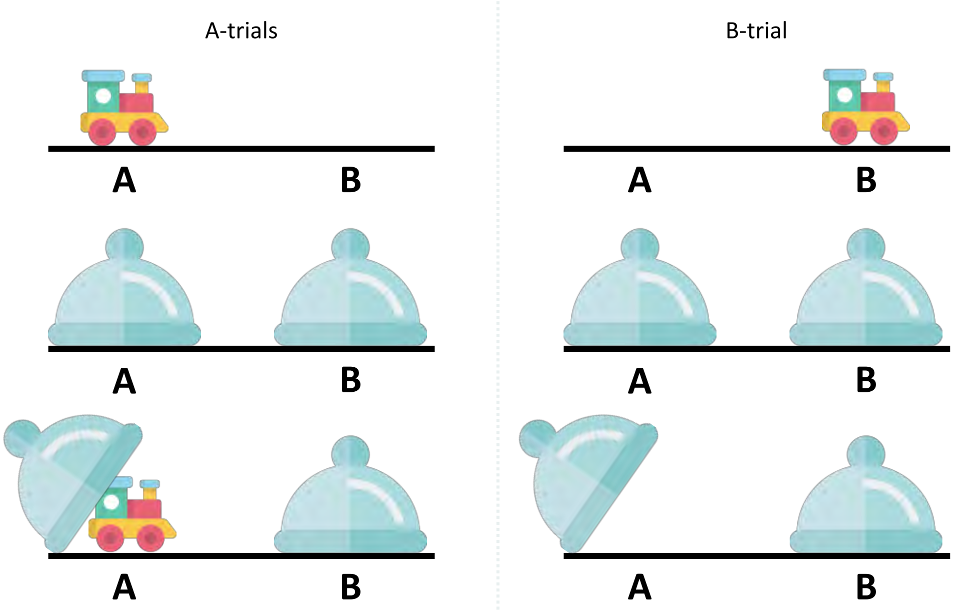

Figure 6. Visualization of the A-not-B-error task. In this task, initially a toy is repeatedly hidden at location A (the A-trials). The child correctly finds the toy at location A. After a number of A-trials, the toy is hidden at location B instead of location A (the B-trial). Yet the child still continuous to search for the toy at location A.

inspired by complex dynamical systems theory, Smith et al., (1999), Spencer et al. (2001), Thelen et al. (2001), Schöner and Thelen (2006), and Schöner and Dineva (2007) showed that particular circumstances make the A-not-B error disappear in 10 month old children, while other circumstances elicit the A-not-B error in older children. To be specific, a salient visual difference between the locations, as well as a change in posture (i.e. sitting vs standing) made younger children correctly search at location B during B-trials (Smith et al., 1999), for example. Furthermore, a longer waiting time between hiding the toy at location B and searching for the toy elicited the A-not-B error in children who were older dan 12 months (Spencer et al., 2001). This example of the A-not-B error again shows that the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances are just as useful for capturing cognitive development as they are for capturing motor development.

This dissertation

I started this General Introduction with describing that hand movements in general, and gestures in specific, have been found to lead cognitive development in children, over speech. Furthermore, I highlighted that a satisfying explanation for this phenomenon has been lacking, both from the perspective of gestures' and speech's typically tight integration, as well as from the theoretical perspectives of complex dynamical systems, coordination dynamics, and affordances. Throughout the General Introduction, I showed how these three theoretical perspectives are very powerful in explaining a broad range of phenomena in many diverse systems and organisms, including human behavior, (cognitive) development and skill acquisition. As described before, my goal in this dissertation, based on these theoretical perspectives, is 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 scales-, and how their hand movements and speech relate to each other, and to the physical and social environment. By researching cognitive development in children, I will move beyond the topics which traditionally have been investigated from these perspectives, such as early motor development, and motor coordination, hereby following the footsteps of many inspiring researchers before me (e.g. Stephen et al., 2009; Thelen & Smith, 2007; Van Geert, 2019). Together with my supervisors and several collaborators I carried out four studies.

In Study 1 (Chapter 2) we investigated the **stability and variability of the coupling between children's gestures and speech**, in terms of level of understanding during a hands-on Science & Technology task, which children did together with an adult who provided support. We also investigated how these within-task measures of gesture-speech coupling predicted general measures of cognitive performance.

In Study 2 (Chapter 3) we investigated students' **gesture-speech synchronization** in an easy and a difficult cognitive task. We specifically researched gesture-speech synchronization in terms of temporal alignment (phase synchronization), semantic similarity (gesture-speech mismatches), and complexity matching (multiscale synchronization).

In Study 3 (Chapter 4) we investigated how children performed hands-on Science & Technology tasks with **different spatiotemporal** properties, and how these different properties of the environment were related to **differences between children's variability of hand movements and speech**. We conceptualized variability in terms of Diversity and Complexity.

In Study 4 (Chapter 5) we investigated how **dyads of children coordinate their speech, hand movements and head movements**, when they solve cognitive problems **together**. We researched the coherence and relative phase angle (which informs about leader- and follower-patterns, and in- and anti-phase coordination) of dyads' speech, hand movements and head movements at multiple timescales, and analyzed whether these measures predicted task performance and dyadic agreement.

In the **General Discussion** (Chapter 6) I will discuss what these studies have contributed to my aim of understanding of how children move their hands and speak during cognitive tasks, and how their hand movements and speech relate to each other, to the physical and social environment, and to cognitive development. Furthermore, I will address what our findings mean in light of the more theorical backgrounds of complex systems, coordination dynamics, and affordances. Lastly, I will discuss what our findings mean for educating children in primary education and inducing cognitive development.