Article

Scaffolding Embodied Access for Categorization in Interactions between a Blind Child and Her Mother

Carolyn Rickard, Mara Strother, Barbara A. Fox (ID) **and Chase Wesley Raymond** *

Department of Linguistics, University of Colorado, Boulder, CO 80309, USA; Carolyn.Rickard@colorado.edu (C.R.); Mara.Strother@colorado.edu (M.S.); Barbara.Fox@colorado.edu (B.A.F.)
* Correspondence: Chase.Raymond@colorado.edu

check for updates

Abstract: During language acquisition, sighted children have immediate and temporally stable access to the 'gestalt' of an object, including particular features that suggest its categorization as part of a class of objects. Blind children, however, must effectively and productively constitute the whole object from its constitutive parts in order to categorize them. While prior studies have suggested that varied experience and appropriate sensory access can contribute to this process, little attention has been given to how this is accomplished. The present study aims to address this issue by using conversation analysis to explore embodied understanding and categorization work between a 26-month-old congenitally blind child and her sighted mother as they play with various animal toys. Here we provide an analysis of a segment involving a particular toy (a cow plush), and ask two questions: (1) During play, how does Mother scaffold embodied routines for the identification of criterial information about a category, and (2) How is knowledge of varied exemplars, not directly accessible within the current activity, then made available to the child? Detailed examination of the linguistic and embodied practices employed by this mother–child dyad provides a concrete example of how non-visual modalities help scaffold the learning of categorization techniques, as well as illustrates the import that the examination of naturally occurring social interaction can have for theories of language and embodied cognition.

Keywords: categorization; acquisition; situated interaction; conversation analysis; embodiment; multimodality

1. Introduction

Interacting appropriately with objects fundamentally depends on humans' ability to categorize (Borghi 2005). Traditional cognitivist views posit that categorical knowledge consists of propositional symbols that are arbitrarily related to their referents, and that such knowledge is represented and manipulated in brain-specific computational modules in an amodal fashion, separated from any sensorimotor activity (de Vega et al. 2008). In contrast, theories of embodied cognition propose that knowledge is grounded in the human body and its interaction with the environment, with meaning emerging through collaborative, situated sensorimotor processes. As Semin and Cacioppo (2008, p. 121) put it, cognition is embodied in the sense that it "consists of affective, cognitive, behavioral, and neurophysiological processes whose organization and function within brains and bodies are promoted by an analysis *across* brains and bodies; and is manifested in synchronization, coordination, and co-regulation of behaviors" (emphasis added).

According to this embodied cognitive perspective, then, the way in which the world is understood is directly linked to perception and action (de Vega et al. 2008; Pecher and Zwaan 2005): When bodies act in different ways, they create the conditions for detecting affordances (the actions suggested by a

particular object), as well as discovering cross-modal associations that facilitate organizing the world and which thereby shape future action (Edelman 1992; Gibson 1979; Zukow-Goldring and Ferko 1994).

While research on embodied cognition has focused primarily on the ways in which bodily actions and perceptions interact with mental representations, the current study explores the links between embodiment and cognition in real-time interaction. Work in conversation analysis (CA) has illustrated the importance of the connection between language and the body, particularly with regard to how interactants build co-operative actions in making sense of the world (Goodwin 2018; Goodwin and Cekaite 2018; Streeck et al. 2011). Through a detailed examination of a blind child and her mother engaging in toy play (see Section 2), we will describe how this dyad makes use of multimodal practices within a shared environment, with each attending to how the other is interpreting (perceiving) and operating (acting) on the objects at hand. In this way, the participants scaffold access (Vygotsky 1978) to *critical features* (Section 3.1) and *varied exemplars* (Section 3.2) crucial to the understanding of categories in the world. On the basis of these data, we posit that it is the collaborative production of such embodied, co-operative actions—that is, the integrated use of language and the physical body to make sense of, and in, actual situated interaction—that provides for the development of embodied understandings and categorizations.

The analysis that we offer seeks to build upon prior work on embodied cognition in three primary ways. The first is that while most work on embodiment theory has centered on cognition in mature adult minds (that is, the representational results of *prior* bodily actions), here we take a developmental approach, focusing on the embodied practices that help initially inform those representations for a child. Given that how young children bring order into their world by categorizing objects is a central concern of research in cognitive and language development, our study aims to draw attention to the moment-by-moment interactional mechanisms through which such categorization develops. The two additional contributions of our study are described in detail in the subsections that follow.

1.1. A Multimodal, Interactional Perspective

It is well established in the developmental literature that initial learning takes place within here-and-now activities. Indeed, developmental theories suggest that young children's situated interactions with caregivers ground the here-and-now referents from which mentation about abstract concepts such as categorization may arise (Bruner 1975; Nelson 1996; Vygotsky 1978). Developmentalists have also shown that initial learning takes place within a shared referential context that includes repetitive multimodal routines which allow the child to utilize current knowledge to predict and extend meaning (Tomasello 2003). Yet despite this widespread agreement in the field as to the import of activities and routines for learning, comparatively little attention has been paid, in embodiment theory research, to *how such interactions actually unfold*. That is, few studies of embodied cognition focus on the real-time embodied practices that ground cognition in situated social interaction. Our study of toy play thus takes up Wilson (2002, p. 635) explicit challenge to embodied cognition researchers to provide more specific explanations of "cognitive activity that is embedded in a task-relevant situation" and cognitive activity in which "sensory and motor resources are brought to bear on mental tasks whose referents are distant in time and space and altogether imaginary". While prior work in this area has focused on the neural pathways and representations in the brain that underpin these seemingly separate aspects of cognition (Barsalou 1999; de Vega et al. 2008; Pecher and Zwaan 2005), here we seek to provide detailed descriptions of the perceptions and actions that give rise to new and increasingly abstract learnings which continuously and contiguously emerge.

Relatedly, theories of embodied cognition stress the importance of multimodal contributions to the generation of knowledge. However, as Vigliocco et al. (2014, p. 2) point out, language studies have focused "predominantly on speech and/or text, thus ignoring the wealth of additional information available in face-to-face communication". In the past, non-verbal aspects of communication were considered ancillary and studied independently; however, more recent work—particularly with visually available gestures—has shown that speech–gesture combinations are ubiquitous and tightly

connected from the earliest stages of development (Kendon 2004; Ozcaliskan and Goldin-Meadow 2005; Tomasello 2003). Moreover, Church and Goldin-Meadow (1986) provide evidence that co-speech gestures and speech combinations index changes in conceptual knowledge. As such, we argue that multimodal analysis of face-to-face interaction must factor into any comprehensive conceptualization of embodied cognition.

1.2. Beyond the Visual

In order to underscore the relevance of *embodied* practices, specifically, to the negotiation of categories, we analyze interactions between a blind child and her sighted mother. While visibly available co-speech gestures have been significantly targeted in prior research, other (i.e., *non*-visually available) multimodal and embodied resources have not yet received the same attention (but see Rickard 2013).

Studies of embodied cognition strongly implicate and favor the visual modality in exploring how perception and action guide crucial understandings such as those linked to knowledge about categorical organization (Barsalou 1999; Gibson 1979; Noe 2004; Semin and Smith 2008). Indeed, for young children exploring their world, vision is the primary motivator of learning, the integrator of the senses, and the most efficient access to pertinent environmental and social information (Haith and Benson 1998; Langton et al. 2000). For example, sighted children have immediate and temporally stable visual access to multimodal resources that are crucial to linking language to categorical affordances in the world—e.g., the 'gestalt', particularly the shape and motions of objects in their environment (Smith and Samuelson 1997), and the specific placement and positioning of objects (Clark 2003). Sighted children can also respond to gestural and facial cues indicating the location of near and distal objects, as well as parts of objects that may highlight critical features and provide a variety of exemplars (Goldin-Meadow 2003; Golinkoff et al. 1994; Kendon 2004; Vigliocco et al. 2014). In addition, as Borghi (2005, p. 12) points out, visual input "potentiates the affordances associated with the object".

Congenitally blind children, however, must make use of other—i.e., non-visual—modalities and environmental resources as they learn to locate, parse out, and categorize objects in the world (Perez-Pereira and Conti-Ramsden 1999). Accordingly, they provide a unique opportunity to expand our understanding of embodied cognition.

It bears mention that, while there is controversy in the literature with regard to how well blind children in general succeed in understanding categorical information, clearly some congenitally blind children without concomitant cognitive issues are successful (Andersen et al. 1993; Perez-Pereira and Conti-Ramsden 1999). Moreover, work by Rickard (2013) suggests that young children and their caregivers tend to make use of the modalities that are most efficient and effective in gaining access to pertinent information; indeed, caregivers for both blind and sighted children consistently meld language with gesture and other non-visual forms of indicative and demonstrative actions to establish joint attention and to highlight important information within their play activities. Thus, rather than focusing on blind children from a *deficit* perspective, we follow the example of scholars such as Avital and Streeck (2011) in analyzing the sense-making practices that blind children *do* make use of to successfully navigate the world. Doing so allows us to underscore how a wider multimodal repertoire of interactive and communicative resources can foster embodied categorical understandings (Akhtar and Gernsbacher 2008).

In what follows, we provide some background on the data, participants, and methods for our study. The majority of the paper is then dedicated to the detailed analysis of excerpts from the data. We conclude by discussing the theoretical implications of our analysis, as well as outlining some potential avenues for future research.

2. Materials and Methods

The dyad selected for this analysis was part of a large longitudinal study of language development in blind and sighted children aged 18 to 42 months (Rickard 2013). The original study was approved

by the Institutional Review Board (IRB) of the University of Colorado, and parental permissions were granted for ongoing educational and research use of the video recordings. Recent confirmation of permission for the current analysis was obtained from both the parent and the child, who is now a junior in high school.

The child on which we focus here, Maddie, was selected based on her meeting of typical developmental and language milestones during the time course of the original study. She had no health or cognitive issues concurrent with her blindness that might adversely affect her language development. In addition, her mother was highly attuned to her needs and was regularly engaged in extended family and community support systems. While this nonrandom participant selection of course precludes generalizations, in this exploratory study we have chosen a different goal, one of rich description of a naturalistic situated interaction that examines all collaborative activity in detail.

At 26 months of age, Maddie had only minimal light perception resulting from Leber's congenital amaurosis (LCA), a rare genetic disorder. Precision in specifying visual acuity in young children is problematic because (1) there is no definitive neurological measure, and (2) due to typical linguistic and cognitive developmental status, self-report is either impossible or unreliable. In addition, particularly in early development, observable acuity levels of blind children can fluctuate (Perez-Pereira and Conti-Ramsden 1999). In the original study, observational assessments (parental and teacher observations, and play-based assessments) were regularly employed to describe acuity levels germane to the study of language development. These included: (1) spontaneously lifting and/or orienting the head toward silent objects near or far; (2) spontaneously moving or reaching toward silent objects near or far; (3) spontaneously orienting to gestures such as pointing toward objects near or far; and (4) spontaneously bringing objects close to the eyes for examination. Maddie exhibited none of these behaviors. Her light perception was assistive only minimally in orienting for mobility through light and dark shadowing such as along hallways and doorways. Accordingly, Maddie did not orient to objects through sight in a manner that would be assistive in the canonical looking/gesturing/labeling scenario commonly observed in interactions with sighted children and employed in experimental designs to link words and referents (Akhtar and Tomasello 2000).

Maddie and her mother were videotaped in their home in order to provide a naturalistic and familiar setting in which to examine their typical modes of interacting during toy play. On a daily basis they spent extensive time in dyadic play activities of the sort reflected in the data we present here. Particular toy sets were selected based on prior observation of their play and discussion with Maddie's parent regarding the types of toys she enjoyed.

To examine these playtime activities, we employ conversation analysis (CA) in an attempt to uncover the multimodal methods and practices—both verbal and non-verbal—that participants use to make sense of, and in, naturally-occurring social interaction (for an overview, see Clift 2016). We argue that this methodology offers a perspective on linguistic and cognitive development that is both more naturalistic as well as more comprehensive when compared to laboratory experimental research designs that focus primarily on the input of parental language forms and the supportive use of deictic gestures to elicit convergent eye gaze.

Moreover, we have elected to take a *single-case analysis approach*, in which findings from previous collections-based analyses of interactional phenomena are brought to bear on the examination of a single interaction (cf. Clift and Raymond 2018). Such an analysis, in Schegloff (1987, p. 101) words, has as its goal to "assess the capacity of [CA as an] analytic enterprise" by applying past results to new data. This is well-fitted to the aims of the present study, in which we seek to explore what the "analytic enterprise" of CA can offer the study of language and embodied cognition from an interactional perspective.

Of particular relevance to this goal is CA's conceptualization of meaning-making as an *emergent* process that unfolds *collaboratively*, in and through participants' moves in interaction. Here we apply this way of thinking to the activity of toy play between Maddie and her mother. As we will show, multimodal practices entail analysis of the interweaving of a wider range of interactional

forms that encompass reciprocal language use, the participants' interacting bodies, the social and cultural environment in which actions unfold, and the particular tools (toys in our data) that provide affordances for learning (Goodwin 2000).

3. Results

In what follows, we provide an analysis of a segment involving a particular toy—a cow plush—and aim to address two primary questions with regard to Maddie and her mother's collaboratively constructed processes for categorization.

1.　During the play activity, how does Mother scaffold multimodal routines for the identification of *criterial information* about a category?
2.　How is knowledge of *varied exemplars*, not directly accessible within the current activity, then made available to the child?

We will address each of these questions in turn.

3.1. Orienting to What It Criterial

While immediate and temporally stable access to the typical shape and features of an object may organize sighted children's categorization activities, in the case of the blind child in our data, categorization work starts from a manual search of the parts of an object, such as a beak for a bird or a snout for a pig. In other words, while sighted children may make use of perceptual 'gestalts' as resources for categorizing, this blind child makes use of key parts for categorizing. This fact leads us to our first research question: Once Maddie has manual access to an object, how does she learn which properties of that object are criterial for defining it as a member of a particular category?

Consider the segment in (1) below, in which Maddie and her mother are seated on the floor, facing each other, with a pile of stuffed animals between them (mostly between Maddie's legs). The sequence we will be examining starts after Maddie has brought her exploration of a toy frog to completion. She initiates the new sequence with a request for another toy: *(I) want the cow* (line 1).[1]

```
(1) Request for the Cow

01    CHI:    (I) want ↑the cow,
02            (0.3) + *(0.4)
      mom              *reaches hand to pile of toys and begins pushing
              toys aside ->
03    MOM:   well we have to ↑find *him,#
      mom    ------------------> *touches cow
      fig                              #fig A
```


Figure A

In the silence at line 2, Mother moves her hand to the pile of toys and begins a manual and visual search through them. At line 3, rather than simply verbally complying with the request (such as saying "okay") and finding the toy *for* Maddie, Mother lets Maddie know that finding the cow will be

---

[1]　Transcripts follow the multimodal conventions outlined in Mondada (2018).

a collaborative activity (through the use of the pronoun *we*) and that it will require some effort on their parts (*have to find him*). On the last word of that utterance, Mother's hand finds the toy cow (Figure A).

Rather than announcing that she has found the cow, however, Mother does some manual actions with it to draw Maddie's awareness to that particular toy, and Maddie herself begins a manual search for it. As seen below in (2), Maddie first grasps a different toy (Figure B), releases it, and then continues the search. Finally, Mother picks the cow up (Figure C) and holds it a bit higher, but still does not announce its presence verbally:

```
(2) Search for the Cow

04                *(0.7) + *(1.0) + *(1.3) + *(1.0)
      mom        *jiggles cow
      chi              *reaches into toy pile
      chi                    *grasps toy and feels it#
      fig                                        #fig B
      chi                                *releases toy and moves hand to a
             different part of the pile, searches -->
      mom                              *picks up cow #
      fig                                        #fig C
```


Figure B

Figure C

While Maddie continues to search for the cow, Mother verbally provides a clue in the form of a criterial property of cows (line 5 below, *udders*). At the beginning of that utterance she performs an action that facilitates Maddie's perception of the critical feature. She rotates the toy cow so that its udders are facing up (Figure D). Both the verbal expression and the manual action of placing the toy for maximal access serve to highlight the salience of this particular feature of cows to guide Maddie's search. After an extensive manual search, Maddie brings her hand to her side and momentarily stops the search (line 6, Figure E):

```
(3) Udders

05    MOM:    *↓f<u>i</u>nd the one that has# the ↑<u>u</u>dders,
      mom     *turns cow with udders up and then puts it back down
      chi     ----------------------------->
      fig                         #fig D
06            (3.0) + *(1.0)
      chi     -->     *hand comes down by her side#
      fig                              #fig E
```


Figure D

Figure E

After Maddie's hand comes to a resting position, Mother initiates an insert sequence to offer help (line 7) (Schegloff 2007). Even before Maddie replies, Mother has lifted the toy cow and rested

it on top of the pile of toys. Once Maddie begins to accept the offer of help (line 10, Figure F), her mother performs an indicative action, tapping the cow on Maddie's hand. She then says *what's this guy* (Figure G):

```
(4) Mommy helps

07    MOM:      do you want [mommy to help you?
08    CHI:                  [m:
09                *(0.7)
      mom         *picks up cow, brings it towards child, holds it
10    CHI:      mommy to help *you:#,=
      mom                      *taps cow on child's hand
      fig                         #fig F
11    MOM:      =what's this guy,#
      mom        ------------->
      fig                      #fig G
```


Figure F

Figure G

Mother has produced a question (*what's this guy*), which makes relevant an answer. Immediately upon completing the question, at the beginning of line 12, Mother lifts the toy so that the cow's tummy and its udders are facing Maddie again making the criterial feature perceptually available. As this movement of the toy starts, Maddie reaches for and grasps a leg of the cow. As these actions on the part of Mother and Maddie come to an apex (Figure H), Maddie answers her mother's question—*a cow*.

```
(5) A cow

12                    *(1.1)
      mom         *lifts cow so that tummy is facing child
      chi         *feels leg of cow
13    CHI:      <a# *co:w,>=
      chi          *begins to grab hold of the cow
      fig       #fig H
```


Figure H

Note that while she is answering her mother's question, Maddie begins to take hold of the cow, but Mother does not release her own hold on the toy. This serves as preliminary evidence for Maddie that there is likely some problem with her answer.

Maddie's response (*a cow*) is based crucially on two factors: the context of her own original request for the cow, and her manual contact with the cow's leg. Although Maddie's contextual reasoning turns out to be correct, Mother nonetheless problematizes that reasoning (*are you sure?*) by suggesting that

Maddie has arrived at her conclusion on inadequate grounds—she should perform an action that indicates her understanding of the criterial property of udders (*feel his tummy*, Figure I):

```
(6) Feel his tummy

14    MOM:     =are you sure? *feel his ↑tummy,#
      mom                     *lets go of cow
      fig                                 #fig I
```


Figure I

As Mother's directive to *feel his tummy* comes to completion, during line 15 Maddie brings the cow closer to her (Figure J) and moves her hand up the cow's body, eventually coming to the cow's nose. Mother corrects that direction of exploration at line 16 (*that's his nose*, Figure K), and before that utterance comes to completion, Maddie begins to move her hand back down the cow's body:

```
(7) That's his nose

15                *(2.7)
      chi         *brings cow to touch her body,# moves hand up the cow's
                  body
      fig                                   #fig J
16    MOM:     ↑that's his no*:s:e,#
      chi                    *begins moving hand down
      fig                 #fig K
```


Figure J

Figure K

Next, during the silence that unfolds at lines 17 and 18, Maddie brings her hand down to the middle of the cow (Figure L), but because of the angle of her hand, she is not actually feeling/perceiving the 'tummy' of the cow. At line 19, Mother once again both verbally and manually guides the child to find the criterial feature: She asks *what's on his tummy*, and while she is saying that, she moves Maddie's hand and places it with the palm on the cow's 'tummy' (Figure M). During the silence at line 20, Mother moves Maddie's hand along the 'tummy' to make haptically available the presence of the udders. And finally, at line 21, Maddie answers the question with the correct term, *udders*.

```
(8) Maddie feels the udders

17                  (3.5)
18    CHI:     *at the tummy,=#
      fig                      #fig L
      mom      *reaches towards cow
      chi              *pulls cow closer to her
19    MOM:     =(show me something. wha-) *feel, what's on his tummy.#
      mom                               *takes child's right hand, places it
               on cow's tummy
      fig                                                      #fig M
20               *(0.7)
      mom          *moves child's hand over the cow's tummy
21    CHI:     udders,
```

Figure L　　　　　　　　　　　　　　　　　　**Figure M**

Through this sequentially complex interaction, Mother guides Maddie to an accurate perception of a criterial property of cows—udders. She first names udders in a way that implicates that udders are criterial for cows, or at least distinguishes the cow toy from the other toys in the pile (*find the one that has udders*, example 3). When Maddie struggles to find the cow, Mother first makes the toy available non-verbally; when that is not sufficient, she taps Maddie's hand with the cow and verbally instructs her to find the category of that toy. Crucially, even though at this point Maddie correctly guesses the category of the toy, Mother interactionally treats her process of reasoning as inadequate, because it has not located the udders as the source of the categorization, and she sets out to guide Maddie in finding the udders, both verbally and by the embodied action of moving Maddie's hand over the cow's "tummy". Through her verbal and multimodal actions across this sequence, then, Mother underscores the need to base categorization decisions on *criterial* features—not on manual inspection of a body part that inadequately distinguishes one type of animal from the others present (e.g., legs), and not merely on contextual reasoning—and thereby scaffolds Maddie's ability to appropriately categorize objects in the world, in particular with regard to the embodied practices necessary to make such categorical determinations in the immediacy of real-time social interaction.

3.2. Orienting to Varied Exemplars

Categorization requires that a child be able to apply the category label to varied exemplars, including those that are not present in the immediate surroundings. This leads to our second research question: How do verbal and multimodal practices in interaction make available to the child knowledge of varied exemplars that are not directly accessible or present in the activity at hand, and which could contrast with the present exemplar in various ways?

Continuing the examination of our target sequence, after Maddie successfully establishes the categorical label for the toy cow via a scaffolded manual search focused on the criterial property of udders, Mother continues discussion of the cow, now guiding Maddie to categorize and label the cow utilizing the acoustic criterial property for cows—the traditional "moo" sound. In the extract below, which occurs just a few seconds after case (8) above, Maddie correctly labels the cow once more (line 8), and her mother employs a question form that seeks an assessment of the cow (line 9, *do you*

like that cow?). In response, Maddie produces not an assessment, but rather the label *Carolyn's cow* (line 11). In this sequence—which is verbally and non-verbally guided, with the addition of the sense of sound—Maddie attends to the use of a deictic determiner and the lack of specificity in referring to the stuffed cow as merely "that cow".

```
(9) Carolyn's cow

01    MOM:    Should we squeeze this- (0.2) ↑animal and see what he is?
02               *(2.0)
      mom         *moves cow toward child's hand,# touches child's right hand
              with cow
      fig                                              #fig N
03    CHI:    *Squeeze *this animal,#
      chi     *touches cow's head
      chi            *moves hand to middle of the cow and squeezes
      fig                     #fig O
04    COW:    *[moo:::# (0.8) mo[o::: (0.8) moo::: (0.8)
      chi     *moves hand up and down body of cow
      fig           #fig P
05    MOM:     [(.hhh)
06    CHI:                      [He says moo: moo: moo:=
07    MOM:    =He *doe:s? (0.6) ↑Who says moo moo moo?
      mom         *releases grip on the cow so that child may hold it
08    CHI:    ↑A cow::,
09    MOM:    Oh::: (1.6) ↑Do you like that cow:?#
10               *(2.0)
      chi         *tilts head downwards toward cow in right hand
      fig                                         #fig Q
11    CHI:    Caro↑lyn's cow,=
12    MOM:    =°That's [Carolyn's cow:,°          ]
13    CAR:             [Yea:h, that's ↑my cow,]=
14    MOM:    =She's ↑sharing it with you,
15    MOM:        (1.2)
16    MOM:    She's letting you pla:y with it for a little while.
```


Figure N

Figure O

In continuing the conversation around the cow, Mother guides Maddie to utilize sound in line 1 (*should we squeeze this—(0.2) ↑animal and see what he is?*) as a feature for not only identifying but also learning more about an object. Notably, Mother's question is not smoothly through-produced, as her voice cuts off and a short pause emerges before the production of the word *animal*, produced with emphasis on the initial vowel sound and heightened pitch. The production of this turn suggests that Mother opted for the superordinate term *animal* rather than the specific categorical label *cow*, as use of the latter term would have obviated the need for Maddie to make use of auditory input to identify the animal.

In this moment, Maddie is not holding the cow in her right hand, so her Mother introduces the animal as though for the first time, initiating yet another scaffolding sequence aimed at categorizing the stuffed cow. In order to encourage Maddie to use her sense of touch to find the animal's sound box, Mother physically positions the cow such that Maddie may immediately grasp it in her right hand (Figure N), only releasing the cow entirely in line 7 after Maddie manually examines the cow's body. As Maddie physically searches for the middle of the cow (the location of the sound box) by moving her hand along the plush, she verbally repeats *squeeze this animal* (line 3, Figure O).

Figure P

Figure Q

While moving her right hand up and down the body of the cow during the cow's sound production, Maddie overlaps the cow's sounds in line 4 with her own statement *he says moo moo moo*, indicating that she recognizes the sound (Figure P). Desiring a specific label for the producer of the sound, her mother uses the question form *who says moo moo moo?* (line 7), borrowing the sound description from Maddie's production in the turn prior. Upon identifying the individual cow present in this particular interaction with an indefinite determiner (*a cow*, line 8), Maddie reveals that she understands that there are many members of the category 'cows', and that the stuffed cow to which she has immediate physical and auditory access is but one exemplar from that category. The use of the indefinite determiner here is the first indication of Maddie's awareness that the categorical label 'cow' extends beyond the single stuffed cow present in the interaction.

After being explicitly prompted by her mother to provide an assessment of the cow in line 9 (*do you like that cow?*), Maddie asserts the possessive *Carolyn's cow* (line 11). Mother's question underscores that the cow in question is a single exemplar in the category through her use of the specifying spatial deictic *that*, referencing the cow in Maddie's hands. Such a deictic term, while grammatical, is nonetheless inherently multimodal in its dependence on the physical and spatial organization of the participants in conjunction with their built environment. The deictic term focuses Maddie's attention on the object in her immediate perceptual space and also makes relevant an assessment response relating only to the cow present in the interaction. Maddie's response in line 11 suggests that she closely attended to the turn-final phrase in mother's production (*that cow*) such that she felt inclined to provide a more specific label for the cow.[2] Maddie's response *Carolyn's cow*, which occurs after an attributable silence of two seconds (see head tilted downwards in Figure Q), may constitute a correction of Mother's prior phrase, problematizing her labelling and use of the distal form, or alternatively it could be working to add specificity to the label *that cow* by the use of the possessive form. The possessive label not only differentiates this particular toy cow from others that may exist, but it also demonstrates a further specification from the use of the general indefinite determiner earlier in line 8. Whether Maddie's turn *Carolyn's cow* serves as a correction of her mother's *that cow* or not, it highlights Maddie's understanding that categories have multiple members in them, and that those members may differ from each other in various ways: The cow in this interaction differs from others in belonging to Carolyn. Over the course of the interaction, Maddie demonstrates the layered nature of her category knowledge of the cow in her ability to move interactionally between *a cow* in line 8 (referring any member of the category using the indefinite determiner) to *that cow* (deictically more specific) to *Carolyn's cow* (possibly the 'maximally' specific form in this interaction), with each shift being facilitated by her mother—via an assessment prompt in line 9, and then later through an explicit repetition of Maddie's turn in line 12.

2　　In addition, Mother's use of the proximal deictic *this* form very recently in the interaction (line 1) may cause some confusion for Maddie, who herself repeated the proximal form in line 3.

4. Discussion and Conclusions

Much research on blind children has tended to follow a deficit model; that is, the focus has been on what these children *lack* compared to their sighted peers, as opposed to an attention to how they make use of the resources to which they do have access in order to achieve interactional and communicative goals. Relatedly, an attention to the relevance of visual (as well as aural) input has predominated in research on language and categorization development. Barsalou's (1999) conception of "perceptual symbol systems", for instance, is entirely visually based and thus cannot account for the development of those children who, although congenitally blind, nonetheless reach the same developmental milestones as their sighted peers. This reality suggests that a more comprehensive model of the development and acquisition of embodied cognition must necessarily include haptic and other sensorimotor pathways as partially contributory or singular alternatives to the visual modality (on which, see, e.g., Struiksma et al. 2009).

In this exploratory investigation, we have offered a case study detailing how multimodal, co-operative actions can provide non-visual alternative access for a blind child—that is, for a child who cannot take in the visual 'gestalt' of an object for the purposes of categorization in language development. Through our examination of Maddie engaging in routine toy play with her mother, we argued that the child's categorical understanding of *criterial information* and *varied exemplars* was constituted through reciprocal elaboration of varied embodied and multimodal resources. Just as visually available gestures and actions ubiquitously co-occur with speech directed toward sighted individuals, physical movements, indicative and demonstrative actions synchronously supported and enhanced Maddie's understandings about how to locate, categorize, and interact with objects, and thereby effectively and productively parse out her world. We hypothesize that this sort of interactional scaffolding significantly contributes to Maddie consistently meeting the same developmental milestones as her sighted peers. Indeed, longitudinal data of Maddie shows her embodied understandings of objects soon begin to apply to "referents that are distant in time and space and altogether imaginary" (Wilson 2002, p. 635) at the same rate as sighted children: While she was bound closely in both proximity and thought to her local context at 26 months, by 34 months she was fluidly moving about her living room mentally, taking her categorized groups of animals on imaginary trips to the moon.

The detailed interactional analysis we offer here provides documentation of the organization of multimodal sensorimotor activity that may provide grounded routes to the building of an early embodied mind. That is, we have targeted non-visual modalities through which children might come to have internal representations from which further action and thought might flow—internal representations which must emerge, we maintain, through physical bodies operating collaboratively in the immediacy of social interaction.

If how embodied mentation is instantiated in the neural system and represented in the brain is part of the goal of work on language and embodied cognition, we maintain that rich description of when and with what multimodal and social resources shifts in mentation might take place should inform such research efforts; and of course it is in naturally occurring, moment-by-moment social interaction that such communicative resources are to be found and examined, as participants jointly make sense with one another. In our data, Maddie and her mother are *doing* things together in the conduct of their play to build up categorical understandings. Maddie's mother interactionally directs her attention, showing her *what* to *see* (making available an appropriate perception), and *how* to *see* (how to act on a toy to obtain pertinent category information) while talking about the details and import of what is being *seen*. It is essential to note here that *seeing* in Maddie's case is not just translated into haptic access, in the same way that seeing for sighted individuals is not simply a retinal image. *Seeing* what is perceived is inherently embodied through the coordination of talk and non-verbal behaviors, combined with the larger contextual configuration of the surrounding. We therefore conclude that any comprehensive theory of language and embodied cognition must necessarily incorporate analysis of the dynamic properties of conduct in real-time, naturalistic interaction. This is because it is in and

through participants' collaboratively embodied behavior *in interaction* that we are able to bear witness to the "synchronization, coordination, and co-regulation of behaviors" that occurs "*across* brains and bodies" (Semin and Cacioppo 2008, p. 121)—at the very moment of their (re)production, in the service of building co-operative actions.

References

Akhtar, Nameera, and Morton Ann Gernsbacher. 2008. On privileging the role of gaze in infant social cognition. *Child Development Perspectives* 2: 59–65. [CrossRef] [PubMed]

Akhtar, Nameera, and Michael Tomasello. 2000. The social nature of words and word learning. In *Becoming a Word Learner: A Debate on Lexical Acquisition*. Edited by Roberta Micknick Golinkoff and Kathryn Hirsh-Pasek. New York: Oxford University Press, pp. 115–35.

Andersen, Elaine S., Anne Dunlea, and Linda Kekelis. 1993. The impact of input: Language acquisition in the visually impaired. *First Language* 13: 23–49. [CrossRef]

Avital, Sharon, and Jurgen Streeck. 2011. Terra Incognita: Social interaction among blind children. In *Embodied Interaction: Language and Body in the Material World*. Edited by Jürgen Streeck, Charles Goodwin and Curtis LeBaron. Cambridge: Cambridge University Press, pp. 169–81.

Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660. [CrossRef] [PubMed]

Borghi, Anna M. 2005. Object concepts and action. In *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Edited by Diane Pecher and Rolf A. Zwaan. New York: Cambridge University Press, pp. 8–34.

Bruner, Jerome S. 1975. From communication to language: A psychological perspective. *Cognition* 3: 255–87. [CrossRef]

Church, R. Breckinridge, and Susan Goldin-Meadow. 1986. The mismatch between gesture and speech as an index of transitional knowledge. *Cognition* 23: 43–71. [CrossRef]

Clark, Herbert H. 2003. Pointing and placing. In *Pointing: Where Language, Culture and Cognition Meet*. Edited by Sotaro Kita. Hillsdale: Erlbaum, pp. 243–68.

Clift, Rebecca. 2016. *Conversation Analysis*. Cambridge: Cambridge University Press.

Clift, Rebecca, and Chase Wesley Raymond. 2018. Actions in practice: On details in collections. *Discourse Studies* 20: 90–119. [CrossRef]

de Vega, Manuel, Arthur M. Glenberg, and Arthur C. Graesser. 2008. *Symbols and Embodiment: Debates on Meaning and Cognition*. New York: Oxford University Press.

Edelman, Gerald M. 1992. *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: Basic Books.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Goldin-Meadow, Susan. 2003. *Hearing Gesture: How Our Hands Help Us Think*. Cambridge: Belknap Press of Harvard University Press.

Golinkoff, Roberta Michnick, Carolyn B. Mervis, and Kathryn Hirsh-Pasek. 1994. Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language* 21: 125–55. [CrossRef] [PubMed]

Goodwin, Charles. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32: 1489–522. [CrossRef]

Goodwin, Charles. 2018. *Co-Operative Action*. New York: Cambridge University Press.

Goodwin, Marjorie Harness, and Asta Cekaite. 2018. *Embodied Family Choreography: Practices of Control, Care, and Mundane Creativity*. New York: Routledge.

Haith, Marshall M., and Janette B. Benson. 1998. Infant Cognition. In *Handbook of Child Psychology*. Edited by William Damon, Deanna Khun and Robert Siegler. New York: John Wiley & Sons, pp. 199–254.

Kendon, Adam. 2004. *Gesture: Visible Action Utterance*. Cambridge: Cambridge University Press.

Langton, Stephen R. H., Roger J. Watt, and Vicki Bruce. 2000. Do the eyes have it: Cues to the direction of social attention. *Trends in Cognitive Science* 4: 50–59. [CrossRef]

Mondada, Lorenza. 2018. Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language & Social Interaction* 51: 85–106.

Nelson, Katherine. 1996. *Language in Cognitive Development: The Emergence of the Mediated Mind*. Cambridge: Cambridge University Press.

Noe, Alva. 2004. *Action in Perception*. Cambridge: MIT Press.

Ozcaliskan, Şeyda, and Susan Goldin-Meadow. 2005. Do parents lead their children by the hand? *Journal of Child Language* 32: 481–505. [CrossRef] [PubMed]

Pecher, Diane, and Rolf A. Zwaan, eds. 2005. Introduction to Grounding Cognition: The role of perception and action in memory, language, and thinking. In *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. New York: Cambridge University Press, pp. 1–7.

Perez-Pereira, Miguel, and Gina Conti-Ramsden. 1999. *Language Development and Social Interaction in Blind Children*. East Sussex: Psychology Press.

Rickard, Carolyn. 2013. Multimodal Cues in the Socialization of Joint Attention in Young Children with Varying Degrees of Vision: Getting the POINT Even When You Can't See It. Ph.D. dissertation, University of Colorado, Boulder, CO, USA.

Schegloff, Emanuel A. 1987. Analyzing single episodes of interaction: An exercise in conversation analysis. *Social Psychology Quarterly* 50: 101–14. [CrossRef]

Schegloff, Emanuel A. 2007. *Sequence Organization: A Primer in Conversation Analysis (Vol. 1)*. Cambridge: Cambridge University Press.

Semin, Gün R., and John T. Cacioppo. 2008. Grounding social cognition: Synchronization, coordination, and co-regulation. In *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches*. Edited by Gün R. Semin and Eliot R. Smith. New York: Cambridge University Press, pp. 119–47.

Semin, Gün R., and Eliot R. Smith. 2008. *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches*. New York: Cambridge University Press.

Smith, Linda B., and Larissa K. Samuelson. 1997. Perceiving and Remembering: Category Stability, Variability, and Development. In *Knowledge, Concepts, and Categories*. Edited by Koen Lamberts and David R. Shanks. Hove: Psychology Press, pp. 161–95.

Streeck, Jurgen, Charles Goodwin, and Curtis LeBaron. 2011. *Embodied Interaction: Language and Body in the Material World*. Cambridge: Cambridge University Press.

Struiksma, Marijn E., Matthijs L. Noordzij, and Albert Postma. 2009. What is the link between language and spatial images? Behavioral and neural findings in blind and sighted individuals. *Acta Psychologica* 132: 145–56. [CrossRef] [PubMed]

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Vigliocco, Gabriella, Pamela Perniss, and David Vinson. 2014. Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B* 369: 20130292. [CrossRef] [PubMed]

Vygotsky, Lev Semenovich. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.

Wilson, Margaret. 2002. Six views of embodied cognition. *Psychonomic Bulletin and Review* 9: 625–36. [CrossRef] [PubMed]

Zukow-Goldring, Patricia, and Kelly R. Ferko. 1994. An ecological approach to the emergence of the lexicon: Socializing attention. In *Sociocultural Approaches to Language and Literacy: An Interactionist Perspective*. Edited by Vera John-Steiner, Carolyn P. Panofsky and Larry W. Smith. New York: Cambridge University Press, pp. 170–90.