THE UNIVERSITY OF CHICAGO

OBVIOUSLY I DON'T KNOW BUT WHATEVER: EMBLEMATIC AND PRAGMATIC

USES OF SHRUG GESTURES IN EARLY CHILDHOOD AND ADOLESCENCE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPARATIVE HUMAN DEVELOPMENT

BY

NATALIE DOWLING

CHICAGO, ILLINOIS

AUGUST 2022

For Audrey, Gilroy, and Lucas, not necessarily in that order.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

more than a decade ago, she has not only kept the LDP afloat (sometimes single-handedly, it seems) but more than a few times kept me personally afloat, and I am so grateful.

I am lucky to have spent these last years as a member of the Susan Goldin-Meadow Lab, the Chatter Lab, and the Department of Comparative Human Development. These three communities – which are far from mutually exclusive – create welcoming spaces for intellectual exploration and academic rigor without sacrificing decency, empathy, passion, or humor. Our lab managers Kennedy Casey and Jess Breeze have been essential supports to me in the final year of writing. CHD's student affairs administrator Janice Pavel has provided unrelenting patience and organization.

I am grateful for so many friends I have found here, but especially Hilary Tackie, Zoe Berman, Maddy Oswald, and Jimmy Waller. Hilary and Zoe, you represent all the best parts of CHD and UChicago. I am so thankful for our friendship and for your unflagging efforts to make our communities better for those who follow us. Maddy, thank you for your unparalleled generosity and joy. Jimmy, you are incredible and indescribable, and my love for you is immeasurable.

I am grateful to the members of my dissertation committee, Susan Goldin-Meadow, Marisa Casillas, John Lucy, and Federico Rossano, for their support both academically and personally.

Federico, thank you for lending your expertise to this project and for giving me the confidence to stick my toes into the world of conversation analysis.

John, thank you for being a guiding light for not just me but so many of us in CHD. I have never left a meeting with you without both a renewed inspiration for my work and a reminder that being challenged up to (or well past) my limits is in fact a good thing.

ABSTRACT

In everyday interaction interlocutors use pragmatic co-speech gestures to cooperatively construct conversation. Shrugs, one of the most common pragmatic gestures, communicate a remarkable array of seemingly unrelated or even contradictory meanings – agreement and disagreement, ignorance and obviousness, interest and disinterest, among others. Although shrugs are often among the earliest gestures acquired by American English-learning infants, children do not immediately use shrugs with the same variability of form and function as adults. Because shrugs are used both emblematically and interactively, they provide unique insight into pragmatic development. How do shrugs simultaneously function as conventionalized symbols and contextualized indices? How do children develop this pragmatic flexibility? Although there is a wealth of literature concerning gesture's role in communicative development, existing research has primarily focused on topic gestures and left pragmatic gestures relatively understudied. Within the limited existing literature, pragmatic gestures are rarely given the same nuanced consideration afforded to topic gestures. There is an implicit assumption that meaningful differences in the role gestures play in early communication are present across categories of gesture but not within the class of pragmatic gesture.

In this dissertation I use a corpus of spontaneous parent-child interaction and an annotation scheme grounded in principles of conversation analysis to explore how pragmatic gestures like shrugs operate in early communication. In three studies, I describe changes in form and function across early childhood and into adolescence. First, I show that the developmental trajectory of palm-up gestures in early childhood is functionally distinct from that of beat gestures. Second, I describe young children's use of shrug gestures with and without speech. From these analyses, I propose that children initially produce shrugs as ignorance emblems and

develop shrugs' characteristic many-to-many form-function mapping in later language development. The third and final study explores this transition from emblematic to pragmatic use in early adolescence. Together, these studies demonstrate the necessity of accounting for gesture variation within the broad functional category of "pragmatic gestures" in communicative development. By looking both within and beyond early childhood, this dissertation contributes to our understanding of pragmatic development as fundamentally multimodal.

1   Introduction and Review of the Literature

1.1   Introduction

Spontaneous dialogue drives early language learning. Infants take communicative turns with caregivers before they speak their first words and continue to use back-and-forth interaction to develop a vocabulary and to practice the structural rules of their language. By school entry at age 5, American English-speaking children know thousands of words, use relational language to organize concrete and abstract concepts, and begin to turn their oral language competence into literacy. Preschoolers have used conversational interaction to master the fundamentals of their language. They are not, however, masters of conversational interaction.

Children continue to develop their conversational competence through adolescence, gradually improving their understanding and execution of the social and linguistic rules that govern cooperative conversation. In everyday dialogue, speakers must determine the collaborative goals of an interaction and coordinate to achieve those goals, all while establishing and maintaining common ground, judging social distance, performing politeness, preserving face, and managing affect. As toddlers, children begin learning to take these conversational actions not only with their words but also with vocal quality, facial expression, eye gaze, posture, and gesture. As teenagers, they continue developing the pragmatic competences needed to co-pilot face-to-face interaction with increasingly sophisticated verbal and nonverbal resources.

In this dissertation I use a corpus of spontaneous caregiver-child interaction and an annotation scheme grounded in principles of conversation analysis to explore how children use pragmatic gestures like shrugs in naturalistic interaction. In three studies, I describe changes in form and function across early childhood and into adolescence. By looking both within and

beyond early childhood, this dissertation contributes to our understanding of pragmatic

development as fundamentally multimodal.

1.1.1 Outline of the Dissertation

In this chapter, I review current literature on the role of gesture in pragmatic development.

I draw from work in gesture studies, linguistics, cognitive psychology, and developmental

psychology to construct an interdisciplinary overview of how pragmatic gestures function

interactively in everyday conversation between adults and how these gestures feature in a

multimodal model of pragmatic development.

In Chapter 2, I discuss the theoretical motivation for my methodologies and review

relevant concepts from conversation analysis. I describe the longitudinal corpus of naturalistic

caregiver-child interaction, the cohort of families, and the annotation systems used in the studies

that follow.

Chapters 3 through 5 present three studies analyzing children's spontaneous production of

pragmatic gestures in the corpus. In the first study, I examine children's use of two pragmatic

gestures commonly produced in early childhood, beats and palm-up gestures. I ask whether

dimensions of form and gesture-speech relation are associated with differences in onset,

frequency of production, pragmatic function, and changes across early language development.

In Chapter 4, I focus on children's shrug gestures before age 5. I compare the form and

meaning of children's shrugs produced with and without speech to identify a candidate for the

"kernel" shrug emblem, mapping a predictable meaning onto a recognizable form before

growing into the highly flexible pragmatic gesture used by adults. I directly build on these results

in Chapter 5, where I discuss how the same cohort of children use shrug gestures in early

adolescence. In addition to describing adolescents' use of kernel and non-kernel shrugs, I explore emerging patterns of use based on compound gesture forms and preference organization.

The final chapter presents an overarching discussion of the results from all three studies. Based on the findings from Chapter 3, 4, and 5, I propose a simple framework for interpreting the many meanings of shrugs in terms of two axes of stance-taking. Finally, I consider limitations of the work and future directions for related research.

1.2 Literature Review

1.2.1 Pragmatic Gestures

Many types of communicative co-speech gestures have easily identifiable forms and predictable functional relationships to speech. Deictic gestures physically and metaphorically point to referents. Iconic gestures visually represent aspects of a referent. Emblematic gestures act as conventionalized signs within a community, like the "thumbs up" emblem meaning "good" to American English speakers.

Other gestures are not so straightforward, but no less important in face-to-face conversation. Many co-speech gestures primarily serve pragmatic functions; that is, they help facilitate communication and mostly benefit the listener rather than the speaker. The forms of these gestures typically do not index or iconically represent a referent in the discourse, nor do they confer a symbolic lexical meaning. Instead, pragmatic gestures refer to the discourse itself. While deictics, iconics, and other topic gestures serve cognitive benefits for the speaker (Goldin-Meadow, 2003; Hadar & Butterworth, 1997; Kita, 2009; Krauss et al., 2000), pragmatic gestures create an extralinguistic context for listeners' meaning-making. All gestures can implicitly or explicitly communicate, but speakers produce pragmatic gestures primarily or exclusively for their listener.

For the purposes of this dissertation, I divide gesture into two functional categories: topic and pragmatic. In essence, this is the distinction Bavelas, Chovil, Lawrie, and Wade (1992) make between *topic* gestures and *interactive* gestures. I use "pragmatic" rather than "interactive" to respect the authors' criterion that interactive gestures take an addressee-indicating form. I open pragmatic gestures as a category to include gestures that perform interactive functions without indicating an interlocutor, like beats, shoulder shrugs, and lateral palm-up gestures. I maintain their term "topic" to describe gestures that add semantic meaning to speech.

Topic gestures are those that indicate or represent a referent in the topic of talk. These include *deictic* gestures like points, which direct attention to a concrete entity in the physical environment or an abstract or non-present entity that has been figuratively "placed" in neutral space. Topic gestures also include *representational* gestures, which depict a literal or metaphorical aspect of the topic of talk, for example wiggling two fingers to iconically mimic running legs or tossing a hand over the shoulder to metaphorically represent the past.

A final category of topic gestures that will be of particular significance in the studies that follow are *emblematic* gestures. Ekman and Friesen (1969) originally defined emblems as "nonverbal acts which have a direct verbal translation…usually consisting of a word or two…well known by all members of a group, class or culture." Like word units in speech, the connection between form and meaning of a given emblem is symbolic and conventionalized within a linguistic or cultural community. Emblematic gestures may less often be referred to as *symbolic*, *quotable*, *autonomous*, *conventional*, or *semiotic* gestures (see Teßendorf, 2013 for a review).

Unlike topic gestures, *pragmatic* gestures neither indicate nor visually represent a topic of talk, nor do they have a conventionally agreed upon symbolic meaning. There is little consensus

about what "counts" as a pragmatic gesture, but in broad terms pragmatic gestures serve a primarily or exclusively discourse-pragmatic function (Ferré, 2012; Kendon, 2004, 2017; McNeill, 1992; Prieto et al., 2018; Shattuck-Hufnagel et al., 2016). In this dissertation a gesture is considered to be pragmatic if it lacks an identifiable referent in the topic of talk. Deictics *necessarily* have a topical referent to index. Representational gestures *necessarily* create a visualization of a referent, whether tangible (iconic gestures) or conceptual (metaphoric gestures). Emblematic gestures symbolize a "direct translation" that has been conventionalized within a community. These gestures reinforce or replace the meaning of a lexical unit in speech.

Because pragmatic gesture is often loosely or inconsistently defined, I work from criteria of indexicality and reference with an eye toward what will be most relevant for developmental pragmatics, following Bates (1976). In its broadest sense, pragmatics is the study of language use in context. Morris (1946) characterizes pragmatics as the relations between signs and their human users. This is a solid starting point, but it overlooks important philosophical distinctions about how different types of signs relate to their contexts of use. A more refined view of pragmatics builds on Peirce's (1991) organization of semiotics, where pragmatics is the study of indices. Unlike icons or symbols, linguistic indices may be interpreted only when contextualized. Indices lack inherent meanings and instead communicate meaning through rules relating the indexical sign with its representation in context.

With this perspective, the study of pragmatic development "occupies the interface between linguistic, cognitive, and social development" (Bates, 1976, p. 3). Research in this domain describes how children acquire and employ rules from all three of these areas of development to make meaning from context. Framed another way, pragmatic development is not concerned with children's acquisition or use of icons and symbols, where a sign's meaning may

be abstracted from its context. Understanding pragmatics as the study of linguistic indices motivates a break from how gesture studies treat nonverbal emblems. Emblems are often considered to fall under an umbrella category of pragmatic gestures because they are explicitly communicative; a conventionalized symbol serves the listener's interpretation of meaning more than the speaker's cognitive processes. Though emblems are undoubtedly intentionally communicative, they are definitionally symbols rather than indices. They carry an arbitrary mapping of form to meaning untethered to context. Because the primary function of emblems is lexical or semantic, they are more appropriately considered topic gestures rather than pragmatic gestures under Peirce's framework.

Defining pragmatic gestures requires going one step further than Peirce's semiotics. Nonverbal indices not only include non-lexical gestures like beats, but also deictic gestures. Deictic gestures index a contextualized referent by quite literally pointing to it. Clearly interpreting these gestures involves understanding the rules that relate signs to contexts, but deictic gestures are distinct from pragmatic gestures in that there is a referent to indicate. There are multiple perspectives on what makes up "reference," where reference may concern either the action of referring or the object of reference. Again, I follow Bates (1976), treating reference as that which "describes the use of a word or phrase by a speaker to stand for an entity-event in the outside world, or an entity-event in his own imagined world" (p.10). Labeling gestures like beats "non-referential gestures" is in line with this interpretation. That is, a co-speech beat gesture may be an action of reference making meaning within its context, but it is non-referential in that it does not literally or metaphorically point to any real or imagined entity.

For consistency, I use the term "pragmatic gesture" rather than "non-referential gestures" to refer to gestures that may lack semantic reference entirely. While the gesture references

interpretable meaning, it lacks an object or entity referent. Instead of commenting on the topic at hand, it comments on the interaction itself.

These categorical distinctions are functional. Many gesture forms fall into multiple categories depending on their usage in a particular discourse context. Palm-up gestures are a clear example of how one form may serve many functions, even simultaneously. A palm with fingers extended outward presenting new knowledge to the addressee is both deictic and interactive. An open palm metaphorically holding information is representational and interactive. The palm-up may be a rhythmic beat emphasizing an element, or it may emblematically mean 'don't know.' The relationships between form and meaning are not fixed, and they cannot – and should not – be ignored. The pragmatic functions of shrugs are not perfectly mapped to their forms, but neither are they entirely independent (Beaupoil-Hourdel & Debras, 2017; Debras & Cienki, 2012; Jehoul et al., 2017). In this dissertation I focus on gesture forms that serve pragmatic functions frequently, but not necessarily exclusively. This approach allows for an exploration of how children employ forms already in their gesture lexicon for new or variant pragmatic functions.

Appropriately classifying gesture types is essential for understanding gesture's role in language development. Different gesture types can play dramatically different roles in early communication, have different onsets and developmental trajectories, and predict different linguistic and cognitive outcomes. Developmental gesture research has primarily focused on topic gestures, typically associated with lexical and syntactic development. Preverbal infants produce pointing gestures with high frequency but seldom produce representational gestures (Acredolo & Goodwyn, 1985; Bates, 1976; Iverson & Goldin-Meadow, 2005; Özçalişkan et al., 2014). Occasionally pragmatic gestures are used as a comparison group for studies primarily

interested in topic gestures. Children's production of iconic gestures, but not non-referential beat gestures, in narrative contexts predicts later narrative outcomes (Demir et al., 2015; Vilà-Giménez et al., 2019); conversely, production of beat gestures, but not iconic gestures, in early naturalistic contexts predicts later narrative (Vilà-Giménez et al., 2021). Miscategorized gestures may mask or distort effects like these.

Considerably less work has explored the role of pragmatic gesture in early dialogic interaction and pragmatic language development, despite common pragmatic gesture forms like shoulder shrugs and palm-up gestures being some of earliest gestures produced by American English-learning infants.

1.2.2   Shrug Gestures

Shrug gestures have been classically considered emblems with the common verbal translation "I don't know" (Ekman & Friesen, 1969; Johnson et al., 1975; Kendon, 2004; D. Morris, 1994/2015). This emblematic use is frequently observed in caregiver-child interaction, produced by both interlocutors. Children and caregivers use exaggerated wrist rotation to comment on activity completion ('all done'), physical absence ('all gone'), and absence of knowledge ('don't know') (British English: Beaupoil-Hourdel & Debras, 2017; American English: Harris et al., 2017a; Iverson et al., 2008). These form-meaning pairings are consistent with Kendon's (2004) observation that across cultures, emblems typically express meanings related to the expression of personal attitudes and states of mind. In this case, shrugs announce an epistemic state or contextual condition.

However, categorizing shrugs as emblems is an oversimplification that obscures their wide range interactive uses. Unlike emblems, shrugs do not typically adhere to one fixed form in adult speech and may more accurately be described as a family of gesture forms or a composite

gesture (Debras, 2017; Streeck, 2009). While adults occasionally produce exaggerated or canonical shrugs, both shoulders raised and both palms up, more often speakers produce non-canonical variants of the shrug composite gesture. Speakers may involve one, both, or neither shoulder and/or wrist; rotate the wrist(s) between 0 and 180 degrees; produce one swift movement or hold a static position; vary handshape and include a head nod, shake, or tilt (Givens, 1977; Jehoul et al., 2017; Streeck, 2009).

Finally, it should be noted that shrugs have been observed in unrelated languages and cultures in both spoken and signed languages (see Cooperrider et al., 2018 for a review of the cross-linguistic use of palm-up gestures). While the present studies are limited to American English speakers, the cross-linguistic prevalence of the shrug is another complication for categorizing shrugs as emblems, which are definitionally conventionalized by individual linguistic communities (Ekman & Friesen, 1969).

1.2.2.1   Form

The *shrug family of gestures* includes a range of component forms, any number of which may be co-produced to create a recognizable "shrug." Smaller gestures performed by the shoulders, head, face, and hands create the *shrug complex* (Morris, 1994/2015), the shrug *compound enactment* (Streeck, 2009), or the *shrugging composite* (Givens, 1977). See Figure 1.1 for depictions of component forms.
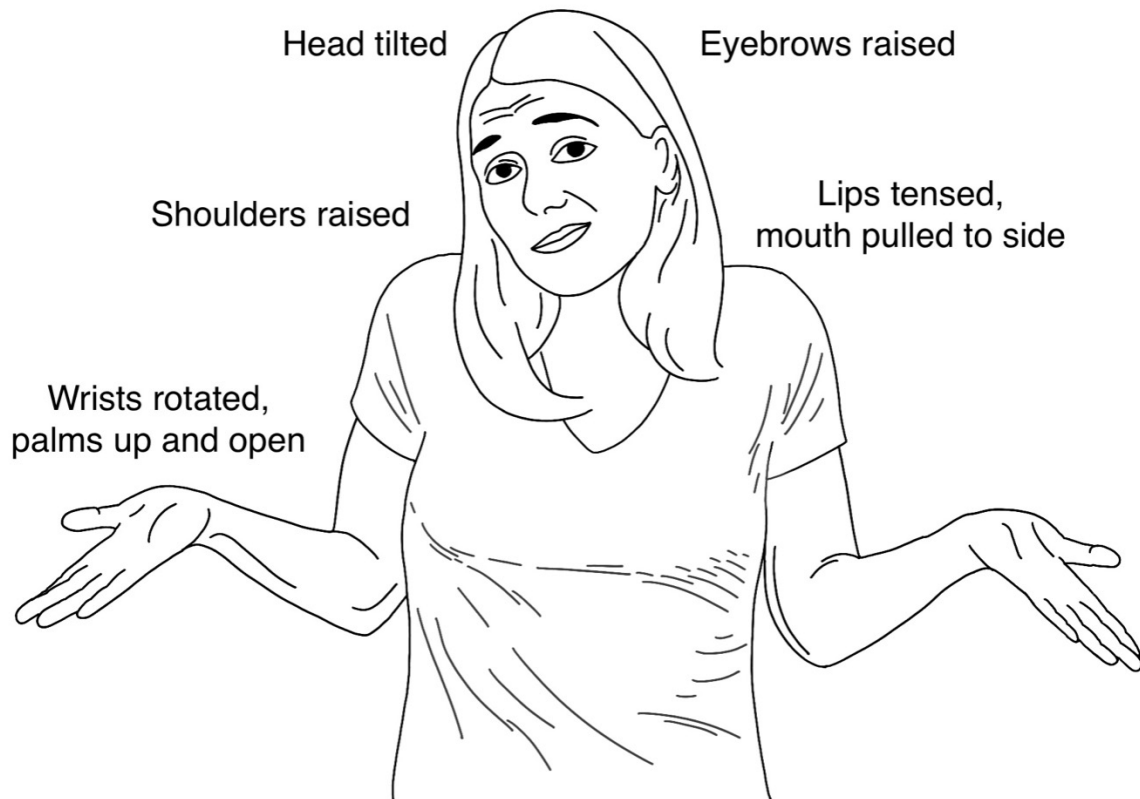
Figure 1.1: Component forms of the shrug composite gesture. Frequently referenced component forms include the shoulder raise, palm-up gesture, head tilt, brow raise, and mouth tension.

Although the word "shrug" may be colloquially synonymous with "shoulder shrug," many researchers have called specific attention to the wrist rotation as a meaningful gesture in its own right. Variations of the shrug component involving rotation of the wrist(s) with an exposed, upward-facing palm have been referred to as *palm-up lateral* or *palm-up epistemic* (Cooperrider et al., 2018), *palm up open hand* (Müller, 2004), *hand shrugs* (Ekman & Friesen, 1969), and *open hand supine* gestures (Kendon, 2004). This form, with or without other shrug components, is often classified functionally as interactive (Bavelas et al., 1992), recurrent (Bressem & Müller, 2017; Müller, 2017), or conduit metaphoric (McNeill, 1992). I refer to this component form as the "palm-up gesture", to highlight the most distinctive characteristic.

This wide range of possibilities for describing palm-up gestures' form and function complicates how we approach studying shrugs. Palm-up involvement co-occurs with shoulder raises and other shrug components so frequently and seamlessly that it should not be ruled out as part of the core shrug gesture. At the same time, if the isolated palm-up is highly versatile in meaning it is likely to affect functional use when acting as part of a shrug composite. Furthermore, both shoulder raises and palm-ups frequently combine not only with each other but with other classes of gesture, like conventional head movements (e.g., head nods), points, and beats (Debras & Cienki, 2012; Jehoul et al., 2017). Form components of the shrug composite gesture are neither entirely interchangeable nor entirely differentiable.

1.2.2.2   Meaning

Like form, shrugs can take many meanings in everyday conversation, even seemingly contradictory ones (see Table 1.1). For example, a shrug may express either ignorance ('I don't know') or certainty ('obviously'), either investment ('I mean it!') or disinterest ('whatever'), and either affiliation ('me too') or distancing ('you're on your own!'). The commonality among all these meanings is not immediately apparent, but gesture researchers have proposed several candidates for a core meaning. Notably, Cooperrider et al. (2018) suggest ignorance as the kernel meaning for the palm-up gesture; in their words, meanings grow from an 'absence of knowledge' and expand to other metaphorical absences, such as certainty, ability, or concern. Others propose openness (Müller, 2004), incapacity (Darwin, 1872/1998), and submissiveness (Boutet, 2018; Givens, 1977) as core meanings of shrugs and palm-ups.

Table 1.1: Observed meanings of shrug gestures

Meaning	Verbal equivalent	Citations
ignorance	"I don't know."	Barakat, 1973; Brookes, 2004; Calbris, 1990; Chu et al., 2014; Debras, 2017; Debras & Cienki, 2012; Graziano, 2014; Johnson et al., 1975; Jokinen & Allwood, 2010; Kendon, 2004; Payrato, 1993; Sparhawk, 1978
uncertainty	"I'm not sure.", "Maybe."	Barakat, 1973; Chu et al., 2014; Debras, 2017; Debras & Cienki, 2012; Ferre, 2012; Gawne, 2018; Graziano, 2014; Jokinen & Allwood, 2010; Payrato, 1993
obviousness, shared or common knowledge	"Obviously!", "As you know..."	Bavelas et al., 1992; Debras, 2017; Debras & Cienki, 2012; Graziano, 2014; Jehoul et al., 2017; Johnson et al., 1975; Kendon, 2004; Muller, 2004
disinterest, indifference	"Who cares?", "Whatever."	Calbris, 1990; Debras, 2017; Ferre, 2012; Streeck, 2009; Jokinen & Allwood, 2010; Payrato, 1993; Sparhawk, 1978; Streeck, 2009
agreement, affiliation	"Exactly right!", "Don't you agree?"	Bavelas et al., 1992; Calbris, 1990; Creider, 1977; Debras & Cienki, 2012; Streeck, 2009
disagreement, disaffiliation, distance	"That's not how I feel.", "I don't agree."	Barakat, 1973; Bavelas et al., 1992; Calbris, 1990; Debras, 2017; Debras & Cienki, 2012; Ferre, 2012; Muller, 2004; Streeck, 2009
submissiveness, incapacity, non-responsibility	"Don't ask me!", "Not my problem."	Bavelas et al., 1992; Debras, 2017; Debras & Cienki, 2012; Ekman & Friesen, 1972; Gawne, 2018; Givens, 1977; Graziano, 2014; Kendon, 2004; Muller, 2004
investment, certainty, commitment	"Of course!", "I'm sure."	Barakat, 1973; Calbris, 1990; Ferre, 2012; Jokinen & Allwood, 2010; Muller, 2004
absence	"I have nothing."	Brookes, 2004
inquiry, interrogatives, requests	wh-questions	Chu et al., 2014; Creider, 1977; Gawne, 2018; Kendon, 2004; McNeill, 1985; Rector, 1986
transferring or handling information, conduit metaphor	"For example...", "Let me tell you..."	Bavelas et al., 1992; Chu et al., 2014; Graziano, 2014; Gawne, 2018; Muller, 2004; Parrill, 2008; Streeck, 2009
beats, emphasis, exclamation	"Wow!", "It took for-ev-er."	Kendon, 2004; Ferre, 2012; Rector, 1986
turn-taking, floor negotiation	"Go ahead.", "Let me interrupt..."	Bavelas et al., 1992; Muller, 2004; Streeck, 2009

While this list of meanings is highly varied, there are some notable categories of meanings contained within. Several of the most common meanings fall on a spectrum of epistemic expression. Ignorance, uncertainty, and certainty reflect the speaker's assessment of their own knowledge; obviousness and some conduit meanings reflect the speaker's beliefs about common ground and shared knowledge. Another group of meanings share information about the speaker's affect or attitude. Disinterest, incapacity, and non-responsibility distance the speaker from their topic; conversely, interest, emphasis, and exclamation commit the speaker to their proposition. A final group serves to mark the speaker's degree of alignment with their interlocutor rather than their conversational contributions, like submissiveness, agreement, disagreement, and turn negotiation.

Co-speech gestures may serve multiple functions at once (Hostetter, 2011), and it is not difficult to imagine a single instance of a shrug serving both emblematic and pragmatic functions in the moment. A "don't know" shrug may simultaneously express disinterest. An "all done" shrug may simultaneously open the floor. If you ask me, "Who was the last winner of American Idol?" and I respond with a shrug, you could safely assume my response to mean "I don't know who won." However, if you have additional context about my interest in and exposure to that particular competition, you may read additional meaning in my gesture. I do not have the knowledge necessary to answer your question *and* "I couldn't care less" (disinterest), "I wish I knew!" (interest), "I can't help you there" (incapacity), or "I'm not the right person to ask" (non-responsibility). When a shrug performs two functions at once, it may be that we read two meanings independently associated with a shrug form; we recognize an ignorance emblem and a second pragmatic gesture expressing the speaker's affect or stance. Alternatively, we may be

reading ambiguity or variation in one core function rather than two distinct, convergent meanings.

We can recognize categorical similarities within these meanings, but this is not simply a case of subtle variations on a single core meaning. Even within functional groupings meanings are not interchangeable in a communicative act. For example, consider two clusters of function, epistemics and affiliation, and two meanings within each. Obviousness and ignorance refer to epistemic state; agreement and disagreement indicate a degree of affiliation between interlocutors. Different combinations of these four meanings actually result in different pragmatic effects.

1. obviousness + agreement = "Yeah of course! Duh!"
2. obviousness + disagreement = "You're crazy. Nobody thinks that."
3. ignorance + agreement = "I don't remember either."
4. ignorance + disagreement = "You seem certain, but I just don't know."
5. ?obviousness + ignorance = "Nobody knows why."
6. ?agreement + disagreement = "I'm right there with you; I don't agree with this either."

If shrugs can convey such a wide range of meanings, alone and in combination, how do listeners interpret a speaker's intended meaning? Clearly if the speaker communicates corresponding meaning in speech, the meaning of a co-speech shrug reinforcing that meaning is not difficult to decipher. A listener easily recognizes the difference between the utterance "I don't know" with a shrug and the utterance "I disagree" with a shrug. The shrugs may be meaningful, but they are not strictly necessary for effective communication.

The communicative burden of a shrug increases when it supplements meaning in speech rather than reinforcing it. As versatile as shrugs are, they can carry meaning independently of speech. You ask me, "What is the only scaled mammal?" I shrug and say, "The pangolin." Do I mean "maybe the pangolin, but I'm not sure?" Possibly I mean "I'm confident it's the pangolin," "obviously the pangolin," "the pangolin, which I remember you mentioned a minute ago," or

"the pangolin, isn't that right?" If you are able to infer which of these I intend to communicate, you have made that inference from nonverbal signals.

Additionally, shrugs can be produced as complete communicative acts, carrying the full communicative burden nonverbally. They can convey multiple, differentiable meanings without co-produced speech and in exactly the same conversational context. You ask me, "Where do you want to go to dinner on Friday?" In silence, I raise my eyebrows, tilt my head, and briefly raise one shoulder – "Up to you.". I quickly raise both shoulders and shift my gaze away from you – "I don't really care." I raise my eyebrows, make eye contact, and sustain a shoulder raise and palm-up with my fingers extended in your direction – "Nowhere in particular, what do you want?" In these nonverbal acts, properties of the extralinguistic context and the gesture itself create meaning.

In the above examples, there is likely a relationship between form and meaning, but the mapping between the many forms and functions of shrugs is unclear. There are at least two possible explanations for this variability: (1) many unique gestures have converged on the same general form or (2) there is a kernel, emblematic shrug gesture that maps outward into many forms and many meanings. The first case supposes the existence of a shrug ignorance emblem, a second bodily action of submission or openness, and a third metaphoric handling gesture and suggests these distinct gestures coincidentally share similar formal features. This explanation is plausible but difficult to demonstrate conclusively. However, a developmental perspective could give evidence to support or refute the alternative explanation. If the shrug's form-function mapping grows from a kernel or "seed" emblem, children should rely heavily on this kernel in the earliest stages of communicative development.

1.2.3 Multimodal Pragmatic Development

Previous research in the field of gesture studies has given rich description of how adults use pragmatic gestures like shrugs in language communities across the globe (see Table 1.1 and Cooperrider et al., 2018). Less is known about how children use these gestures or how shrugs fit into a multimodal understanding of pragmatic development. Although the domain of "pragmatic development" is expansive, this dissertation follows Ninio and Snow's (1996) focus on three cornerstone topics within research on the acquisition of verbal pragmatics. These include the acquisition of communicative intent and rules of communicative speech, the development of conversational skills, and the development of the ability to produce organized, extended discourse that is both cohesive and genre-specific. Nonverbal communication plays an important part in how these skills are borne out in face-to-face interaction between adult speakers, but many open questions remain about how nonverbal pragmatic skills develop alongside verbal pragmatics.

Ninio and Wheeler's (1986) taxonomy for coding verbal communicative acts at the levels of interchange, utterance, and discourse is generally relevant to gestural "utterances" as well. Stand-alone gestures that are produced without speech ("no-speech" gestures) are essentially translatable into analogical speech. Additionally, such codification systems may be applied to both child and adult no-speech gesture (Guidetti, 2002). Taxonomies of verbal acts, like Ninio and Wheeler's, have facilitated research on the transition from the pre-linguistic stage to single-word speech, demonstrating their potential for understanding the organization of nonverbal resources in other aspects of early development.

Appropriately using these systems to code the communicative intent of co-speech gesture is more complicated. How can we isolate the meaning of a gesture when it is intertwined with

meaning in speech? But this problem cuts both ways. Co-speech gesture often adds essential meaning to simultaneously produced speech, for example pointing to an object to disambiguate the meaning of an otherwise unspecified demonstrative. "I want that" may simply communicate desire or, with the right gesture, may call on the listener to hand over the indicated referent. Painting a complete picture of how children develop pragmatic competences requires consideration of how speech and gesture work together to draw on discourse and extralinguistic contexts to design turns, cooperate in conversation, and extend discourse.

It is not a novel proposal to suggest gestures play a role in pragmatic development, and there is a wealth of research on how topic gestures feature in early pragmatics. Infants use pointing gestures preverbally with multiple communicative intents. Early points can be proto-imperative requests for objects or proto-declarative acts of joint attention (Bates et al., 1975). Later, infants in the one-word speech stage combine pointing gestures with one non-redundant or "mismatched" word to create gesture-speech "phrases" before combining two words in speech (e.g., Butcher & Goldin-Meadow, 2000). Infants rely on iconic gestures to communicate actions before they have the words to do so but replace this use with speech as they acquire corresponding vocabulary (Acredolo & Goodwyn, 1985; Bates et al., 1975). Beyond this stage, infants continue to produce iconic gestures, but the relationship between speech and gesture changes. Iconic gestures add new meaning to an utterance not present in the simultaneously produced speech, adding non-redundant information about a topic (Özçalışkan & Goldin-Meadow, 2005).

Preverbal and early-verbal infants use gestures for basic pragmatic functions. Head shakes are not only used for proto-negation (verbally equivalent to "no") before the onset of speech but also to augment the pragmatic intent of "no" in assertives, directives, and refusals in the earliest

days of the one-word stage (Volterra & Antinucci, 1979). Pre-linguistic gesture systems allow children to acquire interactional routines (Behne et al., 2005; Tomasello, 2003; Tomasello et al., 2005). Between 1 and 2 years of age during the transition to speech, toddlers align intonation with gesture and other nonverbal communicative acts to effectively communicate pragmatic intent (Balog & Brentari, 2008).

There is already an abundance of work exploring pragmatic development across multiple modalities, but most of this research concerns multimodality broadly rather than focusing on the role of gesture specifically. In the section that follows, I review literature from several domains of discourse-pragmatics in which the importance of a multimodal perspective is well-established.

1.2.3.1 Epistemics

Children begin to use mental state words around age two (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Harris et al., 2017b). When verbs of cognition first onset (in English, typically "know" followed by "remember" and "think") they are used almost exclusively to refer to the child's own knowledge. It is rare for children under three to refer to their interlocutor's knowledge (references to a third party's knowledge are virtually nonexistent) or to explicitly contrast states of knowledge and ignorance (Bartsch & Wellman, 1995; Harris et al., 2017b). Children do not accurately report their own ignorance until age three, and they continue to over-report their own total and partial knowledge until the school-age years (Rohwer, et al., 2012; Ruffman & Olson, 1989; Sodian & Wimmer, 1987; Wimmer et al. 1988). Words like "know" seem to initially present a perceived truth about the world rather than clearly representing an assessment of knowledge, though it is difficult to tease apart cognitive and linguistic capacity in these measures of speech alone. Looking beyond speech tells more of this story. Behavioral measures suggest young children are able to correctly assess their own

ignorance by age 2 (Balcomb & Gerken, 2008; Bernard et al., 2015; Call & Carpenter, 2001; Lyons & Ghetti, 2011; Paulus et al., 2014) and can communicate this evaluation nonverbally before they can articulate it in speech (Harris et al., 2017a; Hübscher et al., 2019).

Notably, expressions of ignorance and expressions of uncertainty differ in their required skills as well as their ages of acquisition by children. Assessing one's state of ignorance recognizes an outright absence of knowledge; it is a binary state of having or not having access to information. On the other hand, accurately assessing uncertainty requires recognition of incomplete, insufficient, or unreliable knowledge, which first necessitates recognition of access to any relevant information. Developmentally, children express ignorance well before evaluating and expressing their degree of certainty. Children's earliest verbal and nonverbal metacognitive evaluations and expressions are typically limited to states of full knowledge and full ignorance. In speech, factive mental state verbs like "know" and "remember" are produced before non-factive verbs like "think" or "believe" (French: Bassano, 1996; English: Shatz et al., 1983). Likewise, by age two English-speaking children produce modal auxiliaries and adjuncts to express certainty or necessity, while modal expressions of possibility typically emerge in the third or fourth year of life (Hickmann & Bassano, 2016; Leahy & Carey, 2020; O'Neill & Atance, 2000; Papafragou, 1998).

The developmental importance of the distinction between ignorance and uncertainty is highlighted by the interaction of verbal and nonverbal channels. Kim, Paulus, Sodian, and Proust (2016) assessed 3- and 4-year-old children's ability to communicate full and partial knowledge states. Although 4-year-olds have the capacity to evaluate and verbally express uncertainty, the authors found that even the oldest children overestimated their partial knowledge state when verbally reporting it and were significantly worse at reporting partial knowledge than ignorance.

However, the authors also coded the children's production of uncertainty gestures (e.g., shoulder shrugs) and found these gestures to more accurately reflect both partial knowledge and ignorance states. Hübscher et al. (2019) made similar observations, finding that 3- to 5-year-old children frequently overestimated their knowledge verbally while simultaneously producing gestural and prosodic uncertainty signals more reflective of their partial knowledge state. These results suggest observing children's nonverbal behavior may be a window into their capacity to evaluate degree of knowledge and certainty and a preview of epistemic expression not yet available in speech.

1.2.3.2 Extended Discourse

The complementary relationship of speech and gesture in pragmatic development is also apparent in children's extended discourse. By "extended discourse" I refer to organized units of interaction that follow culturally conventionalized principles of appropriate use, like narratives, explanations, definitions, and arguments (following Ninio & Snow, 1996). Children produce both linguistic and gestural devices for cohesion within discourse units with increased frequency and complexity across the pre-school and school-age years, a reflection of both cognitive and communicative capacities (Alamillo et al., 2013; Casillas et al., 2016; Colletta et al., 2010; Levy & McNeill, 2013; Peterson & Dodsworth, 1991).

While verbal and nonverbal measures of narrative development can be tightly intertwined, they are differentially affected by contexts of culture and language. In a comparison of French, American, and Italian children's narrative retelling, Colletta et al. (2015) found differences not only on linguistic measures like overall length of narrative but also gesture measures like frequency of use or preference for a particular class of gesture. These effects of language may be explained by grammatical differences, cultural differences around storytelling practices and

schooling, or both and are supported in other research on narrative retelling in children and adults (Gullberg, 2006; Iverson et al., 2008; Kunene, 2010; Yoshioka, 2008).

The importance of nonlinguistic pragmatic skills is particularly well-established in the domains of polite stance and persuasion strategies. Adults mark politeness multimodally, simultaneously drawing on vocal and gestural strategies together with linguistic strategies for politeness in vocabulary and syntax. Prosody and gesture work together to take polite stance, as speakers adapt their speech to use slower rhythm, quieter voice, and lower pitch and mitigate nonverbal signals by producing smaller and slower co-speech gestures (Bolinger, 1986; Brown & Winter, 2019; Brown, et al., 2014; Hübscher et al., 2017; Winter & Grawunder, 2012). Not only are prosody and gesture interrelated in adults' expression of politeness, they also operate as "sister systems" in the development of polite stance across early childhood (Hübscher & Prieto, 2019).

Toddlers use both linguistic and paralinguistic strategies to mark politeness and differentiate types of control acts, but the two modalities do not share a single developmental trajectory. For example, Ervin-Tripp, Guo, and Lampert (1990) found that young children used three types of speech acts to make requests or otherwise influence others' behavior, each determined by both vocal quality and specific linguistic features. Vocally, children used "pushy" voice with urgent control acts, soft voice and longer pauses for polite acts, and "whiny" voice for retries. In speech, children frequently used intensifiers and name-calling for urgent acts, hedge phrases for polite acts, and syntactic changes for retries (e.g., retrying interrogative structure with imperative, or vice versa). Wellman and Lempers (1977) found that when toddlers (ages 2;2-3;0) did not receive a response they reformulated and tried again 50% of the time by adding

communicative features like vocatives, pointing gestures, and both verbal and nonverbal attention-getters.

Children use these vocal and verbal request strategies as early as 2 years of age, but they do not onset concurrently. A closer look at very early development of politeness and control acts sheds light on how careful examination of nonlinguistic resources – comparing nonlinguistic to linguistic resources and comparing within more refined subcategories of each – can disentangle how multiple modalities are integrated in pragmatic development. Children typically use lexical and syntactic strategies for politeness and control before vocal and behavioral strategies (Ervin-Tripp et al., 1990). A more complete multimodal picture of early polite stance sheds light on the relationships between these skills and perhaps mechanisms of acquisition. At the highest level, without differentiating nonverbal and verbal communication, we see that children have specific strategies to mark polite stance as young as age two. Separating speech and non-speech, we see that verbal strategies come before nonverbal strategies. Within the speech modality, linguistic politeness forms onset before vocal resources. Narrowing further, we see that it is not simply that linguistic precedes nonlinguistic. Instead, "taught forms" (e.g., saying "please" or echoing permission-seeking syntax) account for the earliest linguistic strategies and are often used inappropriately (Ervin-Tripp et al., 1990; Gleason & Weintraub, 1976; Greif & Gleason, 1980; Read & Cherry, 2009). Nonlinguistic vocal resources appear next with gestures following soon after, so that by age four children use nonlinguistic markers of politeness in both voice and gesture to appropriately take polite stance and mark relevant social relationships (Ervin-Tripp et al., 1990; Hübscher, Garufi, & Prieto, 2019). Although at a surface level linguistic skills precede related nonlinguistic skills, as children enter the school-age years they have sophisticated, effective, and adaptive strategies of politeness with both voice and gesture while verbal strategies

for taking polite stance remain mostly limited to fixed forms (Baroni & Axia, 1989; Bates, 1976; Nippold et al., 1982; Pedlow et al., 2004). The critical conclusion here is that fully understanding the development of polite stance requires fitting together all these pieces of the acquisition timeline.

1.2.3.3   *Later Childhood and Adolescence*

The potential to overlook subtler points of pragmatic development remains true beyond early childhood. While lexical and syntactic skills "explode" in the first years of life, pragmatic development notably extends well beyond early childhood. In particular, many pragmatic skills essential to monitoring and negotiating dialogical conversation see dynamic growth in adolescence and young adulthood as children are also exposed to and immersed in new types of social contexts (Larson & McKinley, 1998; Nippold, 2000; Ochs & Schieffelin, 2016; Stivers et al., 2018). Although we therefore would not expect adult-like usage of pragmatic gesture to appear in the pre-school years, we may see some initial patterns in shrug gesture use in early childhood that shape later developmental change.

Although children produce gestures to express ignorance and uncertainty before they acquire the resources to do so verbally, children still do not use nonverbal epistemic signals like shrugs with adult proficiency even in middle childhood or early adolescence. Krahmer and Swerts (2005) elicited and video-recorded responses to factual questions from 7- and 8-year-old children and adults in a feeling of knowing paradigm, then in a second study had separate groups of children and adults estimate the speaker's certainty in their response ("feeling of another's knowing"). Although the children both produced and comprehended nonverbal and paralinguistic expressions of uncertainty, they were significantly worse than adults at both expressing their own feeling of knowing and judging feeling of another's knowing. These

multimodal epistemic markers continue to develop through late childhood, increasing in frequency and complexity between ages 8 and 11 years and beginning to adapt to social setting at age 11 (Visser et al., 2011).

Across the school-age years and into adolescence, children develop more sophisticated linguistic strategies for coordinating in dialogue (see Table 1.2 for a selection relevant publications). There is a shift away from primarily egocentric talk toward strategic employment of allocentric talk in multiple kinds of dyadic interaction, including activity coordination, narrative retellings, and persuasion attempts (e.g., Asher, 1976; Clark & Delia, 1976; Short-Meyerson & Abbeduto, 1997). Although toddlers use taught forms (e.g., "please") for politeness early, the acquisition of lexical and syntactic strategies for politeness is a slow process that continues at least into late childhood. (e.g., Baroni & Axia, 1989; Ervin-Tripp et al., 1990; Nippold et al., 1982).

Table 1.2: Selection of research on discourse-pragmatics in adolescence

Publication	Subject ages (years)
Perspective-taking, egocentric vs. allocentric talk	
Anderson et al., 1994	7-13
Asher, 1976	7-11
Clark & Delia, 1976	7-14
Papafragou & Tantalou, 2004	4-6
Short-Meyerson & Abbeduto, 1997	4-5
Politeness and persuasion	
Axia & Baroni, 1985	5-9
Axia, 1996	4-8
Baroni & Axia, 1989	5-7
Bates, 1976	4-7
Ervin-Tripp et al., 1990	2-11

Table 1.2 Continued: Selection of research on discourse-pragmatics in adolescence

Publication	Subject ages (years)
Nippold et al., 1982	3-7
Pedlow et al., 2004	10-11
Piché, Rubin, & Michlin 1978	11-15
Extended discourse, narrative	
Berman & Slobin, 1994	3-9
Colletta et al., 2010	6-10
Kernan, 1977	7-14
Martin, 1983	6-10
McClements, 1976	5-11
Romaine, 1984	6-10
Scott, 1984	6-12
Social status, ritual talk	
Anderson, 1990	4-7
Berko-Gleason & Weintraub, 1976	2-10
Hollos, 1977	7-9
Conversation	
Brinton & Fujiki, 1984	5-9
Dorval & Eckerman, 1984	7-18
Larson & Mckinley, 1998	12-18
Rafaelli & Duckett, 1989	10-14
Schober-Peterson & Johnson, 1993	8-10
Turkstra, 2001	13-21

Children continue to develop strategies for organizing extended discourse at least into adolescence, particularly in terms of connectedness and cohesion. Before age 6, children primarily use additives and temporal connectives in narrative retellings, rarely marking relationships across utterances with other conjuncts or attitudinal expressions. In middle and late

childhood, children increase their overall use of conjuncts, use a wider range of items including causal forms, and decrease use of the earliest acquired forms (e.g., "and"). Even by age 12, however, children rarely use attitudinal conjuncts like "in fact" or "obviously" (e.g., McClements, 1976; Scott, 1984). Children's use of evaluative devices in narratives and explanations changes over this period as well (7-14 years old), moving from implying feelings without direct quotation to an increasing reliance on explicit declaration of feelings with supporting direct quotations (e.g., Kernan, 1977; Romaine, 1984).

1.2.4   Dissertation Contributions

This dissertation focuses on children's production of pragmatic shrug gestures due to their early onset and their paradoxical property of being both highly emblematic and pragmatic. The existing research reviewed so far highlights the importance of studying early epistemic expression with a multimodal perspective, demonstrating children's capacity to transmit accurate information about knowledge state above and beyond their speech production. However, differentiating between two broad categories of modality – verbal and nonverbal – cannot speak to the pragmatic functions of gesture specifically. Sensible everyday conversation between adults relies heavily on epistemic work expressed in a variety of means, but gesture plays a unique role due to the range of possibilities for how it relates to speech. Like prosody, gestures can be produced with speech to emphasize or supplement meaning, but gestures need not accompany speech at all (Ekman & Friesen, 1969; Payrató, 1993; Teßendorf, 2013). Like filled and unfilled pauses, gestures can fill gaps in speech to hold the floor, but gestures may also be produced simultaneously with speech, efficiently adding to or reinforcing a verbal meaning (Goldin-Meadow et al., 1993; McNeill et al., 1994; Özçalişkan & Goldin-Meadow, 2005). Like behavioral displays such as waiting to respond or averting eye gaze, gestures may accentuate the

speaker's hesitance or inability to provide a response, but gestures are typically limited to a set of possible meanings and are unlikely to be motivated by something outside the communicative context (Gerwing & Bavelas, 2013; Krauss, 1998).

It remains unclear how children integrate gestures specifically into multimodal epistemic expression and whether there are meaningful distinctions of form and function within the category of "uncertainty gestures." Furthermore, while it is evident that children use these gestures for some kinds of epistemic functions early on, it is unclear when or how they begin to use these same forms to express other epistemic meanings (e.g., obviousness, common ground) or non-epistemic meanings (e.g., interest, negation, turn-taking).

The topics that form the theoretical basis for this dissertation are not rooted in a single literature or field of study. Epistemic expression, multimodal pragmatics, gesture as a tool for language learning, stance-taking – these are areas of interest to developmental and cognitive psychology, gesture studies, and linguistics. In this dissertation I build on research from across these disciplines. I aim to partially reconcile the substantial variation of terminology and methodology in order to put these conceptual frameworks in conversation with one another through an interdisciplinary perspective.

The Language Development Project corpus offers a unique opportunity for a mixed-methods approach to exploring early childhood interaction. The corpus includes thousands of hours of video of spontaneous and elicited caregiver-child interaction, following children from infancy to young adulthood. These recordings and simultaneously collected standardized assessments offer the quantitative outcome measures ubiquitous in developmental research, such as mean length of utterance, productive vocabulary, and narrative comprehension scores. Additionally, the videos offer rich, naturalistic interaction that may be analyzed qualitatively,

describing characteristics of the physical, social, and linguistic contexts in which children are immersed in the home.

To reiterate, we question how to carve up the pragmatic gesture space not simply for the sake of doing so, but because it has practical significance to how we approach the study of multimodal language development. Children's acquisition of speech and gesture go hand in hand, but not all types of referential gestures are similarly integrated within language development. Just as deictic and iconic gestures have distinct roles to play in lexical and syntactic development, different types of pragmatic gestures may play different roles in pragmatic development. For example, given that palm-up gestures are often produced without speech and are produced by preverbal infants, children might use them to express meanings they cannot yet communicate in speech, allowing for successful communicative actions like making requests or calling attention to objects and events. On the other hand, because beats are produced with speech by definition, we should not expect early beats to precede or predict productive vocabulary. If these predictions are accurate, failing to distinguish between palm-up gestures and beats, or between gestures with and without speech, may misattribute these effects or miss them entirely.

One current limitation in our understanding of how gesture factors into pragmatic development is a tendency to divide concepts of multimodal communication into linguistic and nonlinguistic components, where the latter category may encompass not only gestures but also intonation, laughter, eye gaze, facial expression, and a host of other paralinguistic vocal and nonverbal resources. In some respects this approach is actually more appropriate than studying these resources discretely, given how tightly interwoven they are in face-to-face interaction. Gestures and prosody, for example, are often thought of as two sides of the same sociopragmatic

coin (Bara, 2010; Perniss, 2018). Speakers combine intonation and nonverbal cues like beat gestures to mark topical focus in speech (Dohen & Lœvenbruck, 2009; J. Kim et al., 2014) and take epistemic stance (Dijkstra et al., 2006; Krahmer & Swerts, 2005; Roseano et al., 2016). This link begins in the first few years of life, with gesture and prosody operating as "sister systems" of pragmatic development (Esteve-Gibert & Prieto, 2018; Snow, 2017). Infants and young children use gestures and intonation together for lexical emphasis, expressing ignorance, and mitigating requests with a simple politeness register, all before the onset of parallel pragmatic resources in speech (Hübscher & Prieto, 2019).

There is clear motivation for and merit to a generalized perspective on nonverbal pragmatics, but it remains worthwhile to understand the individual roles of component parts. Prosody and gesture both support the development of some pragmatic skills in speech, but it is unclear how these two developing systems may support each other. Some "nonlinguistic" communicative resources are dependent on linguistic systems (e.g., intonational contour, pauses, beat gestures) while others can be fully expressive without speech (e.g., shrug gestures, facial expressions). All of these nonlinguistic resources may work together within a conversational context or over development, but in order to recognize individual contributions of prosody, gesture, or anything else we must distinguish between pragmatic strategies that necessarily and optionally make meaning together with speech. Relatedly, recognizing discrete components of nonverbal pragmatics allows us to ask whether features that do not appear to operate within the conventionalized structures of a given language still operate systematically.

The studies reported in Chapters 3, 4, and 5 attempt to home in on the role of pragmatic gestures – and shrug gestures in particular – in the conversations of children and adolescents. In

the next chapter I discuss how principles of conversation analysis may be useful in a corpus

analysis of child language and review the annotation systems used in the studies that follow.

2   Methodology

2.1   Methodological Motivation

Much of our current understanding of multimodal pragmatic development has come from experimental methodologies eliciting language or non-verbal communication from children. Experimental conditions prompt children to produce particular kinds of communicative acts outside of naturalistic dialogue. There are clear advantages to researching language development within such controlled conditions. Experimental techniques allow researchers to systematically manipulate variables to tease apart interconnected features of communication. Additionally, a targeted approach allows for larger-scale data collection than observational approaches and the potential for more representative sample of participants.

On the other hand, experimental conditions are, by design, fundamentally different from spontaneous communication. Although these methods allow for unique insight into children's communicative capacities and mechanisms of acquisition, it is difficult to generalize these findings to children's real-world experience of communication and language learning. Experimental contexts can have vast differences from the interactions children actually participate in. Children may be in an unfamiliar location with an unfamiliar interlocutor performing tasks they would rarely or never encounter at home or school. Because conversational turns are designed within the situational and discourse contexts they occur, it is imperative to assess the appropriateness of experimental contexts.

Corpus analysis with observational data and elicited conversation tasks can counter some or all of these limitations of experimental methodologies. Corpus data are often collected in familiar contexts, commonly capturing children interacting with family in their own homes as

part of the same everyday activities they would engage in without an observer present. Recordings of extended, unguided interactions offer opportunities to explore how a multitude of linguistic and non-verbal communicative devices are integrated within a single communicative action, as part of a broader discourse context, or across development. The value of corpus methods in language development research is best exemplified by the CHILDES database (MacWhinney, 2000). Over 3,000 publications used CHILDES corpora in the project's first 25 years (1983-2008) with topics as diverse as phonemic perception distinctions, socio-moral behavior, politeness strategies, kin terms, and caregiver book-reading style.

However, corpus analysis is not without its own drawbacks. Although a defining characteristic and primary advantage of corpus analysis is its operation at a large scale, acquiring and handling sufficiently large quantities of natural language data creates other limiting factors. The processes of transcription, annotation development, and annotation by multiple trained coders can be extremely time and labor intensive. As a result, there are practical limitations to the amount of data that can be collected, processed, and analyzed in any single study using a corpus. This creates a problem of trade-off between the participant sample size, the frequency or length of sessions for longitudinal collection, and variables to include in annotation and analysis.

This problem is amplified for more qualitatively driven work and work with young children. Although resources like the CHILDES databases supply ample opportunity to analyze children's natural language, only a portion of these corpora are suitable for studying children's spontaneous co-speech gesture. There are at least three limitations in this regard. First, gesture research requires audio-video recordings, which can be both more expensive and more intrusive than audio-only recordings. Second, gestures need to be adequately captured on video even in highly uncontrolled environments. Recording video of young children in the home often means

disruptions by siblings and erratic or surprising actions by young participants which obscure view or require strategic intervention by the data collector. Finally, analyses of nonverbal communicative resources like co-speech gesture remain particularly dependent on manual annotation. There have been significant advancements in automation for transcribing audio recordings, analyzing vocal features of speech, and processing large-scale text data, but most research on pragmatics and pragmatic development continues to require subjective human judgments from trained coders. Motion-capture technologies offer promising potential (Brown et al., 2021; Dodane et al., 2019), but are not yet suitable for studies of naturalistic gesture outside of a controlled, laboratory setting or studies with young children.

2.1.1 Conversation Analysis

Conversation analysis (CA) methodologies offer a unique perspective in the study of dialogical conversation. True conversation analysis cannot be used with a large-scale corpus analysis, and the application of formal CA methods to large databases of language data runs counter to their intended purpose. Classical methods of conversation analysis selectively sample corpora for small-scale analysis, focusing on one or several conversational dyads and limiting analyses to a small slice of dialogue, rarely more than a few minutes in length. Still, there is a growing body of work using these methods to study children's conversation (Bateman & Church, 2017; Gardner & Forrester, 2010; Sidnell, 2016). These methods arguably give the most comprehensive qualitative description of how speakers cooperate to organize and maintain conversation, but the specificity of these micro-analyses can be problematic in attempts to integrate the CA perspective with a developmental psychology perspective. Typically psychological studies of language development demand much larger participant samples and a sufficiently large amount of quantifiable data to perform at least basic descriptive statistics.

The developmental insights "true" CA offers can, at a minimum, be useful points of entry for selectively using CA-based analyses for contextualized, naturalistic child interaction at a larger scale (see Casillas & Hillbrink, 2020 for a review of communicative act theories in language development research). Specifically, the broad frameworks for sequence and preference organization have immediate relevance to the questions at hand. Stivers, Sidnell, and Bergen (2018) demonstrate how these practices are both theoretically and practically applicable to a corpus suitable for a quantitative description of children's question-answer sequences. In this work, the authors collected 48 hour-long recordings of interactions among 95 school age children and annotated question sequences using an existing coding procedure developed for a cross-linguistic comparison of adults' question-answer sequences (Stivers & Enfield, 2010). This work demonstrates not only how application of CA-based coding systems to large childhood corpora is feasible, but additionally how coding systems based on CA principles may be designed for use with both child and adult language corpora.

As a very high-level description, *sequence organization* is a central tenet of conversation analysis, operationalizing the normative conversational rules interlocutors abide by as they coordinate turn-taking in dialog. At its core, sequence in conversation is built on expectations of "nextness" between one social action and the reflexive action that follows (Enfield et al., 2010; Sacks, 1992a, 1992b; Schegloff & Sacks, 1973; Stivers, 2012). A speaker's utterance selects the next speaker, who then has an obligation to take the next turn. The selected speaker additionally has an obligation to construct this next turn in the form of a relevant response. Both the selection of next speaker and the relevance of a responsive action to an initiating action are determined by the structural format of the initial turn, and what counts as relevant for a particular initiating format may be determined by categories of adjacency pairs. As defined by Schegloff and Sacks

(1973), *adjacency pairs* are characterized as two turns produced by different speakers. The two turns are adjacently placed, relatively ordered, and pair-type related. That is, first-pair part actions always precede a specific type of second-pair part actions (or set of actions). Greetings should be met with greetings, offers with acceptances or refusals, requests for information with informative answers or an admittance of ignorance, etc.

Preference organization, as it relates to sequences of social actions, builds on this framework by describing preferred and dispreferred formats of responses when multiple relevant responses are normatively acceptable (Holtgraves, 2000; Pomerantz & Heritage, 2012; Robinson, 2020; Sacks, 1992b). For example, an offer may be met with acceptance or refusal. In a linguistic community that observes the common preference principle "accept, affirm, and agree when possible," the acceptance is the preferred format. Refusal is a dispreferred, though equally relevant, response.

Although systems of sequence and preference are far more complex than presented here, these fundamental properties may be extended to domains that have received relatively little attention from conversation analysts. Specifically, a high-level view of these systems can be used for a better understanding of how children design turns and how speakers of all ages incorporate gesture into their turn design. In both cases, though for different reasons, we expect pragmatic meaning in speech to be less complex or incomplete. We may not expect the kinds of nuanced features of turn design to be present or clearly discernible, but we should still expect children's turns and turns which are largely reliant on gesture to be organized along these simplified lines.

Stivers and Rossano (2010) add to models of turn design by asking not how speakers solicit a relevant response from their listener, but instead how they solicit any response at all. In their conceptualization of *response mobilization*, speakers employ multiple resources of turn

design within a social action, creating a scalar model of response relevance. Briefly, they highlight several response-mobilizing features typical of questions and other actions that strongly pressure a response: interrogative syntax, interrogative prosody, epistemic asymmetry, and sustained eye gaze. Speakers may add one or more of these features to a "telling" (e.g., announcements, assessments) that might not otherwise receive a response. For example, a speaker's negative assessment at the dinner table – "The chicken is a little dry." – may not receive an uptake from the listener. The speaker may then indicate the expectation for a response by looking to their listener or by adding a tag phrase (e.g., "isn't it?"). The speaker may opt to employ both eye gaze and an interrogative tag rather than one or the other to create even more response pressure.

Although Stivers and Rossano do not directly address the contribution of gesture to response mobilization, they do recognize that it must have a role. The functional class of interactive gestures can serve precisely this purpose. Interactive gestures are specialized for dialogue. They indicate the addressee, typically with a loosely formed point or palm-up, in order to facilitate coordination within the interaction. Interactive gestures call upon the listener to continue taking conversational turns by referencing contributions to the dialogue, recognizing shared knowledge, deferring to the listener, and seeking agreement or uptake (Bavelas et al., 1995; Bavelas et al., 1992). Other pragmatic gestures that are not directed toward the addressee often share critical characteristics with the features presented by Stivers and Rossano. Shoulder shrugs and lateral palm-up gestures can signal listener-tilted epistemic asymmetry by indicating a speaker's ignorance or uncertainty. Additionally, these forms commonly accompany other response mobilizing resources, in particular rising or interrogative intonation. Although conversation analysis does not traditionally include a comprehensive breakdown of co-speech

gestures, it is clear that pragmatic gestures like shoulder shrugs, lateral palm-ups, and addressee-indicating interactive gestures play a role in response mobilization and sequence organization more generally.

I combine principles of conversation analysis with formal coding to describe patterns of communication in these naturalistic data. By "formal coding," I refer to the process of annotating verbal and nonverbal features of an interaction using a formalized and replicable coding scheme, following Stivers' (2015) argument for increasing acceptance of social and behavioral formal coding within practices of conversation analysis. Stivers demonstrates in that article, and in an abundance of work before and after, that these practices are not "heretical" to conversation analysis methods. Rather, she argues, coding social interaction is a natural extension of the characterizations of conversation making up the core of CA. These characterizations already serve to indicate patterns of interaction and imply generalizability. Traditional micro-analyses of conversation are undoubtedly essential to identifying highly nuanced patterns of use in dense, complex, truly spontaneous language. Formalized coding is one way to use these impressionistic claims as a springboard in searching for the distributional evidence necessary to bring findings from conversation analysis to a broader audience. It is not dismissive of the value of purely qualitative work standing on its own merit. Rather, it translates this value into the language of psychologists, educators, medical professionals, etc. Further, a systematic way of understanding features of language as quantifiable variables creates an avenue to studying associations with variables outside interaction, such as demographic information, standardized assessment outcomes, and survey data.

This kind of formal coding is essential for studying pragmatic development with naturalistic data, particularly for questions about nonverbal resources. Using easily quantifiable

measures of verbal pragmatic abilities along the lines of mean length of utterance or productive vocabulary is insufficient for a comprehensive understanding of face-to-face dialog as inherently contextualized. Conversation is uniquely human and requires human judgment to adequately consider multiple layers of extralinguistic context.

This dissertation uses a manual annotation system inspired by conversation analysis and applies it to data from the Language Development Project, described in detail below. This longitudinal corpus includes naturalistic interaction between parents and children across twelve time points in early childhood and another two sessions in early adolescence. It offers a rare opportunity to study communicative development not only across infancy and toddlerhood but also into later stages of language acquisition. The data in this corpus were collected to study both speech and co-speech gesture by design, creating an ideal opportunity for exploring the development of multimodal pragmatics.

The Language Development Project corpus allows for descriptions of gesture use both within and across the early childhood and early adolescence observational periods. The twelve early childhood sessions offer insight into productive onsets of early acquired gesture forms, with implications for how we talk about "pragmatic" and "non-referential" gestures as generalized classes, discussed further in Chapter 3. Additionally, the form-meaning mappings of children's earliest pragmatic gestures can offer evidence for or against a core or "kernel" meaning or emblem, such as 'ignorance' (Cooperrider et al., 2018) or 'openness' (Müller, 2004). Finally, a longitudinal comparison of gesture forms in the pre-school years can potentially distinguish between changes in formal use driven by communicative development and those that follow as consequences of motor development.

Describing gesture use in early adolescence, as well as across the early childhood and early adolescence periods, can offer additional context for interpreting results at both stages of development. For example, Chapter 4 pursues the proposal that 'ignorance' is a kernel meaning for the shrug gesture and that this meaning is emblematically linked to a kernel shrug form. This line of inquiry is predicated on the premise that very early in pragmatic development children are not yet able to convey nuanced pragmatic intents drawing on extralinguistic contexts. By isolating young children's gestures produced without speech – and so without the supports of meaning in a speech context – we may identify a plausible kernel form and meaning. By early adolescence, children have moved through later stages of pragmatic development and are more proficient at incorporating meaning from social and discourse contexts. If adolescents' use of emblematic shrugs is no longer contingent on the presence of speech, this would support the interpretation that pragmatic functions of shrugs grow from an early emblem. This proposition is explored in Chapter 5. On the other hand, if there was no significant difference in functional use between the pre-school and middle school years, this interpretation would be called into question. It might imply that pragmatic and emblematic shrugs are in fact two gestures that converge onto a single form, or that conversational skills in the gestural modality develop at a slower pace than those in speech.

Annotation and analyses are limited to the communicative acts produced by children that include a pragmatic gesture. That is, these studies do *not* include annotated data from caregivers' speech or gesture, nor do they include annotation from children's utterances that do not have a co-produced shrug or beat gesture. These omissions limit our ability to draw conclusions about mechanisms of acquisition and gesture's relationship to speech over development.

The verbal and nonverbal input children receive from their parents is an essential piece of all early language acquisition and almost certainly plays a critical role in the development of nonverbal pragmatics. However, the intention of the studies that follow is not to understand mechanisms of acquisition. Rather, these annotation systems enable us to answer questions about children's functional use of shrug gestures. This descriptive work of children's productive communication serves as an essential first step. It is necessary to first ask whether, when, and how children employ nonverbal pragmatic resources before delving into the effects of input.

The decision to focus on children's gestures and co-produced speech follows similar reasoning. Looking more broadly at all of children's communicative acts, including those without gesture or with topic gestures, would allow for targeted questions about the timing of co-developing verbal and nonverbal conversational skills like response mobilization, strategic interruption, and backchanneling. For example, detailed sequence organization coding of all child speech and gesture creates the opportunity to examine the relationships among linguistic and nonlinguistic response mobilizing features (e.g., interrogative syntax, sustained gaze), including interactive gestures. This might provide insight into both understudied aspects of pragmatic development as well as adding evidence to the debate of response mobilization as gradient (Couper-Kuhlen, 2010; Schegloff, 2010; Stivers & Rossano, 2010). Do gestures operate as part of a response mobilization gradient, layering pressure for response on top of interrogative syntax, sustained gaze, or rising intonation, or do they convey similar meanings without layering pressure? When do these gestures emerge in development in relation to other response mobilizing features? Do they emerge alongside verbal markers of epistemic asymmetry, or – as is the case in many other aspects of early language development – do children use nonverbal resources like gesture and eye gaze before analogous verbal resources?

Although these questions cannot be fully with gesture acts alone, we can set the stage for future inquiry. Children produce shrugs extremely early, but do they use them interactively in early childhood *at all*? If it seems interactive shrugs grow from a kernel or seed emblem, what properties of this emblem allow interactive functions to blossom? Is there more merit to an approach that treats palm-up gestures as one element of a shrug composite gesture or to isolating this form as a unique gesture that functions independently from the shoulder raises it often accompanies? Again, it is necessary to lay a descriptive foundation answering these questions around whether and how children produce pragmatic gestures before exploring complex entanglements with speech.

In this dissertation, I intend to use a focus on children's gesture to lay groundwork for the questions that follow naturally. With this in mind, the annotation schemes used in Chapters 4 and 5 are designed to be equally applicable to coding properties of sequence organization and pragmatic function to both speech and gesture for speakers of all ages.

2.2   Overview of Methods

2.2.1   Language Development Project

2.2.1.1   Participants

Subjects were a subset of families participating in an ongoing longitudinal study of language development at the University of Chicago, which includes 64 typically-developing children and their families. Participants were recruited from the greater Chicago area through mailers to targeted zip codes and advertisements placed in a free, monthly parenting magazine. Responding parents were interviewed for background characteristics and to confirm a monolingual, English-speaking household. The final sample was demographically representative of the greater Chicago area, as reported in the 2000 U.S. Census, in terms of race/ethnicity,

41

household income, and parent education. See Rowe (2008) for additional information regarding participant recruitment and demographics of the full sample.

Capitalizing on annotation from prior research, these analyses are limited to a subsample of 18 children (8 girls). This cohort was first selected by Cartmill, Hunsicker, and Goldin-Meadow (2014) to maximize range of early verbal skill. Inclusion was determined by averaging children's mean length of utterance (MLU) across the first five observations (between 14 and 30 months), then selecting the 6 subjects with highest MLU, ($M = 2.04 \pm 0.10$; 3 girls), lowest MLU ($M = 1.22 \pm 0.06$; 2 girls), and median MLU ($M = 1.52 \pm 0.06$; 3 girls).

Table 2.1: Participant demographics and visit completion

Subject ID	Sex	Race/ethnicity	MLU group	Household income	Maternal education	Completed visits EC (max 12)	EA (max 2)
42	M	White, Non–Hispanic	Low	$15,000-$34,999	Some College or Trade School	11	2
48	M	White, Non–Hispanic	Low	>$100,000	Advanced Degree	12	2
77	F	Black	Low	<$15,000	Some High School	12	2
78	M	White, Non–Hispanic	Low	$35,000-$49,999	Advanced Degree	12	2
84	M	White, Non–Hispanic	Low	>$100,000	Some College or Trade School	12	2
105	F	White, Non–Hispanic	Low	$50,000-$74,999	Bachelor's Degree	12	2
24	F	Black	Middle	>$100,000	Advanced Degree	12	2
33	M	Black	Middle	$50,000-$74,999	Some College or Trade School	12	0
37	F	White, Non–Hispanic	Middle	$75,000-$99,999	Bachelor's Degree	12	2
62	M	White, Non–Hispanic	Middle	>$100,000	High School or GED	12	1
74	F	White, Non–Hispanic	Middle	>$100,000	Bachelor's Degree	12	2
88	M	White, Hispanic	Middle	$75,000-$99,999	Advanced Degree	12	2
29	F	Mixed/other race	High	>$100,000	Advanced Degree	12	0
43	M	White, Non–Hispanic	High	$50,000-$74,999	Bachelor's Degree	12	2
44	F	Black	High	$35,000-$49,999	Some College or Trade School	12	2
50	M	White, Non–Hispanic	High	$50,000-$74,999	Bachelor's Degree	12	2
92	M	White, Non–Hispanic	High	>$100,000	Bachelor's Degree	12	2
103	F	Mixed/other race	High	$75,000-$99,999	Bachelor's Degree	12	2

The subsample of families was diverse in terms of household income, parent education, and race and comparable to the larger sample of 64 families (Table 2.1). The participants included 11 White Non-Hispanic, 1 White Hispanic, 4 Black/African-American, and 2 children of mixed/other race. Household income was reported in six brackets and ranged from less than $15,000 to more than $100,000 per year. Based on the bracket midpoints, approximate average yearly household income was $74,000. Maternal education ranged from less than 12 years (no high school diploma or equivalent) to more than 18 years (advanced or professional degree). The most commonly reported education level was completion of a bachelor's degree.

2.2.1.2 Data collection

Data used in this dissertation are from two stages of the longitudinal study beginning in 2001, which occurred in early childhood (before age 5) and early adolescence (ages 11-14).

In the early childhood (EC) stage, families were visited in their homes every 4 months when children were between 14 and 48 months of age. At each of the 12 home visits, 90 minutes of spontaneous interaction between children and their primary caregiver(s) was captured with audio and video recording. Families were instructed to behave as usual, as though the experimenter was not there. The videos capture a wide range of typical day-to-day activities from early childhood, such as reading books, playing with toys, doing jigsaw puzzles, watching television, and eating meals. One family did not complete the home visit at 50 months. The remaining 17 families completed all 12 early childhood sessions.

In the majority of sessions, caregivers engaged with children for all or nearly all of the visit. Because families were not guided towards any kind of interaction, videos also include

children playing by themselves or with other family members while the primary caregiver is off-camera. Rarely, caregivers left children to play by themselves for most or all of the 90-minute session. These cases reflect the variation in caregiver-child interaction across and within families and are typically not treated differently than any other visits. However, extremely low caregiver participation in a session constrains children's opportunities for dialogic interaction and alters the quality of children's communicative acts. These sessions are excluded in some analyses where this dramatic discrepancy is problematic. The nature of the problem and exclusion criteria are discussed in Chapter 4 and Appendix B.

Sixteen of the 18 families in the subsample continued on to the early adolescence (EA) stage (Table 2.1). In this stage, families were visited in the summers following 5th - 8th grades. Visits were timed to coincide with the ending of the school year rather than subject birthdays, resulting in larger age range at each visit. This study uses data from the post-5th grade visit (mean age = 11.44 years, SD = 0.41) and post-7th grade visit (mean age = 13.44 years, SD = 0.38).

In these two sessions following 5th and 7th grade, caregivers and children were recorded working together on a series of 12 tasks. Tasks were designed to encourage cooperation and elicit higher-order thinking talk in four domains: art, science, math, and social. Families were allowed to move on to the next task at any time but were prompted to begin the next task after five minutes, so that each session lasted approximately one hour or less. Additional description of the tasks and discussion of families' engagement during the sessions is provided in Chapter 5 and Appendix B.

2.2.1.3 Base transcription and gesture annotation

The studies described within this dissertation began with existing, prepared transcripts. For the early childhood (EC) stage, transcripts included speech and basic gesture annotation. For the early adolescence (EA) stage, transcripts included speech but not gesture.

All spontaneous speech by participant children and primary caregivers was transcribed in Microsoft Excel. Speech was transcribed verbatim but not phonetically and included conventionalized communicative sounds (e.g., "mmhm", "ouch"). In the EC visits, caregiver speech directed to other adults was not transcribed unless the child was clearly attending to it. Speech was transcribed at the utterance level, with breaks between utterances decided by multiple criteria including pause length, grammatical structure, and intonational contour. To ensure high inter-coder reliability, agreement was calculated for both word units and utterance boundaries. Before independently transcribing videos, coders were required to reach 95% agreement with model transcripts for word and utterance metrics. Approximately one-third of transcripts were partially double-coded by a second expert transcriber. Transcripts with less than 90% agreement on either metric were rejected and re-transcribed until satisfactory agreement was reached.

For the EC visits, transcribers simultaneously annotated communicative gestures from both caregivers and children alongside speech transcription. This first-layer gesture annotation including codes for form (e.g., "point", "thumbs up", "iconic", "beat"), body part(s) and side(s), and approximate gloss. Agreement for gesture annotation followed the same procedures as transcription reliability described above.

2.2.2   Additional Annotation and Analysis

Using these existing transcripts as a base, five additional layers of coding were performed only for child-produced gestures: (1) pragmatic function of communicative acts (EC only), (2) palm-up form completeness (EC + EA), (3) shrug gesture form (EA only), (4) shrug gesture function (EC + EA), and (5) sequence organization (EC + EA). The first layer is used as the basis of analysis in Chapter 3, describing differences in utterance-level pragmatic use of two non-referential gesture forms. The second, fourth, and fifth layers are used in Chapter 4 to identify an emblematic kernel shrug form and meaning. The third, fourth, and fifth layers are used in Chapter 5 to extend these findings and describe gesture function in more advanced stages of pragmatic development.

All analyses were conducted using R (Version 4.1.3; R Core Team, 2022) and the R-packages *lme4* (Version 1.1.28; Bates et al., 2015), *nnet* (Version 7.3.17; Venables & Ripley, 2002), *rstatix* (Version 0.7.0; Kassambara, 2021), and *tidyverse* (Version 1.3.1; Wickham et al., 2019).

2.2.2.1   *Annotation schemas*

Here I describe the general intention for each layer and offer an overview of the essential annotation. Full coding manuals with examples are included in Appendix A. Reliability metrics are included in Appendix C.[1]

---

[1] Other properties that were coded but ultimately not used in this dissertation include palm-up orientation, facial and vocal features of response mobilization, gesture-speech relationship (for co-speech gestures), repair sequences, performative acts, and reported speech. These codes are not discussed here but are included in the full coding manuals provided in the appendix.

First, all communicative acts containing a palm-up[2] or beat gestures were coded for *pragmatic function of the communicative act* (Table 2.2)

Second, palm-up gestures in the early childhood stage were coded for *completeness of presentation* (adapted from Hundertmark, 2016). Palm-ups were coded as *complete* if they were produced with pronounced (~180 degrees) rotation of the wrists and had an overall stiff or crisp form with a visible "freeze" at the peak of movement. Palm-ups were coded as *reduced* if they did not meet the criteria for complete forms. Palm-up completeness was coded referencing video files without audio.

Third, the early adolescence sessions were coded for *presence and form of shrug gestures*. This process included marking the presence of a gesture in the shrug family as well as the component forms used, including shoulder raises, palm-ups, beats, head shakes, head nods, or head tilts. Palm-up forms were coded for completeness (as above).

Fourth, for both EC and EA visits all shrug gestures with palm-up or shoulder raise component forms were annotated for *pragmatic meaning* and whether they were produced together with speech or as a gesture-only communicative act (Table 2.3). Additionally, head gestures produced without other shrug forms were coded for meaning in the EA visits, but not the EC visits.

Finally, all shrug gestures were annotated for features of *sequence organization*, including positioning in the turn sequence (Table 2.4), request type (Table 2.5), communicative intent

---

[2] Gestures coded as "flip" or "shrug with flip" in the original first-layer gesture annotation were considered "palm-ups" for pragmatic function and palm-up completion coding.

(Table 2.6), and preference format (Table 2.7). Again, head gestures produced without other

shrug forms were coded for sequence organization features in the EA visits, but not the EC visits.

Table 2.2: Annotation, pragmatic function

Pragmatic function	Example
Unbiased assertions	
Explanations, declaratives	*I know how to make a towel with paper.*
Information responses	(following inquiry) *She's at school today.*
Epistemic uncertainty	
Confirming questions	*Does this go on my head?*
Information-seeking questions	*Where is – sister and Mama?*
Expressing absence or ignorance	*I don't know!; All gone.*
Epistemic agreement	
Acknowledgments	*Ok Daddy.; Uhuh.*
Affirmations, agreements	*Yes, you're right.*
Negation	
Corrections, contradictions	*But I'm switching now.*
Flat negations, disagreements	*No, you're wrong.*
Requests	
Imperatives	*Put it here so Mommy can see.*
Action requests	*Can we go upstairs?*
Object requests	*Mama binkie.; Can I have more juice?*
Expressive acts	
Exclamations	*Wow!*
Markings, performatives	*Thank you!*

Table 2.3: Annotation, shrug meaning

Meaning	Description	Verbal equivalent
	Investing actions	
Investment	commitment to proposition, obviousness, certainty	*I'm telling you, obviously, clearly, I'm sure, I feel strongly about this, as we all know, of course*
Disinvestment, full	ignorance, inability	*I don't know, I can't answer that, how should I know?, can you help me?*
Disinvestment, partial	uncertainty, unwillingness, disinterest, distance from topic, lack of commitment, non-responsibility, disinterest	*I'm not sure, I think so, I guess so, maybe, I'm no expert, if I had to guess, I don't care, whatever, it doesn't matter*
	Affiliating actions	
Affiliation	agreement, acknowledgment, following	*ok, me too, I understand, I agree, you're right, I'm following, keep going, we're on the same page*
Disaffiliation	disagreement, confusion, not following	*I disagree, I don't like what you're saying, I'm confused, you're wrong, I'm not following you*
Affirmation	affirmation, confirmation	*yes, correct*
Negation	negation, disconfirmation	*no, incorrect*
	Other	
Claim floor	claim next turn, hold the floor	*I'll talk next, I'm not done yet, don't interrupt, let me interrupt*
Cede floor	pass turn, open the floor	*go ahead, what do you think?, I'm done talking, feel free to interrupt me*
Literal absence or completion	inability due to absence or completion	*all gone, all done*
	Excluded	
Other	non-interactional meaning	*on the one hand (metaphoric contrast), that one in front of me (deictic with palm or nod form)*
Unclear	cannot infer intended meaning	*unintelligible speech; missing untranscribed context*

51

Table 2.4: Annotation, sequential positioning

Position	Description	EC Example	EA Example
First-pair part (FPP)	initiates a turn sequence and receives a response; non-contingent	CHI: I WANT TO COLOR. PCG: ok, we can color.	CHI: WHOSE TURN IS IT TO READ THE TASK? PCG: you can go next.
Second-pair part (SPP)	responds to a first-pair part and closes a turn sequence; contingent	PCG: let's have an apple for snack. CHI: NO MAMA, COOKIE.	PCG: well you have to take math, right? CHI: UHHUH, MATH AND ENGLISH, TOO.
No turn transition	neither initiates nor closes a turn sequence, may attempt to open a sequence but receives no response; non-contingent	CHI: I WANT DORA. CHI: mama I want Dora.	CHI: I THINK ALEX IS WORSE FOR STEALING. CHI: well actually they are both pretty bad.
Backchannel	offers feedback about comprehension or agreement without claiming the floor	PCG: remember at gymnastics yesterday (CHI: YEAH.) Miss Emily said you did a good job.	PCG: well some kids need more help than you though (CHI: RIGHT.) so maybe not everyone takes the same math.
Scripted exception	protocol artificially limits turn sequence, e.g., a question directed to the experimenter does not receive a response because experimenter is instructed to stay silent	CHI (to EXP): WHY AREN'T YOU TALKING? EXP: (no response)	CHI (to PCG): DID YOU GET TO CHOOSE YOUR CLASSES? EXP: sorry it's time for the next task.
Unclear	positioning cannot be reasonably inferred due to insufficient context, unintelligible speech, etc.	PCG: where'd the last piece go? CHI: (UNINTELLIGIBLE) PCG: about time for lunch I think.	PCG: I guess (inaudible). CHI: (INAUDIBLE) PCG: I think we're done.

Table 2.5: Annotation, request type

Request Type	Description	EC Example	EA Example
Open	interrogative syntax, no closed set of predictable responses; e.g., wh-questions	Where did Daddy go?	How did it work when you were in high school?
Closed	interrogative syntax or intonation, closed set of predictable responses; e.g., yes/no questions, multiple choice questions	Is Daddy at work?	So should students take all the same classes or all different classes?
Imperative	imperative syntax, issues directive	Read it again.	Um, you answer it first.
Non-request	all other acts; e.g., declaratives, assessments	Sissy is at school.	You have to take math to graduate.

Table 2.6: Annotation, communicative intent

Communicative Intent	Description	EC Example	EA Example
Scenario	comments on behavior in the immediate scenario, announces next action, narrates pretend play, does not present fact or opinion	I'm going to go get my Dora. Is the princess is going to the ball, too?	I think we answered this one enough. Can we use a pencil or no?
Objective	presents (dis)provable claim, objective facts, may be correct or incorrect	It's red! Did Grandma say it?	The Beethoven one was faster tempo. You can go to jail if steal from a store.
Subjective	presents (dis)agreeable claim, subjective opinions, assessment, announcements of cognitive or emotional state	I don't like the blue ones. Thirteen is really really old!	Did you like Billy Joel better than Beethoven? But I'd feel bad about lying to the old man.

Table 2.7: Annotation, response preference format

Preference	Description	EC Example	EA Example
Preferred	response takes relevant format *and* affirms, acknowledges, agrees, accepts	PCG: you want some yogurt? CHI: MMHM.	PCG: did you notice anything the same? CHI: THEY HAD THE SAME, LIKE, MELODY KIND OF.
Dispreferred	response does not take relevant format *or* negates, ignores, disagrees, rejects; inability or unwillingness to respond not due to lack of knowledge	PCG: was Robbie at preschool today? CHI: I DON'T LIKE PRESCHOOL ANYMORE. *or* PCG: you better turn that tv off right now. CHI: NO!	PCG: so which is worse, cheating or stealing? CHI: THEY'RE BOTH BAD. *or* PCG: I think everybody needs to know math. CHI: NO NOT EVERYBODY.
Ignorance	special case of dispreferred; inability to provide a preferred response due to lack of knowledge	PCG: where did the scissors go? CHI: I DON'T KNOW WHERE SCISSORS.	PCG: well when do you think Beethoven was alive? CHI: I HAVE NO IDEA.

3   Two Dimensions of Pragmatic Gesture in Early Childhood

3.1   Introduction and Literature Review

Topic gestures are interrelated with speech across the development of pragmatics. Children use pointing gestures and head gestures for multiple communicative intents when they lack the verbal lexicon necessary to encode intent entirely in speech (Bates et al., 1975; Butcher & Goldin-Meadow, 2000; Guidetti, 2005). Similarly, infants use iconic gestures to communicate about actions before they have the vocabulary to do so and later use iconics to augment, rather than replace, verbs once they enter the lexicon (Acredolo & Goodwyn, 1985; Bates et al., 1975; Özçalışkan & Goldin-Meadow, 2005).

However, a multimodal model of pragmatic development based primarily on the functions of topic gestures minimizes the role of an entire class of gestures that are pragmatic by definition. Like "pragmatic development," the term "pragmatic gesture" can be broad and ambiguous. In this chapter, the essential characteristic of pragmatic gestures is their function of communicating information about the interaction itself. They operate outside the topic of talk and provide conversational meta-comments about epistemics, attitude, and turn-taking instead of – or in addition to – contributing semantic meaning about the topic. A thorough discussion of pragmatic gesture as a phenomenon is presented in Chapter 1.

Children begin to produce pragmatic gestures very early in communicative development, but the ambiguity around how we talk about these gestures in adults' conversation likely contributes to a relative lack of research on their role in children's early interactions. Infants typically produce head nods, head shakes, palm-up gestures, and shoulder shrugs before their second birthday and before acquiring corresponding lexical terms (Beaupoil-Hourdel et al.,

2015; Benazzo & Morgenstern, 2014; Guidetti, 2005). While these gestures frequently function pragmatically in everyday conversation, they also function emblematically as nonverbal translations of 'yes,' 'no,' and 'don't know.' Because these gestures commonly function as emblems in preverbal and early verbal communication, it can be easy to overlook where they may also serve pragmatic or interactive functions.

Existing research on children's earliest pragmatic gestures suggests they serve an important role in communicative development, but this work can face two limitations. First, these studies often compare pragmatic gestures to representational gestures but rarely look within the class of pragmatic gestures. (Colletta et al., 2015; Colletta et al., 2010; Esteve-Gibert & Prieto, 2014; Vilà-Giménez et al., 2020; also see Vilà-Giménez & Prieto, 2021 for a review). Second, pragmatic gestures are often subsumed under an umbrella category of nonverbal communication along with prosody, facial expression, pauses, and behavior (Borràs-Comes et al., 2011; Dijkstra et al., 2006; Hübscher et al., & Prieto, 2017). One way to begin addressing these limitations is to highlight two particularly salient dimensions of gesture: gesture form and gesture-speech relation.

3.1.1   Gesture Form

Gesture *form* describes the key visual or physical characteristics of a gesture. Formal properties of gesture are those features which are recognizable and describable without any additional context, such as handshape, orientation in space, and temporal phases of movement (McNeill, 1992). The form dimension exists independently of function, even in cases where the two are tightly linked. For example, an outstretched index finger usually takes a deictic function, visually indicating the literal or metaphorical location of a referent. Less often, one outstretched finger might emblematically stand in for the number one, iconically represent a perch for a bird,

or interactively signal to an interlocutor the desire to take the next turn. Divorcing form and function recognizes the possibility that a gesture form does not enter a child's lexicon via the most typical or predictable form-function associations.

Differences in form have been foundational to the study of topic gestures. Even within functional categories, formal categories matter. For both adults and children, the form of pointing gestures can vary based on the type of referent indexed and the pragmatic intentions of the speaker, such as referent individuation, discourse relevance of location, and imperative acts (Cochet & Vauclair, 2014; Kendon & Versante, 2003; Wilkens, 2003). Iconic gestures presented in character-viewpoint more effectively communicate a referent's relative positioning than observer-viewpoint iconics, while the reverse is true for information about properties like speed and shape (Beattie & Shovelton, 2002). This differentiation between character- and observer-viewpoint is also relevant to the development of narrative production, where both spontaneously produced and trained character-viewpoint gestures are associated with better narrative structure (Demir et al., 2014; Parrill et al., 2018).

The form dimension of functionally pragmatic gestures can be a bit muddled. For example, Bavelas, Chovil, Lawrie, and Wade (1992) introduce *interactive gestures* as an entirely functional category of gesture but at the same time impose strict formal constraints. The authors' limitation of interactive gestures to addressee-indicating forms excludes gestures that perform the same interactive functions but differ in handshape, direction, or both (e.g., lateral palm-ups) and may include gestures that deictically indicate the addressee as topic-relevant rather than referencing the interaction. On the other end of the spectrum, the shrug family of gestures seems to have almost no agreed upon formal constraints, leading to a wide range of interpretations for

their pragmatic functions in conversation (Boutet, 2018; Cooperrider et al., 2018; Debras, 2017; Givens, 1977; Jehoul et al., 2017).

Consequently, gestures labeled pragmatic, interactive, or discursive frequently discount formal features. This is not an unreasonable approach to answering questions exclusively about function, but as we study pragmatic gesture over development there may be much to learn from how gestures' emerging functions relate to their forms. This chapter will focus on two of the most commonly produced forms of pragmatic gesture in early childhood, *beats* and *palm-up gestures*.

McNeill (1992) describes *beats* as gestures that lack discernible meaning and are instead recognized by their prototypical movement. They are "small, low energy, rapid flicks of the fingers or hand" (p80) that emphasize the semantic content of temporally matched speech. Despite lacking semantic meaning independent from meaning encoded in speech[1], beats contribute pragmatic meaning by rhythmically aligning with prosodic cues to frame discourse.

Children begin producing beats around their second birthday alongside the emergence of multiword utterances (Nicoladis et al., 1999) and increase use of beats through the school-age years as discourse-pragmatic and narrative skills increase in complexity (Colletta et al., 2015, 2010; Mathew et al., 2018). A small body of recent work has investigated connections between children's production of beats and narrative. Training children to produce oral narratives with beats improves children's narrative structure and oral fluency (Vilà-Giménez & Prieto, 2020). Longitudinal observational studies of early spontaneous beat production and later narrative abilities have shown mixed results (Vilà-Giménez et al., 2020; Vilà-Giménez et al., 2021), but

---

[1] Though see Yap et al. (2018) for discussion of "hidden meaning" in some beats.

this may be explained by methodological differences in how the formal category of beat gestures was defined. See Vilà-Giménez and Prieto (2021) for a systematic review of beats' predictive role in early language development.

Palm-up gestures go by many names, including *palm-up open hand* (Müller, 2004), *palm lateral* or *palm-up presentation* in the *open hand supine* family (Kendon, 2004), *palm-up epistemic* (Cooperrider et al., 2018), *flips* (Harris et al., 2017; Vilà-Giménez et al., 2021), *palm-revealing* or *conduits* (Chu et al., 2014), and *hand shrugs* (Ekman & Friesen, 1969; Johnson et al., 1975; Morris, 1994/2015), among others. This range of terminology reflects some disagreement in what exactly "counts" as the palm-up form, but all these gestures share the fundamental formal features of outward wrist rotation with loose or extended fingers, with the palm exposed upward or outward at the gesture's peak. For the sake of simplicity, gestures that meet these basic formal criteria are referred to here as *palm-up gestures* or simply *palm-ups*.

Though palm-ups are often considered emblematic or conventional gestures (Johnson et al., 1975; Morris, 1994/2015), they also perform pragmatic and interactive functions in spoken and signed languages across the globe, like epistemic stance-taking, distancing speaker from topic or interlocutor, managing the conversational floor, and metaphorically handling information (see Cooperrider et al., 2018 for a review).

Palm-up gestures are some of the earliest gestures produced by children, typically appearing before two years of age and often before an infant's first words (Acredolo & Goodwyn, 1985; Beaupoil-Hourdel & Debras, 2017; Harris et al., 2017; Iverson et al., 2008). Infants use palm-ups as emblems for literal absence ('all gone') and ignorance ('don't know') (English: Beaupoil-Hourdel & Debras, 2017; Harris et al., 2017; Italian: Caselli, 1983; Graziano, 2014). Toddlers use palm-ups to mark other epistemic states before they do so in speech

(Catalan: Hübscher et al., 2019; German: Kim et al., 2016). By school entry children produce

presentational palm-ups to mark discourse (Graziano, 2014). These studies have established that

young children use palm-ups with multiple meanings, but do not speak to how different

functions arise, persist, or mutate over development.

It is easy to recognize that beats and palm-ups share several key properties, such as acting

in complement with prosody and emphasizing discourse content. As a result, they are often

grouped together functionally as "non-referential gestures" or simply as a looser category of

"beats" (Dimitrova et al., 2016; McNeill, 1992; Shattuck-Hufnagel et al., 2016; Vilà-Giménez et

al., 2020). The logic of treating beat and palm-up forms as one and the same is predicated on the

fact that beats and palm-ups can be extremely similar in both form and function. However, it

fails to account for the fact that despite this potential for similarity, many or most uses of beats

and palm-ups serve entirely different functions with entirely different forms.

3.1.2   Gesture-Speech Relation

A second dimension of pragmatic gestures key to multimodal pragmatic development is

gesture-speech relation, the way meaning in the nonverbal modality interacts with meaning in

the verbal modality within a single communicative act (Özçalışkan & Goldin-Meadow, 2005).

For example, pointing to a table *disambiguates* "that" in the utterance "look at that," *reinforces*

the meaning of "table" in the utterance "sit at the table," and *supplements* the utterance "sit

down" by adding information about where to take a seat.

Perhaps the most basic level of gesture-speech relation is *substitution,* where a

communicative act exists only in the nonverbal modality. A speaker can respond to "which

table?" by pointing to the table without any co-produced speech whatsoever. Both children and

adults frequently substitute gestures for speech, but the substitution relationship may have

special relevance in early childhood, when productive vocabulary is limited, overall speech production is low, and isolated gestures frequently serve as full communicative acts (Bates et al., 1975; Iverson & Goldin-Meadow, 2005).

This substitutive relationship is particularly useful for the comparison of beats and palm-up gestures because it highlights one of the most problematic issues with treating the two forms as one and the same. Palm-ups may be produced either with or without speech, but beats are never produced without speech. As co-speech gestures, palm-ups perform a range of pragmatic functions, some very similar to beats, but they are also meaningful on their own. Using the example above, another communicative and valid response to "which table?" would be to indicate lack of knowledge by performing a palm-up instead of a speech act. Beats, in contrast, must be produced with speech *by definition*. Beats serve to add emphasis to some element of a verbal utterance. A beat without speech would not be a beat at all. Both palm-ups and beats may (at least in theory) take a reinforcing, disambiguating, or supplementing relationship to co-produced speech, but only palm-ups may take a substituting relationship.

3.1.3   Present Study

The study presented in this chapter explores how pragmatic gestures fit into a multimodal approach to the study of pragmatic development. I look at two dimensions of pragmatic gestures likely to reveal meaningful categorical distinctions within pragmatic gestures, if such distinctions are in fact worth making. First, I divide along gesture *form* between beats and palm-up gestures. Second, I divide communicative acts by the *gesture-speech substitution relation*: co-speech gestures produced simultaneously with a verbal utterance and "no-speech" gestures comprising a full communicative act which substitutes for a verbal utterance.

Focusing on these two dimensions of how children construct communicative intent with gesture and speech can illustrate the importance of giving pragmatic gestures the same considerations given to topic gestures. Is it sufficient to treat pragmatic gestures as a unified functional category? If not, what distinctions within this category are significant?

In this chapter I ask three primary research questions. First, do gestures differing across these dimensions have different developmental onsets? Second, do differences across these dimensions relate to differences in frequency of use over development? Finally, do differences across these dimensions relate to differences in the pragmatic functions of communicative acts in early childhood, both overall and across early development?

3.2   Methods

Participant recruitment, demographics, and home visit data collection procedures are described in Chapter 2.

3.2.1   Transcription & Annotation

Home visits were transcribed for speech and simultaneously coded for gesture form and gloss as described in Chapter 2. See Huttenlocher et al. (2010) for transcription procedures and reliability and Rowe and Goldin-Meadow (2009) for first-level gesture coding procedures and reliability.

Children's communicative acts produced with a beat or palm-up gesture were coded for pragmatic function using an annotation scheme adapted from Ninio et al. (1994) and originally described in Vilà-Giménez et al. (2021). Most communicative acts included a verbal utterance with co-speech gesture, but gesture-only acts produced in silence ("no-speech") were also included. Annotation was conducted using transcripts only, without access to audio or video.

This method allowed coders to make judgments about pragmatic function at the utterance level, based solely on the speech and fundamental gesture characteristics of a communicative act rather than cues from prosody, facial expression, or visual information in the physical context.

Coding included four broad categories of pragmatic function with an additional six subcategories, following the commitment space semantics framework from Krifka (2015):

1. *Unbiased assertions.* Communicative acts in this category were those with declarative or explanation illocutionary force and no markers of modality.

2. *Biased assertions or questions.* These acts express a degree of commitment to the truthfulness of the proposition, often in relationship to an interlocutor's contribution. Subcategories of biased assertions were epistemic uncertainty (including questions and requests for information), epistemic agreement, and negation.

3. *Requesting speech acts.* These acts include both imperative and interrogative forms of requests for action (rather than information).

4. *Expressive speech acts.* These acts included exclamations and utterances primarily serving to convey emotion ("hooray!", "ouch!") as well as performative markings (e.g., "the end", "sorry").

Further description and examples are provided in Table 2.2. See also Appendix A for full coding manual and additional examples. All transcripts were annotated for pragmatic function by one coder. Inter-rater reliability was determined by having a second coder double code 20% of the transcripts. Inter-rater reliability for pragmatic function was high (Cohen's κ=.846, p<.001). Annotation disagreements were resolved by coder consensus.

3.2.2 Data Analysis

I performed a longitudinal analysis examining the relationships between children's pragmatic gesture form, co-presence with speech, and pragmatic function between 14 and 58 months of age.

3.3 Results

Children's gestures were compared along two dimensions: (1) gesture form (beat vs. palm-up) and (2) presence or absence of co-produced speech (co-speech vs. no-speech). Theoretically this two-by-two analysis creates four possible constructions for communicative acts. However, a defining characteristic of beat gestures is a rhythmic marking of speech. Beat forms are therefore necessarily co-speech. In order to avoid problems of multicollinearity, the analyses that follow combine the two dimensions into the single variable of *gesture construction* for the communicative act. The three possible gesture constructions were no-speech palm-ups, co-speech palm-ups, and co-speech beats. Figures are formatted to clearly differentiate gestures along both dimensions, but all analyses were performed using the single gesture construction variable. Full models and statistical tables are included in Appendix D.

3.3.1 Age of Onset

A first descriptive analysis compared age of production onset for each of the three possible gesture constructions: no-speech palm-ups, co-speech palm-ups, and co-speech beats. Age of onset was defined as the child's age in months at the session where the child was first observed to produce the gesture.[2]

---

[2] The overall rarity of pragmatic gestures in children's spontaneous interaction necessitated loose criteria for onset. Using stricter criteria (e.g., requiring the gesture construction to be produced at two sessions in a row) resulted in a similar but non-significant trend.
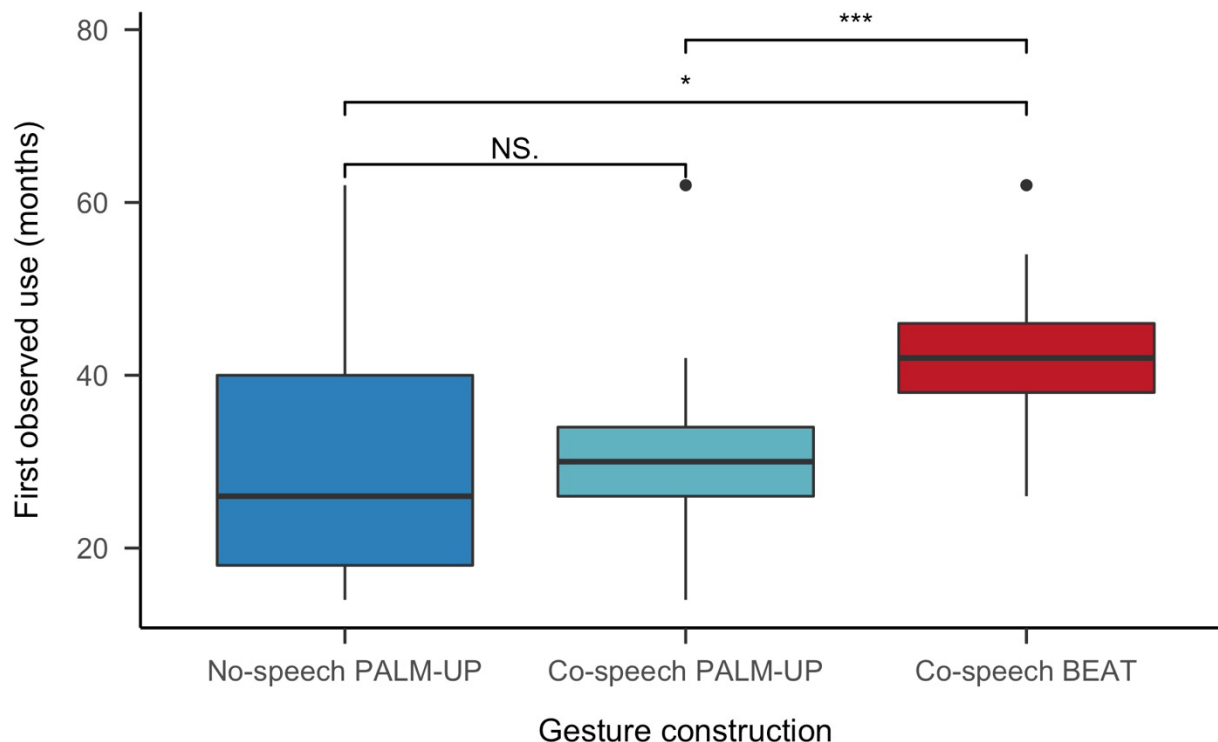
Figure 3.1: Distribution of mean production onsets for pragmatic gesture constructions. Palm-up gestures with and without speech onset one year before beats.

Both co-speech and no-speech palm-ups were typically first produced around the 30-month visit. Mean onset for no-speech palm-ups was 30.44 months (SE = 3.71) and 31.11 (SE = 2.44) for co-speech palm-ups. Average onset for beat production was one year later, at the 42-month visit (M = 42, SE = 2.10; Figure 3.1).

A one-way ANOVA showed significant variation in mean age of onset ($F(2, 51)$ = 5.23, p = .009). Pairwise comparisons using t-tests with Bonferroni adjustment revealed significant differences in mean age of onset between co-speech beats and co-speech palm-ups (p = .004) and between co-speech beats and no-speech palm-ups (p = .030). There was no significant difference in age of onset between the two palm-up forms.

The results indicate the onset of these pragmatic gesture constructions is more tied to gesture form than the presence or absence of speech.

3.3.2 Frequency of Production

I next built a Poisson mixed-effects regression model to compare children's frequency of use for each gesture construction across the twelve sessions. The final model included fixed effects for gesture construction (factorial; no-speech palm-up, co-speech palm-up, or co-speech beat), child age in months (numeric), number of communicative acts produced by the child at each session (numeric, centered) and interactions between gesture construction and both age and number of communicative acts and a random slope for age by subject.
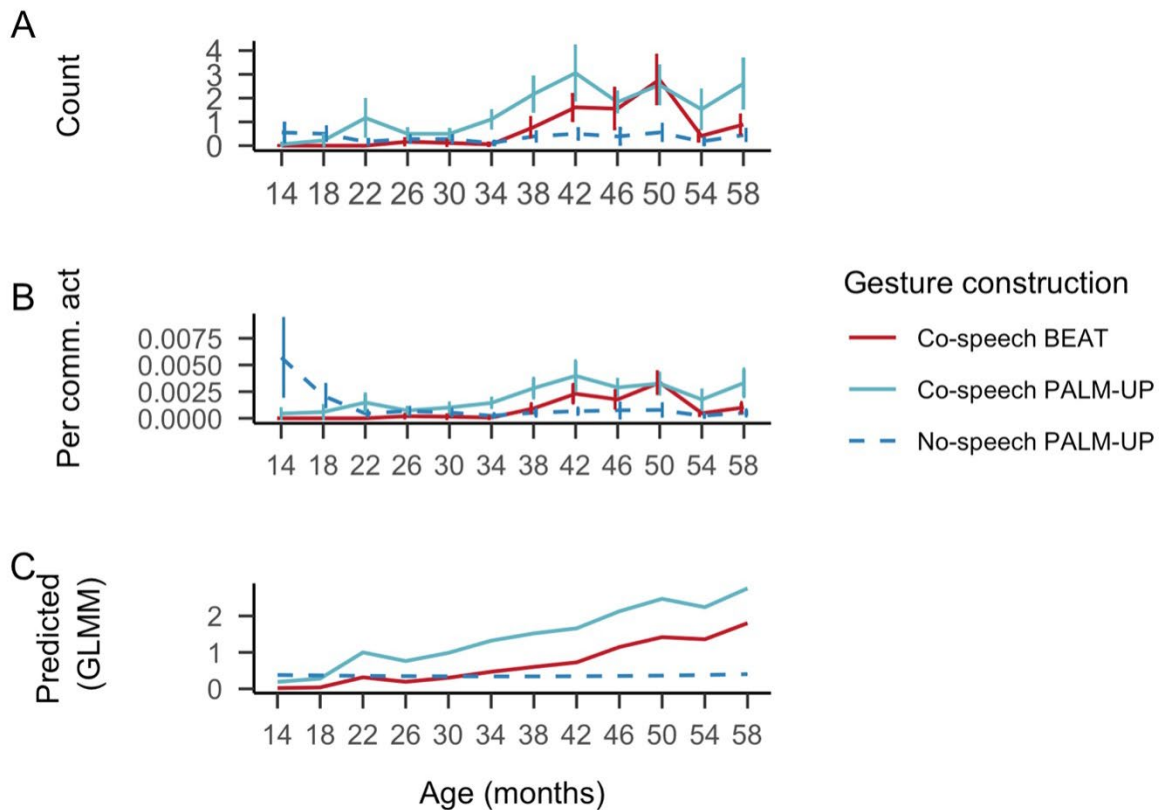
Figure 3.2: Actual and predicted frequencies of gesture across time. Children produce more co-speech beats and palm-ups, but not no-speech palm-ups, between 14 and 58 months. These trends are similar for raw frequency of observed gestures (A), gestures as a proportion of communicative acts (B), and frequencies predicted by GLMM (C).

Generally, children increased production of both co-speech gesture constructions over time and with increasing number of communicative acts but did not increase production of no-

speech palm-ups (Figure 3.2-A). These trends were similar for observed gestures as a proportion

of each child's total number of communicative acts at each session (Figure 3.2-B) and for

predicted frequencies using the GLMM regression (Figure 3.2-C). There were significant

interactions of co-speech beats with age ($\beta = 0.07$, $SE = 0.01$, $p < .001$) and number of

communicative acts ($\beta = 0.96$, $SE = 0.12$, $p < .001$). Similarly, there were significant

interactions of co-speech palm-ups with age ($\beta = 0.04$, $SE = 0.01$, $p < .001$) and number of

communicative acts ($\beta = 0.68$, $SE = 0.08$, $p < .001$). Conversely, production of no-speech palm-

ups did not vary across age or number communicative acts.

Critically, the increases in production across time for both co-speech constructions existed

above and beyond overall increases in amount of talk. That is, children do produce more co-

speech gesture as they produce more verbal utterances, but this predictable increase does not

fully explain the observed upward trajectories of co-speech gesture frequency across early

childhood (see Appendix Figure 2 and Appendix Table 5 for model comparison).

3.3.3 Gesture Construction and Pragmatic Function

The final research question asks whether differences in gesture construction are

associated with particular pragmatic functions in children's early communication. In a first

analysis, I compare overall relationships between constructions and functions. In a second

analysis, I use a multinomial logistic regression to model how these relationships change across

development.

3.3.3.1 *Overall construction-function associations*

I first explored relationships between types of gesture act construction and the pragmatic

functions for all child-produced beats and palm-ups in the data. A Chi-square Test of

Independence revealed a significant relationship between a communicative act's pragmatic

function and gesture construction ($X^2 = 168.91$, $p < .001$). The majority of co-speech beats were produced with unbiased assertions (68%) and the majority of no-speech palm-ups communicated epistemic uncertainty (78.21%). Co-speech palm-ups accompanied acts with a range of pragmatic functions and had no clear primary function. Figure 3.3 visualizes these relationships as an alluvial plot, where band widths represent the frequency that gestures within each category of construction serve each pragmatic function.
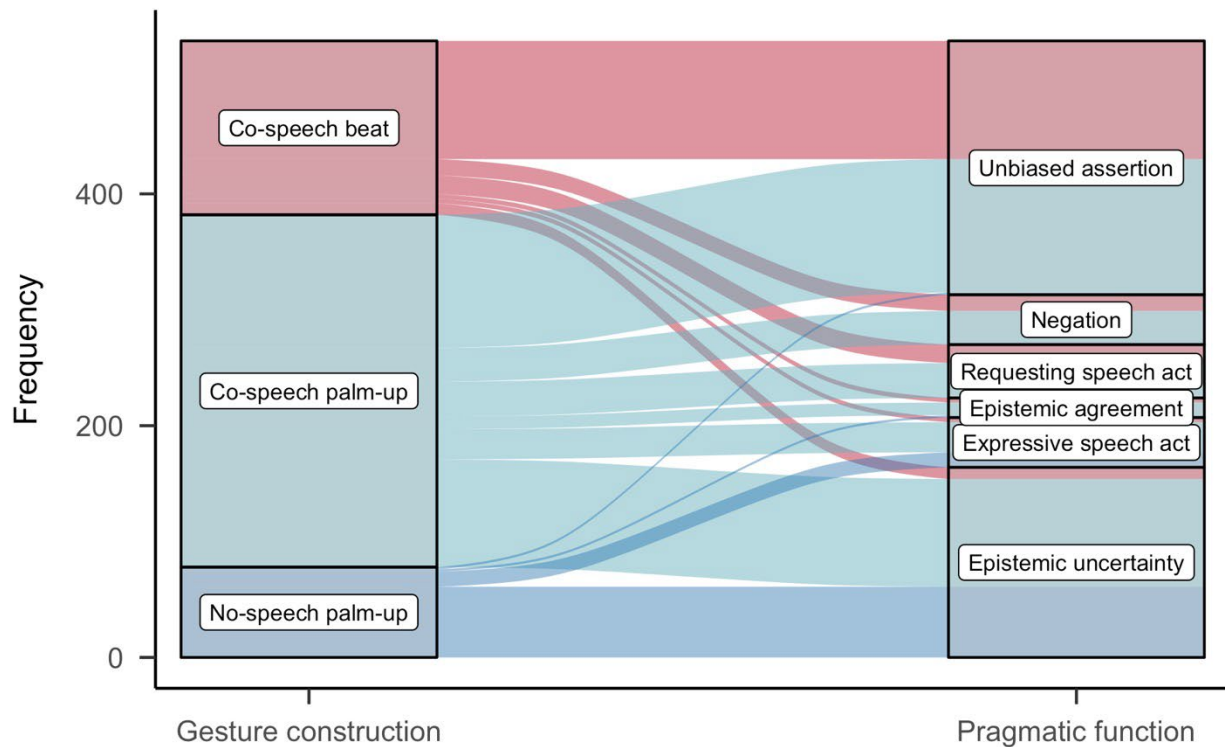
Figure 3.3: Overall associations between gesture constructions and pragmatic functions. Co-speech beats are primarily associated with unbiased assertions. No-speech palm-ups are primarily associated with epistemic ignorance. Co-speech palm-ups are not associated with a primary pragmatic function.

Post-hoc analyses with Bonferroni adjustment confirmed a strong positive association between co-speech beats and unbiased assertions (Std. residual = 7.88, $p < .001$) and a strong positive association between no-speech palm-ups and epistemic uncertainty (Std. residual = 9.81,

p < .001). There was no significant positive or negative association between co-speech palm-ups and any category of pragmatic function (see Appendix Table 6).

3.3.3.2 *Functional change over development*

A follow-up analysis modeled how these overall relationships between gesture construction and pragmatic function changed across early childhood. Figure 3.4 plots how often each of the three gesture constructions was produced with the six pragmatic functions, where frequency is summed across all children at each of the 12 visits. The overall associations between co-speech beats and unbiased assertions and between no-speech palm-ups and epistemic uncertainty appear to hold across time.
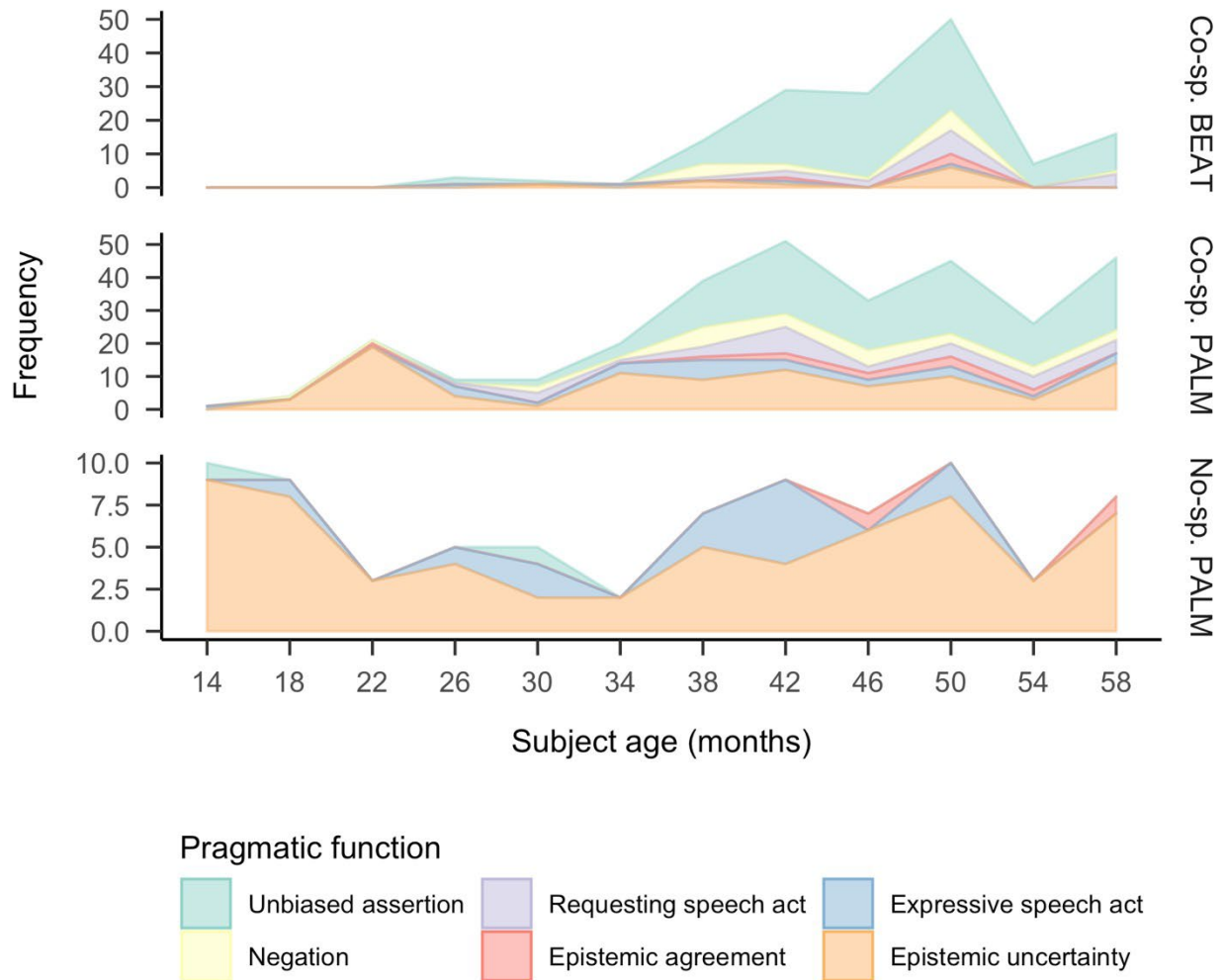
Figure 3.4: Pragmatic function by gesture construction across time. Overall associations between co-speech beats and unbiased assertions and between no-speech palm-ups and epistemic uncertainty hold across time.

I used a multinomial logistic regression to model the likelihood of a communicative act serving each pragmatic function predicted by the type of gesture construction (factor) and the interaction between gesture construction and age in months (numeric, recentered; models included in Appendix D). The predicted trends are modeled in Figure 3.5 where the y-axis represents the probability that a given gesture construction serves each pragmatic function across

the observation period.
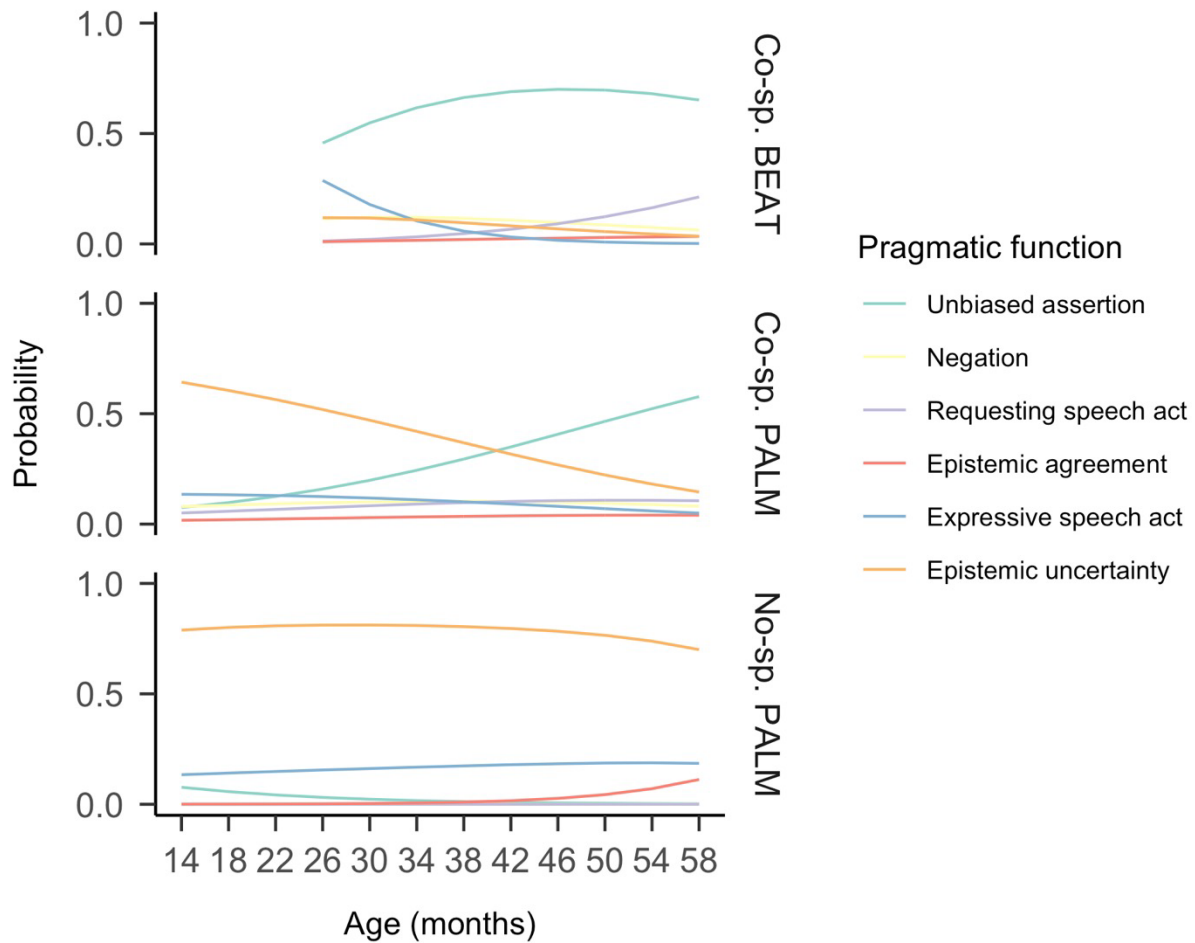
Figure 3.5: Predicted likelihood of pragmatic function for shrugs across early childhood. Curves represent the likelihood a communicative act will serve the pragmatic function based on the act's gesture construction across EC, predicted by multinomial logistic regression. As children develop, their co-speech palm-ups are less associated with epistemic uncertainty and more associated with unbiased assertions. Most associations between gesture construction and pragmatic function are consistent over time.

To test the significance of these trends, I performed a logistic regression for each pragmatic function, predicted by type of gesture construction, child's age in months, and their interaction. Most gesture construction/pragmatic function relationships did not have statistically significant change over time, but there were several exceptions, consistent with the multinomial model.

Between 14 and 58 months, co-speech beats were decreasingly associated with expressive acts (β = -0.16, SE = 0.06, p = .014). Co-speech palm-ups were increasingly associated with unbiased assertions (β = 0.06, SE = 0.01, p < .001) and decreasingly associated with epistemic uncertainty (β = -0.06, SE = 0.01, p < .001).

3.4 Discussion

The analyses in this chapter show that two simple dimensions of gesture use relate to differences in how young children use pragmatic gestures, with consequences for how we understand gesture's role in pragmatic development. I identified a group of gestures produced frequently by infants and young children which are inconsistently defined and relatively understudied. I focused on two dimensions of gesture, form and presence of co-produced speech, which may be clearly recognized and differentiated among gestures in this group. The three analyses in this study illustrate how these dimensions of pragmatic gesture feature in early language development.

First, the form of the gesture construction, but not presence of speech, was related to the age children first produced these gestures. Children began using palm-up gestures about a year before beat gestures. Given that beat gestures must be produced with speech while palm-up gestures may be produced alone, one intuitive explanation for this is that because beats are necessarily co-speech there is simply no opportunity for co-speech gestures before children have sufficiently advanced verbal abilities. However, the fact that children produced co-speech palm-ups as early as they first produced no-speech palm-ups means children do have the opportunity and capability for co-speech gesture acts. The difference in productive onset is driven by form, not gesture-speech relation.

If the difference in onset of beats and palm-ups is not explained by whether or not the forms can be produced without speech, what else might be driving this effect? Perhaps it is due to specific physical features of these forms. If the palm-up form was less complex than beats or required less manual dexterity, infants could produce them earlier in motor development. In reality, however, just the opposite is true. In this system of gesture annotation, the identifying features of a palm-up are significantly more physically involved than a beat. A palm-up is identified by a specific handshape and specific movement, where the fingers are extended away from the palm and there is a visible outward rotation of the wrist. The only identifying characteristic of a beat gesture (in this annotation scheme) is a pronounced and punctuated movement. A beat may take any handshape and the movement may be in any direction. The variation in onset then is not sufficiently explained by dexterity limitations.

One important difference between palm-ups and beats is that palm-ups often serve lexical or emblematic functions while beats generally do not carry semantic meaning. The palm-ups children produced before the onset of beats did not exclusively function emblematically, with co-speech palm-ups in particular serving a range of pragmatic functions soon after onset. Still, these earliest co-speech palm-ups were much more likely to communicate epistemic uncertainty than co-speech palm-ups in the later sessions. It may be that the mere possibility for palm-ups to take topical meanings encourages children to add them more easily to their gesture lexicon, allowing pragmatic functions to grow from early emblematic use. Beats, on the other hand, never function emblematically and are always produced in complement to meaning in speech. It may be more difficult for children to add this exclusively pragmatic gesture to their repertoire from scratch.

Though the *substitution* gesture-speech relation, determined solely by presence or absence of speech, was not associated with productive onset, other relationships between co-speech

gestures and speech might explain the observed differences. Beats are, definitionally, rhythmic with a reinforcing relationship to speech. They are tied to prosody and add emphasis but not meaning. Co-speech palm-ups are often similar to beats in these ways, linked to prosody and emphasizing linguistic meaning, but frequently have a supplementing relationship to speech rather than a reinforcing relationship. Gesture and prosody work as "sister systems" in communicative development (Hübscher & Prieto, 2019) and children's gesture-speech "mismatches" (i.e., gesture supplementing speech with additional meaning) can indicate transitions between stages of language development (Butcher & Goldin-Meadow, 2000; Özçalışkan & Goldin-Meadow, 2005). Given the differences in gesture-speech relation expected between co-speech beats and co-speech palm-ups, a closer look at these relationships might address why these forms differ in onset in ways unexplained by the characteristics of the physical forms themselves.

In general, the results of the first analysis indicate that differentiating along the dimension of form is necessary to avoid misinterpreting observations about functional use of pragmatic gestures. Without this categorical division, the onset for a broad category of "pragmatic gestures" or "non-referential gestures" would be considerably earlier than when children actually begin producing rhythmic, reinforcing beat gestures. Separating along the form dimension allows us to ask what properties of beats lead to them entering children's gesture lexicons later in pragmatic development.

The second analysis in this study compared how frequencies of gesture production changed across early development. Differences in frequency trajectories differed by presence or absence of co-produced speech. Use of co-speech gestures, both palm-ups and beats, increased after onset. Use of no-speech palm-ups decreased across the observation period, with the steepest

decline appearing at the earliest visits, between 14 and 22 months of age. Although it may seem intuitive that co-speech gestures increase across the span of early development marked by dramatic growth in vocabulary and syntax, increases in co-speech gesture use are not fully explained by increases in overall quantity of speech. Co-speech beats and co-speech palm-ups were produced at increasing rates even as a proportion of overall amount of talk. Although it is true that children have more opportunities to produce co-speech gestures as they produce more speech acts, children in this study increasingly integrated beat and palm-up gestures with speech acts above and beyond what would be expected from increased opportunity.

While it may come as no surprise that no-speech gestures did not increase at rates comparable to co-speech gestures, there is no reason to presume they should not increase at all. The rapid increase in the number of speech acts young children produce in spontaneous interaction means an increase in the higher-order category of communicative acts as well. Conversation is not a zero-sum game between verbal and non-verbal turns. As children contribute more to a dialogue, they have every opportunity to contribute more gesture-only communicative acts, but this was not the case. Not only did children not increase their production of no-speech pragmatic gestures, no-speech gestures decreased in use as a proportion of overall number of turns. It seems that the frequency with which children produce these gesture constructions is tied to developing pragmatic skills *in speech*.

Dividing along the dimension of gesture-speech relation brings this stark contrast between co-speech and no-speech pragmatic gestures to light. In this early stage of pragmatic development, children are changing how they integrate verbal and nonverbal modalities, not simply how much they communicate in each. Without the distinction between speech presence and absence, co-speech and no-speech gestures are at odds with one another in an analysis of

frequency of use, obscuring or mitigating the relationship between verbal and non-verbal development. With this distinction, we see that children's use of pragmatic gesture is not simply hitching a ride with speech. Instead, children's use of pragmatic gesture is intertwined with – but not perfectly parallel to – pragmatic changes in the verbal modality.

The final analyses of this chapter showed that both form and substitutive gesture-speech relation relate to the pragmatic functions of communicative acts in early childhood and to how those functions change over time. Co-speech beats and no-speech palm-ups each tended to serve one primary pragmatic function at onset, unbiased assertion and epistemic uncertainty respectively, and maintained a strong bias toward these functions across early childhood.

Co-speech palm-ups were not associated with a single primary function. Instead, they seemed to be influenced by both form and gesture-speech relation. Co-speech and no-speech palm-ups "flocked together" in the first half of the observation period. These two variations on the palm-up form shared a productive onset and both tended to perform epistemic uncertainty. In the later visits, the function of co-speech palm-ups shifted to more closely "flock" with co-speech beats. The association with epistemic uncertainty decreased as co-speech palm-ups became more likely to accompany acts of unbiased assertion.

Despite an increasing alignment with beats along the dimension of gesture-speech relation, co-speech palm-ups never took on a primary function. They continued to frequently accompany acts of epistemic uncertainty and were more likely than either co-speech beats or no-speech palm-ups to serve other pragmatic functions.

Adults use palm-up gestures with tremendous flexibility in function (Cooperrider et al., 2018; Debras, 2017; Jehoul et al., 2017). This analysis hints at what properties of these gestures allow for this flexibility and where in language development it begins to emerge. By the final

observation at 58 months, children's co-speech palm-ups were pulled in two directions by these two dimensions of form and gesture-speech relation. If no-speech palm-ups largely operate emblematically as a non-verbal stand-in for "I don't know" and beats exclusively function pragmatically, co-speech palm-ups find multiple meanings somewhere between the two extremes.

Failing to break apart these pragmatic gestures into constructions based on both form and gesture-speech relation distorts their functional trajectories. Because co-speech palm-ups are much more common than no-speech palm-ups, grouping these pragmatic gestures by form alone masks the strong association between no-speech palm-ups and epistemic uncertainty across early childhood and perhaps emblematic roots of palm-ups gestures in language development. Grouping by presence or absence of speech without a division by form fails to account for how commonly palm-up gestures function emblematically. An analysis lumping together co-speech beats and co-speech palm-ups under a broad label of "non-referential gestures" may not identify early relationships between these gestures and complex pragmatic meanings because it includes gestures primarily operating on a semantic rather than pragmatic level.

Pragmatic development is a long and complicated process, where different pragmatic skills are highly dependent on other linguistic, cognitive, and social skills that continue to develop well beyond early childhood. Dividing along the dimensions of gesture form and gesture-speech relation allows us to examine how pragmatic gestures can perform more specified roles in development than just "doing pragmatics." Epistemic expression (like no-speech palm-ups expressing ignorance) is not the same pragmatic skill as emphasizing selective information (like beats adding emphasis to assertions). Nor is it the same as selecting request-making strategies, soliciting attention, negating, or marking affect (functions frequently performed by co-

speech palm-ups but rarely by no-speech palm ups or co-speech beats). Breaking down how different pragmatic gesture constructions serve different pragmatic functions helps us understand how children develop proficiency putting together many separate contextual puzzles into discourse.

This study carves up a gesture space in an oversimplified way. I considered just two possible dimensions in which pragmatic gestures can differ and just two possibilities within each dimension. I do not claim that these basic divisions give a complete model of pragmatic gesture's role in early language development. Instead, these results indicate that even such simplistic divisions provide compelling reasons to avoid over-generalizations of pragmatic gestures in developmental research. If we afford pragmatic gestures the same fine-grained attention we give to topic gestures – like differentiating between complete and reduced palm-ups or allowing for multiple categories of gesture-speech relation for co-speech gestures – we can more fully appreciate how children use gesture and speech together pragmatically in early communicative development.

In the chapters that follow I pursue a more nuanced analysis of both formal and functional features of palm-up gestures, treating them as one component form of the shrug composite gesture. In Chapter 4, I use a conversation analysis-inspired annotation system to explore emblematic and pragmatic functions of shrugs in the same early childhood corpus used in this chapter. I use the distinction between co-speech and no-speech gestures to identify a kernel shrug emblem, mapping a recognizable shrug form to a predictable shrug meaning. In Chapter 5, I extend these analyses to children's shrugs in early adolescence.

4   Emblematic and Pragmatic Uses of Shrugs in Early Childhood

4.1   Introduction and Literature Review

Children begin using communicative gestures between 8 and 12 months of age, frequently before they speak their first words (Bates, 1976, 1979). Early production of co-speech gestures often precedes, predicts, and facilitates early milestones in verbal language development (Iverson & Goldin-Meadow, 2005; Özçalişkan et al., 2014; Özçalışkan & Goldin-Meadow, 2005; Rowe & Goldin-Meadow, 2009). While co-speech gestures serve critical pragmatic functions in everyday conversation, developmental research on gesture has primarily focused on referential *topic* gestures, like deictic pointing and representational iconic gestures, and their role in lexical and syntactic development.

There is less research exploring the role of non-referential[1] *pragmatic* co-speech gestures in communicative development. Children produce rhythmic beat gestures to highlight prosodic prominence and add emphasis in the pre-school years (see Vilà-Giménez & Prieto, 2021 for a systematic review), and gestures like shoulder shrugs and palm-up gestures are some of the earliest gestures children produce (Acredolo & Goodwyn, 1985; Bartz, 2017; Graziano, 2014a; Harris et al., 2017). These same gestures do sophisticated interactional work in everyday conversation between adults, such as metaphorically handling topics, holding the floor, and taking discourse stance (Bavelas & Chovil, 2000; Bavelas et al., 1992; Dimitrova et al., 2016; Ferré, 2012; Kendon, 2004; McNeill, 1992). Pragmatic gestures like shrugs will become critical

---

[1] Here "non-referential" means the gesture lacks an entity referent, not that it does not perform a referring action. See Chapter 1 for further discussion of this distinction.

supports for pragmatic aspects of everyday language use, but of course – like any part of language learning – infants' shrug gestures do not carry the communicative sophistication of adults'. Rather than commenting on interaction, they are frequently used as emblematic stand-ins for phrases like "all gone" and "don't know".

In order to better understand how children learn to blend gesture and language to efficiently coordinate conversation, we must ask how pragmatic gestures emerge in early productive communication and how these gestures develop from lexical to interactive functions. Given the early appearance and dramatic functional transformation of shrug gestures between toddlerhood and adulthood, I take the shrug gesture family as a starting point for better understanding how communicative gestures fit together with other emergent pragmatic skills as part of fundamentally multimodal processes of language development.

4.1.1   Multimodal Epistemic Expression over Development

Research into the development of metacognition has shown that children are not consistently able to accurately assess and communicate knowledge states with words until the school-age years. Infants begin using cognitive state words around their second birthday (Bretherton & Beeghly, 1982) but do not accurately report total ignorance until age 3, and even then tend to overestimate their own knowledge (Rohwer et al., 2012; Ruffman & Olson, 1989; Wimmer et al., 1988). Children verbally overreport their own partial knowledge states until age 6 (Rohwer et al., 2012; Sodian & Wimmer, 1987).

However, these findings only reflect children's ability to communicate total or partial ignorance explicitly and verbally. Looking at nonverbal communication, it is clear that infants' meta-awareness is more advanced than indicated by their words alone. Before age 2, infants use gestures to signal ignorance (Bartz, 2017; Harris et al., 2017) and opt out of answering questions

when they lack sufficient knowledge (Goupil et al., 2016). Similarly, young children communicate accurate self-assessment of partial knowledge or uncertainty through gestures, filled pauses, and prosody before they do so in speech (Hübscher et al., 2019; Kim, Paulus et al., 2016).

By taking a multimodal perspective, these studies give a more comprehensive understanding of early epistemic expression. What is not yet clear is how epistemic co-speech gestures specifically, rather than broadly defined "nonverbal communication," factor into early interaction.

Although children are able to communicate full and partial ignorance states in nonverbal channels in the second year of life, their use of nonverbal epistemic resources like shrugs and pauses does not appear adult-like even in late childhood (Krahmer & Swerts, 2005). This extended trajectory is characteristic of pragmatic development more generally, including the development of epistemic expression in speech. Young children acquire the vocabulary for verbs of thinking and speaking very early but do not initially use them for highly contextualized and interactive meanings, like marking politeness and deference, hedging uncertainty and disagreement, and differentiating fact and opinion (Bassano, 1996; Bretherton & Beeghly, 1982; Shatz et al., 1983). If gestures play an active and systematic role in pragmatic development alongside speech, there may be a similar mapping shift. That is, a gesture form initially taking a singular epistemic meaning may later appear with variations of that form taking variations of that meaning.

4.1.2   Shrug Gestures

In this study, I focus on one pragmatic gesture – the shrug composite – to ask whether we find a simplified epistemic form-meaning mapping in early pragmatic gesture, complementary to

early speech. Shrugs are uniquely suited for this purpose for a few reasons. First, children produce the shrug form very early, often preverbally, and it remains extremely common in adult conversation (Beaupoil-Hourdel & Debras, 2017; Debras, 2017). Second, shrugs are highly variable in form. The shrug family of gestures includes multiple component forms (e.g., palm-up gestures, shoulder raises, head tilts) which may be combined or used in isolation. Third, shrugs serve both emblematic and pragmatic functions in adults' conversation, sometimes simultaneously (Debras, 2017; Morris, 1994/2015). Many meanings are epistemic, but by no means all. Shrugs' early productive onset in children's communication and their flexibility and variation of use allows many plausible mappings of form and meaning to emerge and reshape over development.

Developmental data is key for illuminating how the shrug gesture's many forms and meanings come to interrelate. For example, if interactive and pragmatic functions grow outward from a first "kernel" emblem, this would support the idea that these complicated relationships arise from changes in form-meaning mapping as the capacities for epistemic expression and stance-taking in discourse develop, rather than the convergence of multiple unique gestures onto a shared set of forms.

The extreme variability in both form and meaning in adult speech as well as the high frequency of production make it difficult to tease apart emblematic and pragmatic use in adult conversation alone. If the shrug's ultimate developmental destination is a gesture made up of one or several formal features used for one or several simultaneous pragmatic meanings, what might its developmental point of origin be?

In order to better understand the many-to-many mapping between form and meaning in adult interaction, we can examine what this mapping looks like in the earliest stages of

communicative development. If the shrug's form-meaning mapping grows from a core emblem, children might rely heavily on this conventionalized kernel gesture before acquiring the pragmatic skills necessary to use the gesture interactively.

4.1.2.1 *Shrug forms*

The shrug gesture may be more accurately described as a family of gestures or *shrugging composite* (Givens, 1977) – also referred to as the shrug *complex* (Morris, 1994/2015) or *compound enactment* (Streeck, 2009) – comprising component gesture forms from the head, shoulders, face, and hands. Figure 1.1 is reproduced below, depicting common form features in the shrug family.

Head tilted

Eyebrows raised

Shoulders raised

Lips tensed, mouth pulled to side

Wrists rotated, palms up and open
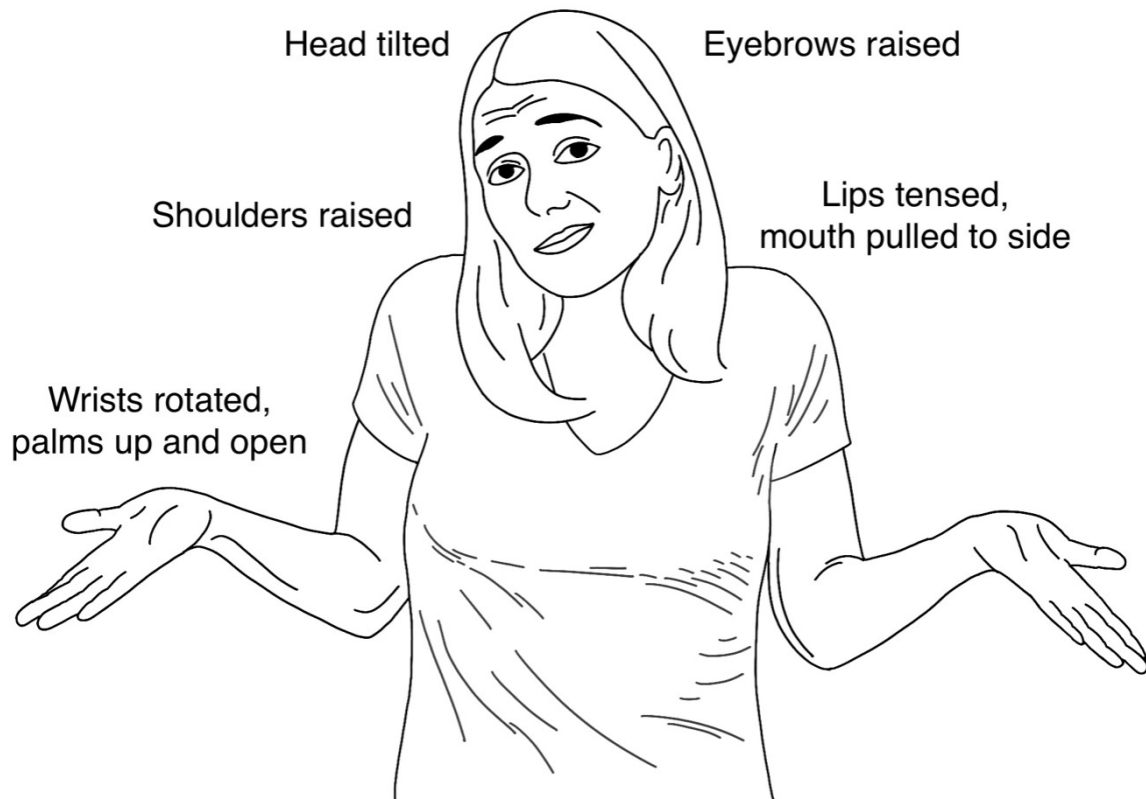
In this chapter I limit analysis to two form components, the shoulder raise and the palm-up. I include the shoulder raise because it is central to the concept of a shrug gesture. It is

arguably the most distinctive feature of the shrug composite, colloquially synonymous with the word "shrug." The palm-up gesture has been a topic of interest as a gesture in its own right (see Cooperrider et al., 2018 for a review). I include the palm-up as a form feature of interest because it is one of the earliest gestures produced by children (British English: Beaupoil-Hourdel & Debras, 2017; American English: Harris et al., 2017; Catalan: Hübscher et al., 2019; Italian: Graziano, 2014b), it is frequently produced with a shoulder raise (Chu et al., 2014; Jehoul et al., 2017), and it conveys the same set of meanings as shoulder raises (Debras, 2017; Ferré, 2012; Jehoul et al., 2017; Müller, 2004).

4.1.2.2   Shrug meanings

Adults use shrugs for a wide array of communicative intents, as discussed in Chapter 1 and outlined in Table 1.1. These meanings do not share an obvious connecting thread and can even be in opposition to one another. In different contexts a shrug may convey certainty or uncertainty, interest or disinterest, agreement or disagreement, leading to an open question of whether there is a singular polysemous shrug gesture or whether there are in fact multiple shrug gestures converging on a set of related forms.

One commonality among many of the shrug's meanings is the intention to express epistemic information. Cooperrider et al. (2018) propose ignorance as a kernel meaning for the lateral palm-up and suggest other meanings grow from this core, even referring to this form as the "palm-up epistemic." Absence of knowledge extends to absence of ability or concern then grows outward to uncertainty, obviousness, hypotheticals, and interrogatives, and finally to exclamatives.

Given the significance of nonverbal signals, including shrug gestures, in children's earliest epistemic expression, we might find evidence for a core epistemic meaning in early

interaction. Moreover, shrugs' flexible many-to-many form-meaning mapping may arise from a core mapping. A kernel meaning may be tied to a kernel form, creating a kernel emblem. Because children's use of shrug gestures to mark ignorance is well established, I pursue "absence of knowledge" as a strong candidate for the shrug's kernel meaning, following Cooperrider and colleagues' suggestion regarding the palm-up form.

4.1.3   Present Study

The present study uses principles of conversation analysis in combination with formal coding of speech and gesture to explore how young children integrate verbal and non-verbal channels interactively (see Stivers, 2015 for more on formal coding in conversation analysis, discussed further in Chapter 2). I use a longitudinal corpus of spontaneous caregiver-child interaction to examine whether the forms and meanings of children's shrugs support the claim for a shrug kernel emblem. If such an emblem exists, we should expect shrugs to have a dominant form-meaning pairing in early communication, particularly when unaccompanied by supporting meaning in speech. Alternatively, if the shrug "family" is really a convergence of multiple gestures, there should not be an association between primary form and primary meaning. We use a developmental lens to further our understanding of both the shrug gesture itself and the multimodal developmental trajectories of epistemic expression and meta-awareness in interaction.

Previous work has established that young children express epistemic state multimodally through speech, gesture, and behavior (Hübscher et al., 2019; Kim et al., 2016). However, it remains unclear how gestures – in particular "uncertainty gestures" like shrugs – contribute to early epistemic communication. We ask two primary research questions to address how shrug gestures fit into a multimodal model of early epistemic expression and add to our understanding

of how these pragmatic skills emerge and function at the earliest developmental stages of discourse-pragmatics.

First, do shrug gestures produced by young children suggest a kernel form or meaning? Although co-speech shrugs can easily take on a wide range of forms and meanings, shrugs without speech carry the full communicative burden of the act. The form must be recognizable and the meaning interpretable using social, physical, and other environmental contextual information outside of the speech context itself. Because children often produce their first shrugs at the pre-lingual stage or shortly thereafter (Beaupoil-Hourdel & Debras, 2017; Harris et al., 2017), early caregiver-child interaction is a rich context for identifying a kernel gesture given the abundant opportunity for children to use both co-speech gestures and gesture-only, "no-speech" communicative acts.

Second, are the forms of shrug gestures in early childhood associated with differences in pragmatic and emblematic function? When a gesture is closer to a conventionalized core form, it should be more easily recognized as an emblem and so more restricted to a corresponding core meaning. If we identify a kernel form and meaning separately and they together make up the kernel shrug emblem, we should expect a strong association between them in early interaction. In particular, following Cooperrider et al. (2018) the *palm-up gesture* may be critical to delineating a kernel form and *ignorance* ('absence of knowledge') may act as a core meaning.

4.2   Methods

These analyses were performed using video and transcript data from the Language Development Project corpus of caregiver-child interaction. Details on participant recruitment, demographics, and selection of the 18-participant subsample are presented in Chapter 2, as well

as methodologies and inter-rater reliability metrics for preliminary transcription and gesture annotation.

4.2.1   Annotation

An overview of annotation from both layers of coding are provided below. Coding methods and reliability measures are discussed comprehensively in Chapter 2. The full coding manual is included in Appendix A.

1. *Palm-up presentation:* the form "completeness" of palm-up gestures, when present. A palm-up was coded as complete when it was produced with a full 180-degree rotation of one or both wrists and a momentary or extended "freeze" in this position. A palm-up was coded as reduced if neither wrist fully rotated or if the palm-up was produced with a single smooth or flicking movement.

2. *Sequential positioning:* the act's sequential positioning in a turn sequence. Acts were determined to be in the first position of a turn sequence if they sought or received a response from the listener and in the second position if they provided a relevant response.

3. *Request sequence type:* the interrogative or imperative form of the first-pair part of any request sequence. Requests were closed if they expected an answer from a limited set of responses, such as yes/no and multiple-choice questions. Requests were open if they expected an answer without a closed set of possible responses, such as *wh*-questions. Finally, requests were imperatives if they issued a command or directive using imperative syntax.

4. *Pragmatic meaning:* the shrug's primary meaning or pragmatic function in the conversation. Options included ignorance, investment, disinvestment, affiliation, disaffiliation, and absence.[2] To distinguish between seeking knowledge and asserting knowledge state, acts that fell in the first position of a closed or open request sequence were assigned the meaning of inquiry in place of the manually coded meaning.

---

[2] A seventh possible category of meaning, directing turn-taking, was not observed in these data. All shrug gestures that met inclusion criteria were coded with one of the seven meaning categories with no option for "other."

4.3   Results

Analyses are limited to child-produced gestures. Although caregivers' non-verbal input undoubtedly influences children's gesture and may help explain the processes by which shrugs emerge and develop, I focus on child-produced gesture to ask whether a shrug kernel emblem is identifiable and how it is employed by young children rather than the mechanisms behind acquisition.

4.3.1   Inclusion Criteria

Sessions were excluded from analyses if the caregiver was not present for most or all of the session, since children had few opportunities to gesture, and the meaning of these gestures was usually ambiguous. Nine of the 216 sessions were excluded due to extremely low caregiver participation, which resulted in a total of 9 gestures excluded (see Appendix B for session exclusion criteria). Individual gestures were excluded from analyses if the transcript lacked sufficient context to determine sequential positioning or meaning, typically due to unintelligible speech, poor audio or visual quality, or the absence of an interlocutor.

4.3.2   Identifying Kernel Meaning and Form

I performed a simple descriptive analysis of all child-produced shrug gestures in the corpus to identify a plausible kernel form and meaning for shrugs. Here I describe proportional frequency of use for shrug meanings and forms, separating gestures co-produced with a spoken utterance from gestures produced without speech, making up a complete non-verbal communicative act. This division between "co-speech" and "no-speech" shrugs allows us to compare gesture production driven by differences in the contextual availability of information. Co-speech gestures add, reinforce, disambiguate, or complement meaning already present in the speech channel. When communicative acts have the potential to simultaneously encode meaning

verbally and nonverbally, the burden of effective communication is shared between both modalities. On the other hand, no-speech gestures must communicate meaning in the nonverbal modality alone. These gestures carry the full communicative burden of the act; there is no speech context to provide redundant or complementary information.

Although adult speakers easily make pragmatic inferences from contexts outside concurrently produced speech, such as complex social relationships, shared knowledge, and verbal information provided much earlier in the discourse, children in the earliest stages of pragmatic development are less adept incorporating these varied contexts into interaction. Infants and toddlers performing entirely nonverbal communicative acts have motivation to produce gestures which are recognizable in form and predictable in meaning, independent of subtle extralinguistic contexts. If pragmatic shrugs grow from a kernel emblem, children's no-speech shrugs are more likely than co-speech shrugs to function emblematically, where a predictable form-meaning pair communicates a direct verbal translation in the absence of speech.

4.3.2.1   Meaning

Shrug meaning was determined by annotation for gloss, sequential positioning, and request type. The meaning 'inquiry' was assigned to communicative acts that initiated or attempted to initiate a turn sequence with either a closed or open request. The meaning of all other acts was determined by the gloss annotation described above, resulting in seven categories of meaning.

These categorical divisions in meaning are primarily motivated by the proposal from Cooperrider et al. (2018) that ignorance or "absence of knowledge" is the kernel meaning for the palm-up lateral form component and the most commonly observed meanings for shrug gestures in language communities across the globe.

4.3.2.1.1 Literal absence

One distinctive category of meaning for child-produced shrugs is literal absence, communicating that an entity is not present or that an activity is not occurring.[3] American English-learning infants and their caregivers often use the palm-up form to emblematically represent the phrases "all gone" or "all done" as well as simple "where?" questions about the immediate context, frequently accompanied by the parallel speech. These early emblems are not consistently included in analyses of shrug and palm-up gesture function. Cooperrider and colleagues (2018) omit these gestures entirely in their theoretical discussion of the palm-up.

One reason for such an exclusion is that there is little evidence suggesting literal absence is a cross-linguistically pervasive meaning. Even within language communities where the absence meaning is observed, it seems to be relatively uncommon in conversation between adult speakers. There is an intuitive assumption that these gestures are more common in infants' communication along with the verbal analogs "all gone" and "all done," but this possibility has not yet been explored empirically.

Before addressing whether epistemic "absence of knowledge" is the shrug's kernel meaning, I examine how children produce shrugs to indicate literal absence or activity completion. These initial analyses allow for the possibility that absence of knowledge grows out of a deeper "absence of entity" kernel meaning. They also offer empirical justification for either

---

[3] Although absence of entity and completion of an activity are not precisely the same concept, they are often used interchangeably by young children and parents. Completion of activities in toddlerhood often goes hand in hand with literal absence. Lunchtime ends when lunch has been eaten, a puzzle is complete when there are no pieces remaining, coloring time is over if there are no pages left in the coloring book, etc.

including these shrugs in the analyses that follow or excluding them as a culturally defined nonverbal emblem distinct from pragmatic shrug gestures.

Ultimately, absence shrugs stand apart from other shrugs in these data in three ways: consistency of form, skewed use by a handful of children, and decreasing frequency of use over time. First, absence shrugs are overwhelmingly produced using the complete palm-up form without a shoulder raise (Figure 4.1). Of the 30 observed absence shrugs, only one used a shoulder raise. All others were produced as a palm-up gesture, and 23 of these were complete palm-ups. Literal absence was the only category of meaning with no observed cases of children combining a shoulder raise and palm-up form.
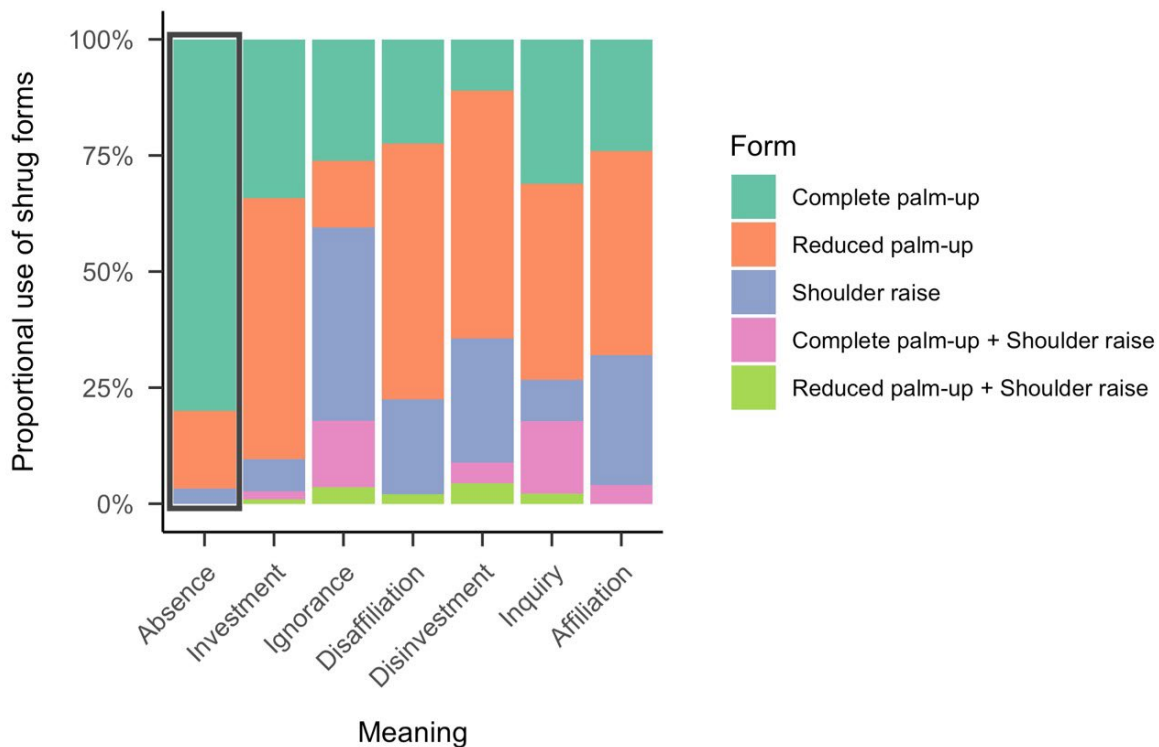
Figure 4.1: Proportion of shrug forms for absence vs other meaning. Absence shrugs primarily take complete palm-up form without a shoulder raise.

Second, although shrugs indicating literal absence were not significantly less common than some other meanings, they were primarily produced by only a handful of children. In fact,

fully half of these shrugs were produced by just two of the 18 subjects. Seven children never

produced a shrug expressing absence or completion (Figure 4.2).
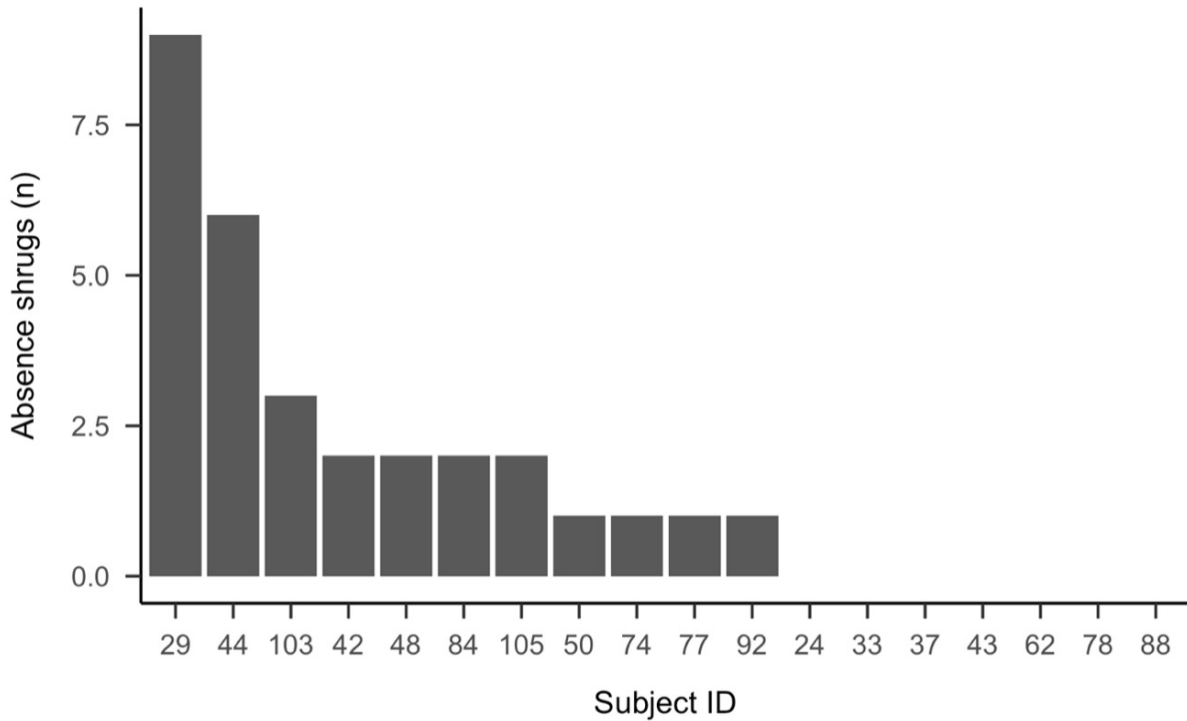
Figure 4.2: Distribution of absence shrugs by subject. Most subjects rarely or never produced absence shrugs.

Third, absence shrugs were unique in their frequency trajectories across the observational

period. While most categories of meaning were observed more frequently as children grew and

produced more language overall, absence shrugs were common in the earliest visits and

decreased in use over time, the last observed at the 46-month session (Figure 4.3).
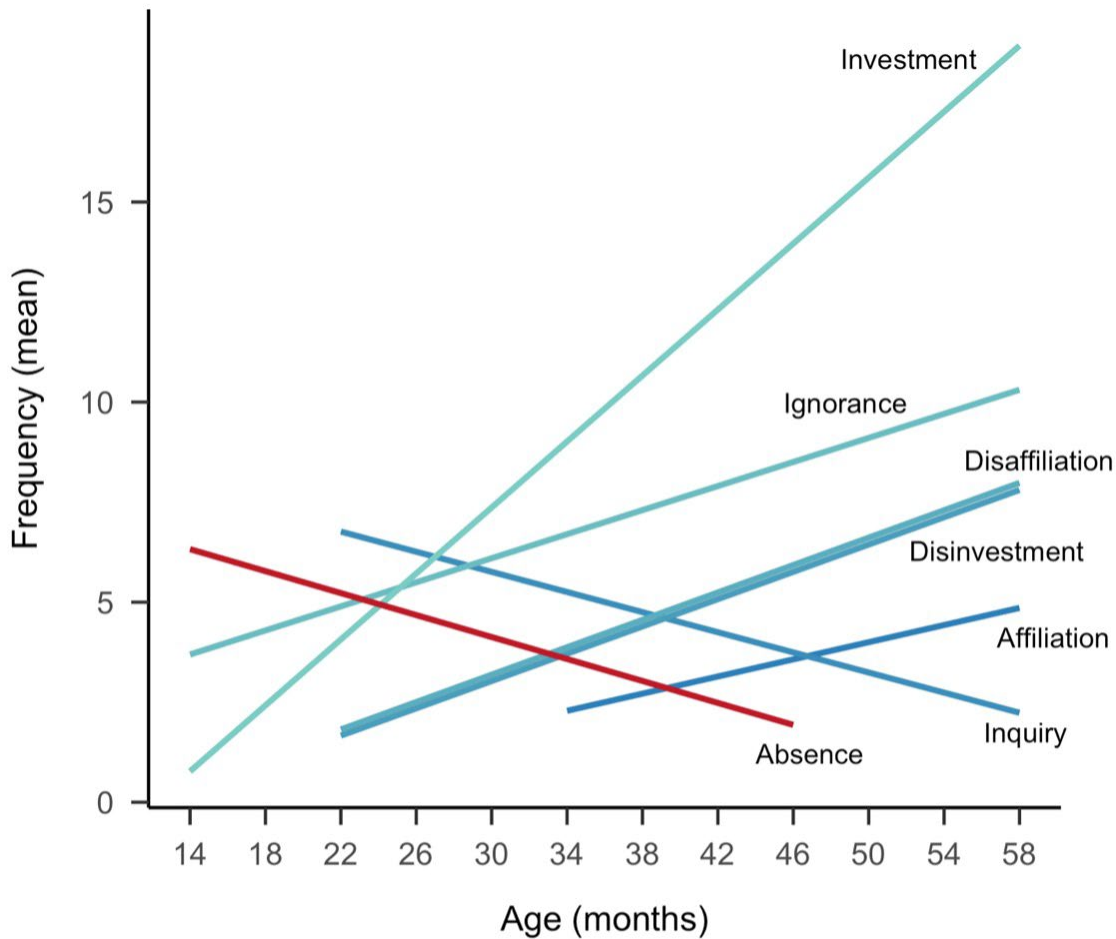
Figure 4.3: Production of absence shrugs over time. Absence shrugs decreased across development and were not produced after 46 months.

Taken together, these observations support the decision to treat young children's "all gone" and "all done" shrugs as a separate phenomenon. This category of meaning displays a distinctive emblematic function, set apart from the diverse range of meanings used by adult speakers cross-linguistically. Although some other categories of meaning did tend to have a primary form, use of the complete palm-up form for absence shrugs was uniquely dominant. The skewed production of shrugs by just a few children was also present for shrugs expressing affiliation, the least commonly produced meaning. Unlike absence shrugs, however, affiliative shrugs did not have a dominant form and emerged both later in development and with increasing frequency. The general decrease in frequency of use over development was also observed in

shrugs expressing inquiry, but inquiry shrugs did not use one primary form, were not predominantly produced by a small number of subjects, were not observed at the earliest visits, and were used even at the final 58-month visit.

4.3.2.1.2 Ignorance

Having ruled out literal absence as a candidate for the kernel meaning of shrugs, I return to pursuing ignorance or "absence of knowledge" as a kernel meaning. A Chi-square Test of Independence revealed a significant relationship between shrug meaning and presence of co-produced speech (X^2(5, $N = 362$)) $= 135.40$, $p < .001$). Ignorance was the only category of meaning more often associated with no-speech shrugs than co-speech shrugs, both proportionally and in raw frequency (Figure 4.4; Appendix Figure 4). Additionally, the vast majority of no-speech shrugs communicated ignorance (72.37%) while ignorance was rarely the meaning of co-speech shrugs (10.14%, Table 4.1). Post-hoc Chi-square analyses with Bonferroni adjustment confirmed this significant negative association between ignorance and presence of speech (Std. residual $= -11.42$, $p < .001$).

Table 4.1: Distribution of shrug meanings by speech presence, early childhood

	Co-speech shrug	No-speech shrug
Ignorance	10.14%	72.37%
Investment	38.11%	6.58%
Disaffiliation	15.03%	7.89%
Disinvestment	13.29%	9.21%
Inquiry	15.73%	0.00%
Affiliation	7.69%	3.95%
Sum	*100.00%*	*100.00%*

Figure 4.4 depicts the relationship between presence of co-produced speech and shrug meaning. Unlike a bar chart depicting proportions for a single dimension, in this mosaic plot both axes are percentage scales. Presence of speech for all shrugs is represented on the x-axis, where shrugs are more frequently produced with speech (79%) than without (21%). Gesture meaning as a proportion of co-speech and no-speech shrugs *separately* is represented on the y-axis. As a whole, the mosaic plot is a visual depiction of the contingency table above (Table 4.1), where the area of each cell represents proportional production of shrugs with the corresponding relationship between co-produced speech and shrug meaning. Overlaid numerals report the raw frequency of use for each box in the mosaic, out of 362 total shrug gestures in the corpus. As a visual reminder that ignorance is the proposed kernel meaning, shrugs communicating ignorance are presented in blue and other meanings in shades of purple.
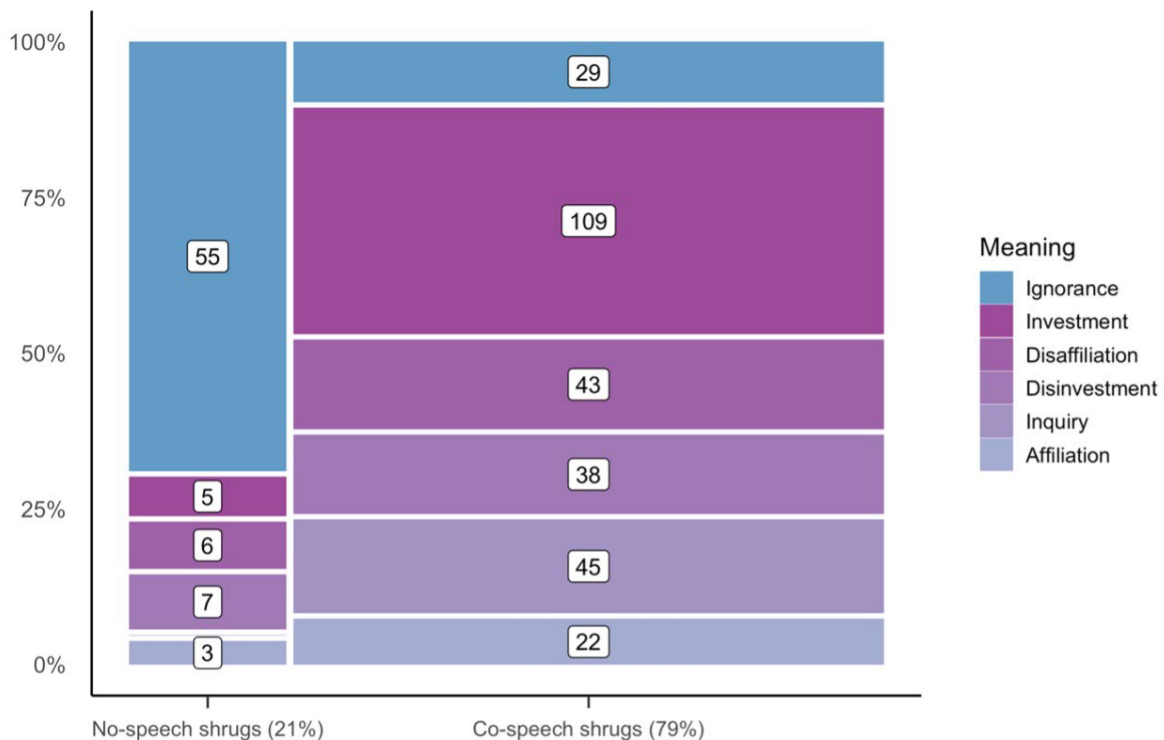
Figure 4.4: Distribution of shrug meanings by speech presence in early childhood. Shrugs are more frequently produced with speech than without. Co-speech shrugs infrequently express ignorance and do

not have one primary meaning. No-speech shrugs express ignorance in a majority of cases. Boxes label raw frequency (N=362).

In sum, when children's shrugs made up a full communicative act without supporting meaning from co-produced speech, there was a strong tendency toward just one predictable meaning: ignorance. When children's shrugs were produced with speech, carrying less of the communicative burden, this tendency disappeared. In line with Cooperrider et al.'s (2018) proposal of "absence of knowledge" as a kernel meaning for the palm-up gesture, these data suggest ignorance may be the kernel meaning of the shrug family of gestures.

4.3.2.2 Form

Identifying a kernel meaning is not enough to establish the existence of a kernel shrug *emblem*. Emblematic gestures are the nonverbal equivalent of a word, where a linguistic or cultural community conventionalizes an otherwise arbitrary mapping between some form and meaning. If the pragmatic shrug gesture arises from an emblematic shrug there must also be a kernel form.

While there is strong theoretical motivation to predict ignorance as a kernel meaning for the shrug, it is less clear what a kernel form might be. The literature offers several hypotheses. As previously discussed, many researchers identify the palm-up form as a gesture in its own right that happens to be frequently co-produced with the distinct shoulder shrug gesture (Cooperrider et al., 2018; Müller, 2004). By this reasoning, the palm-up may be the form with a conventionalized link to the ignorance meaning. Alternatively, if the palm-up is a kinesthetic consequence of shrugging the shoulders (Boutet, 2008, 2018), the shoulder raise component may be the kernel. A third possibility is that children's earliest shrugs operate within the physical constraints of early motor development. This might result in young children tending to produce simpler or smaller movements, like using a reduced palm-up or omitting hand involvement

entirely in favor of a shoulder raise. Yet another possibility is precisely the opposite; children may acquire the emblem by mimicking their parents and so first produce the exaggerated complete palm-up forms that are particularly salient in their caregivers' communication.

Lacking a single strong hypothesis about an emblematic form, I repeat the exploratory analysis performed above to identify a plausible kernel from the same criteria that determined a plausible kernel meaning.

Shrug form was annotated for presence of shoulder raises and both presence and presentation of palm-up gestures resulting in five possible composite forms. A Chi-square Test of Independence showed a significant relationship between shrug form and presence or absence of co-produced speech ($X^2(4, N = 362) = 46.02, p < .001$).

In a simple comparison of forms between co-speech and no-speech shrugs, shrugs that did not include a reduced palm-up component (complete palm-up, complete palm-up with shoulder raise, or shoulder raise alone) were produced proportionally more often when unaccompanied by speech (the parallel effect observed for the ignorance meaning). Conversely, shrugs with a reduced palm-up component, with or without a shoulder raise, were more likely to be produced as co-speech gestures than no-speech gestures (Table 4.2).

Table 4.2: Distribution of shrug forms by speech presence, early childhood

	No-speech shrug	Co-speech shrug
Shoulder raise	43.42%	15.03%
Complete palm-up	28.95%	26.22%
Complete palm-up + Shoulder raise	11.84%	5.24%
Reduced palm-up	14.47%	51.05%
Reduced palm-up + Shoulder raise	1.32%	2.45%
Sum	*100.00%*	*100.00%*

These differences in form across co-speech and no-speech gestures were considerably smaller than the differences observed for meaning. In post-hoc Chi-square analyses with Bonferroni adjustment, only two of the five relationships between form and presence of speech were significant. There was a significant negative relationship between isolated shoulder shrugs and presence of speech (Std. residual = 5.40, p = .399) and a significant positive relationship between isolated reduced palm-ups and presence of speech (Std. residual = -5.72, p > .999).

The mosaic plot in Figure 4.5 represents proportional production of shrugs by both presence or absence of speech and shrug composite form and may be interpreted in the same manner as Figure 4.4 above. To visually highlight the difference between shrug forms positively and negatively associated with the presence of speech, shrug forms without a reduced palm-up (proportionately less often produced with speech than without speech) are presented in shades of red. Those forms with a reduced palm-up (more often produced with speech than without) are presented in shades of yellow. Again, overlaid numerals indicate raw frequency.
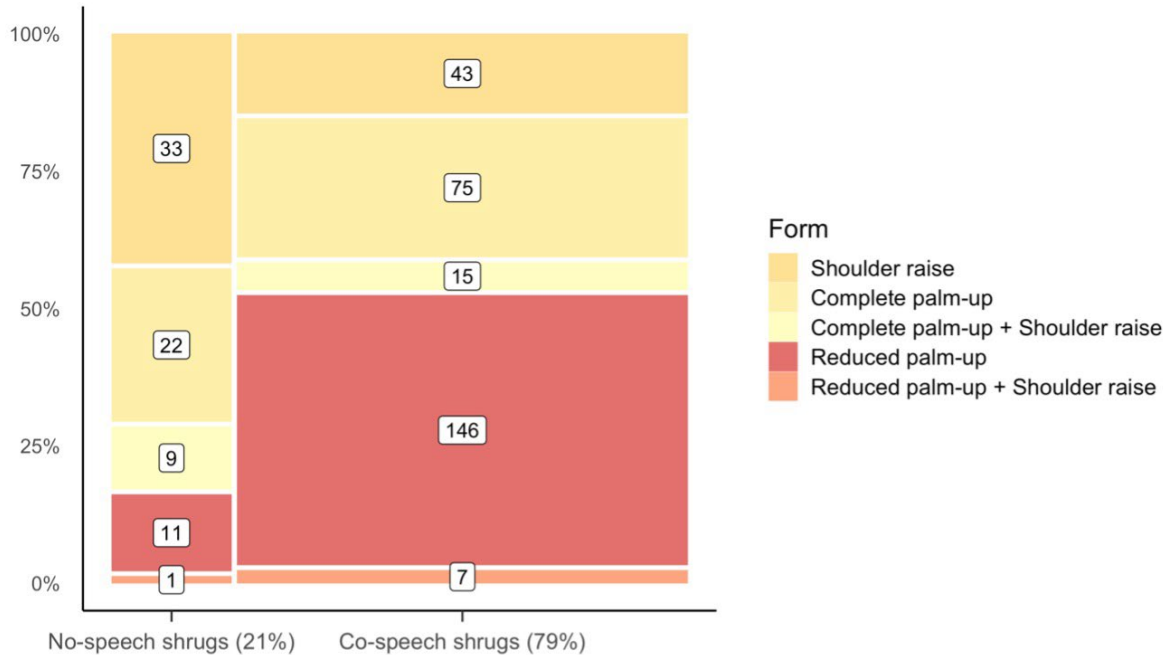
Figure 4.5: Distribution of shrug forms by speech presence in early childhood. Shrugs that include a reduced palm-up component are more often produced with speech than without. Shrugs without a reduced palm-up are more often produced without speech than with speech. These effects are significant for reduced palm-ups (without a shoulder raise) and shoulder raises (without a palm-up). Boxes label raw frequency (N=362).

As demonstrated in the Chi-square test above, there is a relationship between presence or absence of speech and shrug form. Precisely what aspect of form drives this relationship is less clear. Forms with reduced palm-ups tend to appear with speech and forms without reduced palm-ups tend to appear without speech, suggesting kernel form is related to "completeness." On the other hand, given that the only significant associations in post-hoc analysis emerged with isolated shoulder raises and isolated reduced palm-ups, it may be that the presence or absence of a shoulder raise is at the core of the kernel form, easily recognized and salient without speech.

Teasing apart these two possibilities requires additional data. Going forward in this exploratory study, I follow the reasoning supported by the non-significant associations. That is, I define the proposed kernel form as any "unreduced" shrug which does not include a reduced

palm-up component, or those forms that tend to be produced more often in no-speech communicative acts.

4.3.3   Form-Meaning Association

To determine whether the proposed kernel form and meaning together constitute a kernel shrug emblem, I built a mixed-effects logistic regression with gesture meaning as a factorial dependent variable (kernel ignorance meaning/other meaning). I included three predictors of theoretical interest: gesture form (factorial; kernel unreduced form/other reduced form), gesture-speech relation (factorial; co-speech shrug/no-speech shrug), and subject's age in months (numeric). Additionally, I included random effects with random slopes for child's age in months by subject. The inclusion of an interaction between gesture-speech relation and child's age did not result in a model with significantly better fit and so was omitted from the final model. This regression showed a significant positive effect of kernel form on kernel meaning (β = 1.06, SE = 0, p < .001) and significant negative effect of co-produced speech on kernel meaning (β = –2.86, SE = 0, p < .001). There was no significant association between age and shrug meaning across the early childhood observations (14-58 months). An alternative model using meaning to predict form showed similar effects. There was a significant positive effect of kernel meaning on kernel form (β = 1.19, SE = 0.38, p = .002) and significant negative effect of co-produced speech on kernel meaning (β = –1.25, SE = 0.40, p < .001). Full models and tables are included in Appendix D.

Figure 4.6 includes two Venn diagrams depicting the relationship between kernel meaning (ignorance) and kernel form (unreduced) both co-speech and no-speech shrugs. Critically, these figures show that the association between kernel meaning and form is not simply reflecting that both tend to be produced more with no-speech shrugs. For both co-speech and no-

speech shrugs, shrugs with the kernel meaning of ignorance are more often produced with the

kernel unreduced form than with a reduced form.
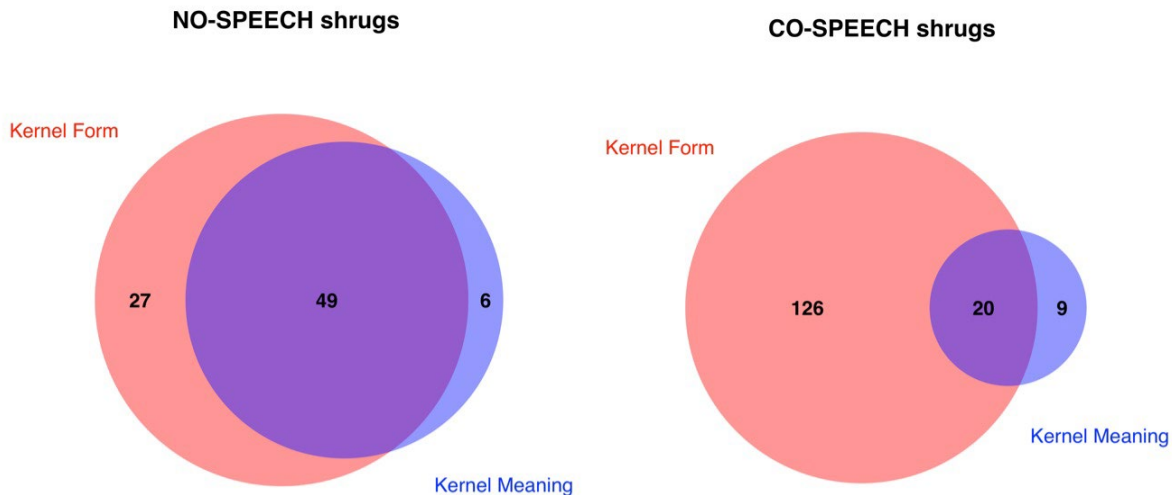
Figure 4.6: Associations between kernel form and meaning in early childhood. Kernel form is associated with kernel meaning in both no- and co-speech contexts. Labels report raw frequency (no-speech N=82; co-speech N=155).

4.4 Discussion

Cooperrider and colleagues (2018) have proposed that the kernel meaning of the palm-up

gesture is 'absence of knowledge' and that other 'absence' meanings extend outward from this

kernel, like absence of certainty or concern. This chapter pursued the authors' suggestion to look

for this theoretical kernel meaning in young children's gesture. After finding descriptive

evidence to support this proposed kernel ignorance meaning for the shrug family of gestures, I

used the same descriptive analysis to propose a kernel "unreduced" shrug form (i.e., the shrug

does not include a reduced palm-up). Kernel form and kernel meaning were associated above

and beyond associations with presence or absence of speech, which may reflect children initially

incorporating a shrug emblem into their gesture lexicon before using shrugs with the pragmatic

flexibility of adults.

To identify a plausible kernel form and meaning, I compared children's shrugs produced with and without speech. Because children have clear lexical and syntactic limitations in the first few years of life, young children's interaction with caregivers is rife with opportunities for communicative gestures both with and without co-produced speech. I hypothesized that early in pragmatic development children would use shrugs differently in these two gesture-speech contexts, since they are less able than older children or adults to incorporate extralinguistic contexts into their communicative acts. If the shrug has emblematic origins, children could be more likely to use this predictable "verbal translation" with a reliably recognizable form when producing shrugs without speech. No-speech shrugs carry the full communicative burden of the conversational turn, while turns with co-speech shrugs can share meaning between the verbal and nonverbal modalities.

Children's no-speech shrugs were significantly associated with the ignorance meaning, while co-speech shrugs more often signaled investment, uncertainty, and other non-ignorance meanings. These findings not only support Cooperrider and colleagues' (2018) proposal of 'absence of knowledge' as the kernel meaning for the palm-up gesture, they further suggest ignorance is the kernel meaning for the shrug composite gesture.

Using the same reasoning, these results also suggest a kernel form. Children's no-speech shrugs were infrequently produced with a reduced palm-up. Instead, when the shrug gesture carried the full communicative burden of the conversational turn, it was produced with a shoulder raise and/or a crisp, visually salient palm-up gesture.

Separately identifying the form and meaning of shrugs children tend to produce without speech does not mean these shrugs are functioning emblematically. A shrug emblem requires a conventionalized mapping between form and meaning, which does not necessarily follow from

the observation that unreduced form and ignorance meaning independently tend to appear in children's no-speech shrugs.

A closer look at the relationship between unreduced shrug form and ignorance meaning reveals the form-meaning mapping we would expect for a shrug emblem. A mixed-effect logistic regression showed a significant association between the unreduced form and ignorance meaning controlling for speech presence. The kernel meaning of ignorance was predicted not only by absence of speech but also by the proposed kernel form without a reduced palm-up. Likewise, unreduced form was predicted by both absence of speech and ignorance meaning.

Identifying an emblematic origin for the shrug allows us to look more closely at how adults' pragmatic use of shrugs develops, with flexible mapping between many forms and many meanings. There are several clear differences between how children in this study produced shrugs and how the literature has described adults' shrugs. Children's shrugs tended to communicate ignorance, but adults' shrugs more frequently communicate obviousness or certainty (Jehoul et al., 2017). Müller (2004) has suggested the core meaning of palm-ups is metaphorical handling of information (the exposed palm requests and presents information in speech) but information-handling meanings (e.g., inquiry, affiliation) were less common than epistemic meanings (e.g., ignorance, uncertainty) in these data. Palm-ups are often used by adults to hold, open, and pass the floor (Bavelas et al., 1992), but children in this sample never used a shrug gesture for turn-management.

It is not typical of emblems to see major shifts in form-meaning mapping over development. A toddler uses a thumbs-up to mean 'good' and a finger pressed to the lips to mean 'quiet'; an adult does just the same. What explains this developmental transformation from emblematic to interactive functions?

103

One possibility has to do with children's ability to produce variations of the shrug form. If patterns of gesture use were merely an artifact of motor development, we would expect an early abundance of reduced palm-up forms with later production of more dexterously demanding complete palm-ups with full wrist rotation and frozen motion at the gesture's peak. Instead children produced complete (N=145) and reduced palm-ups at comparable rates (N=170).[4] This is especially clear in shrugs that communicate literal absence or completion. Absence shrugs were most common at the earliest sessions and decreased steadily before disappearing entirely after 46 months, but 80% were produced with a complete palm-up form. An increasing rate of reduced palm-ups might reflect a tendency for children to produce generally "sloppier" gestures as they become more comfortable using co-speech gestures, but this pattern was not observed with pointing gestures in other research with this same corpus and cohort (Hundertmark, 2016).

Another possibility is that the shrug gesture is functionally changing as part of pragmatic development. The shrug's kernel meaning of ignorance and the high frequency of other epistemic meanings position shrugs as a non-verbal analog to modals and verbs of cognition. Children initially use verbs of thinking and speaking to assert their perceptions of the world, later use them to express an assessment of their own thoughts and knowledge, and eventually use these verbs to serve a variety of discourse functions with only loose or metaphoric ties to knowledge or acts of cognition, like marking politeness (e.g., Abbeduto & Rosenberg, 1985; Harris et al., 2017; Hickmann et al., 1993; Shatz et al., 1983).

---

[4] Because the question of whether children are able to physically produce a gesture form is unrelated to gesture meaning, these frequencies include shrugs meaning absence or completion.

Shrug gestures develop similarly to epistemic linguistic and paralinguistic features. Like modals and verbs of cognition, form and meaning are initially bound together and gain flexibility across early communicative development. Like in speech, children's earliest epistemic expression with shrugs is typically outright ignorance. However, like other non-verbal resources such as prosody, young children also use shrugs as uncertainty markers to indicate partial knowledge. Children's tendency to map a kernel epistemic meaning onto a kernel form, regardless of modality, may be rooted in communicative intent rather than having the form in their verbal or gestural lexicon.

For young children, like adults, the shrug gesture communicates information comparable to and compatible with speech. Speakers can construct their turn with speech only or gesture only, can reinforce speech with a gesture matching spoken meaning, or can use meaning in gesture to supplement meaning in speech. Children's tendency to use no-speech shrugs emblematically suggests that at some level they treat the gesture as a lexical item interchangeable with its verbal gloss ('I don't know'). At the same time, children's co-speech shrugs frequently served interactive or pragmatic functions. In combination with speech, they conveyed meanings without a direct lexical gloss, like non-ignorance cognitive states, affect, and (dis)affiliation.

This study focused on the distinction between no-speech and co-speech gestures, those that substitute for speech and those that are somehow integrated with meaning in speech. These analyses cannot speak to more specified relationships between speech and co-speech gesture (e.g., reinforcing, supplementing), but the difference in usage between no-speech shrugs and co-speech shrugs indicates that children are sensitive to at least some aspects of gesture-speech context in constructing conversational turns.

Describing the developmental use of shrugs not only offers insight into multimodal pragmatic development, it also encourages a multi-functional perspective on the shrug gesture more broadly. Critically, although there is evidence for a kernel form-meaning mapping, I do not suggest that shrugs should be universally classified as emblems. Quite the opposite, the analyses in this chapter show even very young children use shrugs for many of the same kinds of pragmatic functions that adults do. I argue that children's variable use of form and meaning across speech contexts demonstrates it is inadequate to bin shrugs together with other emblems, with beats, with metaphoric conduit gestures, or with interactive gestures. Shrugs are a powerful communicative tool for both children and adults, but we cannot fully understand how shrugs operate so fluidly and flexibly in everyday conversation without recognizing the multiple functions they perform and how these functions change over development.

This study faces several important limitations. While these results suggest interesting parallels between the development of verbal and non-verbal epistemic expression, this analysis is limited to gesture alone. We can only speculate on the broader relationship between speech and gesture in pragmatic development. Future research should ask whether a relationship between the developmental trajectories of both modalities is present at the level of the individual.

Additionally, this chapter describes only children's use of gesture in parent-child interaction. Focusing on child-produced shrugs helps us understand the earliest functions of a complicated communicative resource and how children are mapping form to meaning as they construct their turns. Questions about mechanisms of acquisition or what drives developmental change cannot be addressed without, at a minimum, analyzing how parents gesture with their children.

Finally, the data in this chapter are exclusively from interaction in the early childhood stage. Although young children use shrug gestures very frequently (second only to pointing gestures in these data) these patterns of use are far from adult-like. Like many aspects of pragmatic development, we should expect the formal and functional use of pragmatic gesture to continue developing into adolescence. While age was not significantly associated with form-meaning mapping in the model, this only describes patterns of use in the earliest years of life. Age will unquestionably be a factor across a longer developmental period.

In the following chapter I explore how children use shrugs in early adolescence, at a time when they make great strides in discourse-pragmatic skills essential to sophisticated conversational interaction. I ask how adolescents shrug in conversation, particularly with emerging interactive functions, and I examine longitudinal changes in form and meaning between the early period described in this chapter and early adolescence.

5 Emblematic and Pragmatic Uses of Shrugs in Early Adolescence

5.1 Introduction and Literature Review

The descriptive analyses in the previous chapter showed that young children use shrugs emblematically when they have limited or no contextual support in speech. In these early years children have not yet developed sophisticated strategies for incorporating extralinguistic information into their conversational actions. If the functions of pragmatic gestures in early childhood are dictated by pragmatic competence more broadly, how might these gestures function in the school-age years and beyond? Pragmatic development is a notably long process, and many of the verbal skills that parallel shrugs' pragmatic functions, like stance-taking and negotiating turn-taking, continue to see dramatic developmental changes in late childhood and adolescence.

In this chapter I revisit the same cohort of children later in development. I look for similar evidence of emblematic use of shrugs with limited speech contexts and build on these results by exploring (1) a notable difference in shrug form and (2) use of shrugs in specific discourse contexts.

5.1.1 Later Pragmatic Development in School and Peer Contexts

Pragmatics has a characteristically long developmental trajectory compared to phonological, semantic, and syntactic development. Children become master word-learners and can employ complex syntactic constructions upon school entry. Kindergartners still have a ways to go building their lexicon and organizing increasingly complex sentence structure, but they bring with them enough foundational knowledge to benefit from explicit classroom instruction in these domains.

In the first years of formal Western schooling[1] children make significant gains in metalinguistic competence as they both gain access to written language and receive explicit instruction for how to reflect on, change, and expand their own oral language (Gombert, 1992; Nippold, 1998; Van Kleeck, 1982). Metalinguistic awareness continues to develop into adolescence and young adulthood, allowing for increasingly more sophisticated uses of irony, sarcasm, and metaphor (Billow, 1975; Capelli et al., 1990; Spector, 1996) and abilities to parse and account for phonological, lexical, syntactic, and pragmatic ambiguity (Brodzinsky, 1977; Cairns et al., 2004; Kessel, 1970; Shultz & Pilon, 1973). During this developmental span older children and adolescents rely heavily on intonational and nonverbal signals to detect and interpret irony and figurative language, especially at stages where they can identify nonliteral language but cannot explain it (Capelli et al., 1990; Dent, 1984; Keenan & Quigley, 1999; Milosky, 1994).

Developing metalinguistic awareness enables many oral language skills[2] to be incorporated into formal schooling alongside emerging literacy skills. Some aspects of pragmatics may be explicitly taught in the classroom, for example narrative structure, but most later pragmatic language learning is implicit. Children acquire context-specific interactional skills as they gain experience with new discourse contexts. This applies not only to the new

---

[1] Because the corpus used in this study comprises 16 typically developing English-speaking children in American school systems, I limit my literature review to contexts of Western schooling only. Formal Western schooling is obviously not a universal part of child development, but other styles of formal and informal schooling are beyond the scope of this work.

[2] In this sense "oral language" contrasts with "written language." The term is somewhat of a misnomer, as "oral" language includes not only the verbal components of spoken languages, but also the linguistic structures of signed languages and the communicative nonverbal features of any language.

interactional contexts in which children actively participate, but also exposure to new contexts between other interlocutors. For example, in a study of school-aged Hungarian children, Hollos (1975) found children were able to use some T-V forms (i.e., "tu-vous" systems of polite address) appropriately but not others based on differences in previous contextual exposure. In role-play activities, children were able to use correct T-V structure when enacting conversations between two familiar, but not unfamiliar, adults. Although children would not yet have participated in interactions between two familiar adults, by age 7 they have had many opportunities to observe such conversations between parents, extended family, family friends, etc. At the same age they have likely spent considerably less time in environments with two unfamiliar adults using reciprocal V forms.

At school entry Western children are immersed in new social situations with new communicative goals and consequences. Throughout childhood and adolescence, children continue to be exposed to new social contexts with peers and adults. In most American school systems children begin elementary school in a single classroom with a single teacher and a single group of peers and in middle school or high school transition to multiple classrooms with multiple teachers and rotating groups of peers. Social expectations from both adults and peers change. Social hierarchies develop and relationships must be defined, face saved, reputations managed. Eventually adolescents are motivated to present themselves as adults in some contexts and children in others, demanding increased attention to register. The culturally defined social roles they occupy guide processes of linguistic socialization as they gain entry into some interactional contexts based on social and cultural identities (Ochs, 1982; Ochs & Schieffelin, 2001; Romaine, 1984). In turn, the conversational skills they acquire though accessible social

interactions impact future opportunities to socialize with and influence high-status peers and adults in positions of power (Corsaro, 1979; Evans, 1987; Hemphill & Siperstein, 1990)

Unlike the home environment children are familiar with before school entry, classroom interaction is overwhelmingly adult-initiated and adult-structured, with limited opportunity for fully cooperative teacher-child conversation. Many strategies children have used in the home for both learning and using language will prove less effective or even counterproductive at school. Classroom interaction largely takes the form of "test questions" (alternatively referred to as "display" or "pseudo-" questions) where the correct answer is already known by the question-asker (MacLure & French, 1981; McTear, 1985). In the classroom, teachers ask questions not to seek new information on the topic but to evaluate students' knowledge. This type of interaction is a new experience for many children when they begin school, while other children may enter the classroom with the advantage of some exposure to test questions from parents (Willes, 1981; Yu et al., 2019). Even children who bring some familiarity with them are unlikely to come from households where adult-child interactions are dominated by such test questions to the degree of teacher-student interaction.

The organization of classroom turn-taking is very different from the interactions most children are familiar with from home and family contexts. Children are expected and permitted to contribute much less to interaction in school than at home, as teachers do up to two-thirds of the talking in a classroom (MacLure & French, 1981). Children lack the "rights" to make academic or interactional mistakes or to interrupt or correct a teacher as they might their parent, but the rules of "right" and "wrong" language use in the classroom are often implicit and socially loaded (Dannequin, 1987; Mehan, 1979).

School success depends on both academic knowledge and interactional skills, and children must integrate the two for successful classroom participation (Mehan, 1979). While most children are able to quickly learn the implicit rules of a classroom, others may fail to do so when they perceive classroom interactions as disorderly or unpredictable (Willes, 1981). Individual variation in the interactional exceptions of home environments likely contributes to students' abilities to adapt to classroom norms (e.g., children who hear test questions at home begin school understanding the intentions behind this style of interaction), but variation in teaching styles can also contribute. Teachers' attempts to solicit a particular kind of response can fail when student and teacher do not collaboratively develop the topic of narrative, with children who do not struggle adapting to classroom social rules (Michaels, 1981). Adult-child interaction norms in schools can mislead teachers' assessment of students' abilities. For example, Barnes and Todd (1977) found that 13-year-olds displayed more advanced conversational skills when participating in small-group academic discussion with classmates than were apparent in teacher-student interaction and exceeded teachers' expectations for students' verbal abilities.

New peer contexts mean the development of new pragmatic competences. School-age children are able to identify the social groups they belong to, largely defined by age and gender, and develop language patterns of use specific to the social organization of those groups. For example, girl groups tend to use inclusive language structures that reinforce intimacy and loyalty while boy groups tend to use grammar that asserts a position of dominance, reflecting a tendency for these groups to be tightly hierarchically organized (Goodwin, 1980, 2006; Maltz & Borker, 1996).

As children move through the school-age years, they both spend an increasing amount of time socializing with peers and have more opportunities to interact with new peer groups.

Conversations with peers do not replace interactions in the home, but instead allow children to be selective about which topics belong to which social groups (Raffaelli & Duckett, 1989). Children shift register between peer and adult interactions and in adolescence add new strategies to their repertoire of resources for register shift. Larson and McKinley (1998) recorded adolescents' use of conversational behaviors like using figurative language, smooth topic shifts, and influencing others' actions and beliefs. They found that in addition an overall increase in many of these behaviors between age 12 and 18, adolescents used these behaviors more in peer interactions than conversations with adults. Furthermore, these behaviors indicated register shifts between groups of peers. Adolescents were not only more formal with adults than with peers but were more formal with peer acquaintances than close friends.

As they are continually immersed in new social contexts with both adults and peers, children and adolescents must pick up on many new communicative norms. Each time they participate in an interaction, they must assess the context to determine which linguistic, cognitive, and social rules they should follow and then attempt to adhere to those rules, reassessing and adapting as needed. Pragmatic competences like irony, cohesion, and givenness that have received much attention in research on later language development are certainly relevant, but in the next sections I highlight some later-developing conversational skills explicitly tied to dialogue and cooperative interaction.

5.1.2   Request Strategies and Register

Children develop basic ways of differentiating register early in life. Toddlers begin using contextualized politeness strategies by age 3 and adapt some of these strategies (particularly vocal and gestural) to account for social roles and burden (e.g., Ervin-Tripp et al., 1990; Hübscher et al., 2019). As children enter and progress through schooling, interacting with new

groups of children and adults, the necessity for attention to polite register grows. Across the pre-school years children develop more complex linguistic expressions for presenting and re-presenting requests. By age 5, children have a repertoire of request forms and rules for employing them, like mitigating burdensome requests with politeness markers (Read & Cherry, 2009).

Politeness strategies for requests continue to develop at least through middle childhood. In a study of children's requests and control acts in the pre-school and school-age years, Bates (1976) found that before age 4 children mainly produced requests with direct questions and commands. Children could produce all the appropriate syntactic forms for polite requests between age 5 and 6 but were not able to strategically modulate request syntax until after age 7.

Axia and Baroni (1985) investigated differences children's request strategies with groups of children at 5, 7, and 9 years old. For all age groups, all initial requests tended to be impolite, but politeness for repeat requests differed by age. The youngest children increased volume following refusals (indicating *some* strategy for achieving their goals) but did not give more polite requests, while 7-year-olds did increase polite constructions and 9-year-olds even more so. Only the oldest children used strategies like mitigation, conditional tense, and negotiation following non-responses or motivated refusals.

Infants and toddlers employ many resources to have requests granted, but persuasion strategies see dramatic development across the school-age years. Younger children typically take an egocentric perspective, where both initial requests and reformulations express their own desires. By late childhood, children employ more allocentric strategies and politer request forms (Axia, 1996). Persuasion continues to develop as an interactional skill through adolescence. Between 11 and 15 years of age, adolescents increasingly adjust imperative syntax based on

114

social status and use more "propositional appeals" like conforming to norms or expectations of the social context (Piché et al., 1978).

Some key structures of narrative production do not emerge in children's language until middle childhood or adolescence. For example, it is not until age 9 that children incorporate past perfect verb forms, nonfinite forms, and multiple markings of aspect into oral narratives. Nine-year-olds' narratives are still far from adult-like though, with less lexical diversity, fewer expressions for time-relations, fewer locative prepositions, less temporal conjunctions, and generally less complex syntactic packaging within story episodes (Berman & Slobin, 1994; Clark, 2009).

Later gains in pragmatic competences are also notable in extended dialogic discourse. Larson and McKinley (1998) conducted a longitudinal study of conversational behaviors of eight "average-achieving" American English-speaking adolescents (4 girls, 4 boys) beginning at age 12 and continuing to age 18, when the students graduated from high school. They found that the frequency of abrupt topic changes decreased significantly over the observational period while the frequency of both positive and negative interruptions increased significantly. Increasing rates of interruptions were not a reflection of rude behavior but rather cooperative strategies for maintaining the interaction, for instance with exclamations of empathy and interest or requests for clarification and repair. Other conversational skills that see dramatic changes in the school-age and adolescent years include participating in longer dialogues with longer durations of sustained topic (Brinton & Fujiki, 1984; Schober-Peterson & Johnson, 1993), increasing rates of relevant and factually based comments (Dorval & Eckerman, 1984), adjustments to mirror or accommodate thoughts and feelings of interlocutors (Raffaelli & Duckett, 1989; Rawlins, 1992), differentiation of topic selection and conversation style by relationship to interlocutor (Larson &

McKinley, 1998; Raffaelli & Duckett, 1989; Rawlins, 1992; Turkstra, 2001), and additional use of conversational supports like backchannel feedback (Nippold, 1998; Turkstra, 2001).

Despite this extended trajectory, research into later pragmatic development is much more limited than research on pragmatic development for infants and toddlers. This is true for all aspects of pragmatics, and it is especially true for multimodal pragmatics.

5.1.3   Gesture in Later Development

There is a large body of work examining the role of gestures, especially topic gestures, in early language development. It is well established that children's early topic gestures like pointing to objects in the environment and iconically representing actions are closely connecting to emerging lexical and syntactic skills in infancy and the pre-school years. There is considerably less work studying gestures in later communicative development.

Some of the most notable work on older children's gesture falls more in the domain of cognitive development than language development. Older children's production of gestures that supplement or contradict meaning in speech (speech-gesture "mismatches") in math and science learning contexts are more indicative of children's knowledge than speech alone and can reflect a "ready-to-learn" state (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Crowder & Newman, 1993; Perry et al., 1988). We know gesture mismatches also indicate transitional states in early language learning (Butcher & Goldin-Meadow, 2000; Goldin-Meadow & Butcher, 2008; Özçalışkan et al., 2013), but we know very little about whether gestures may also indicate transitional or ready-to-learn states in later stages of oral language and literacy development.

Capirci, Cristilli, de Angelis, and Graziano (2011) compared elicited narrative retellings from 4-, 6-, and 8-year-olds and analyzed the use of both linguistic devices and representational

gestures. They showed that both formal and semantic aspects of gesture developed over time. Changes in representational gesture forms illustrated children's increasing ability to use gesture expressively. There was also an increasing trend of semantic accuracy for representational gesture, showing that children must not only acquire new gesture forms, they must learn to use these expressive formal components appropriately. Moreover, they found that the gradual mastering of representational gestures in the school-age years occurred in parallel to gradual mastering of narrative devices in speech. Another cross-sectional study of linguistic and gestural features in narrative retellings from Colletta, Pellenq, and Guidetti (2010) compared rates of both representational and non-representational gestures for 6-year-olds, 10-year-olds, and adults. Rates of all gestures increased significantly across all three ages, but non-representational gestures increased at higher rates than representational gestures. Notably, this category of *non-representational* gestures is not equivalent to *pragmatic* gestures. Non-representational gestures included some kinds of pragmatic gestures (e.g., beats) but other common pragmatic gestures (e.g., palm-ups, shoulder raises, and head gestures) were omitted from the analyses entirely.

These findings show that pragmatic functions of gestures in narrative continue to grow beyond the pre-school years and even hint that pragmatic gestures may play a different role than representational gestures (Alamillo et al., 2013; Colletta et al., 2015; Kunene, 2010). However, all of these studies are centered on gestures in monologic narrative retellings. Gestures, particularly pragmatic gestures, serve very different functions in monologic and dialogic discourse. For example, pragmatic beat gestures are useful for organizing narratives and emphasizing important topical content, but there is no use for interactive gestures that invite listener participation unless there is an actively participating listener.

There is some existing research that hints at how gesture might factor into older children's pragmatic development in dialogic interaction, but this work generally concerns multimodality rather than gesture specifically. For example, Krahmer and Swerts (2005) conducted two experiments on how children (aged 7-8) and adults signal and perceive uncertainty with both audio and visual information. They found that nonlinguistic uncertainty markers like pauses, brow raises, and "funny face" were both produced at high frequencies by adults and perceived as highly informative by adults when assessing another speaker's feeling of knowing. The results for children showed the same general patterns but with weaker effects. Children used the same signals but at relatively low rates. They also interpreted these signals as indicating uncertainty but were worse than adults at accurately assessing feeling of another's knowing. This study demonstrates that school-aged children use nonverbal cues to take and read epistemic stance, but that they have not yet developed adult-like proficiency. A follow-up study from Visser, Krahmer, and Swerts (2011) compared production and perception of these same audiovisual markers of feeling of knowing with 8- and 11-year-olds. As predicted, the older children produced more audiovisual epistemic markers and more accurately rated feeling of another's knowing. Furthermore, both groups expressed certainty and uncertainly differently in collaborative and competitive contexts, but in different ways. Younger children more often expressed certainty in collaborative contexts and uncertainty in competitive contexts. Older children did just the reverse. By age 11 children have a more advanced understanding of display rules for different social contexts (Gnepp & Hess, 1986; Saarni, 1984). At age 11, but not age 8, children apply these rules to their own nonverbal self-presentation, prioritizing certainty signals in competitive contexts where uncertainty may be a vulnerability and prioritizing uncertainty signals in collaborative contexts where uncertainty may communicate an invitation for cooperation.

Taken together, research on older children's gesture in monologic narratives and on their multimodal signals in dialogic contexts suggest that pragmatic gestures act as an important piece of later pragmatic development, but there remains a problematic gap. Pragmatic gestures are a critical part of navigating everyday conversation by taking stance, managing turns, and monitoring listeners' engagement and comprehension, but we know very little about how these skills are mastered in the decade of development between entering pre-school and when they begin to reach adult-like competences.

5.1.4   Present Study

The previous chapter described young children's production of shrug gestures in spontaneous caregiver-child interaction. By comparing shrugs produced with and without speech I identified a kernel shrug form (shoulders raised and/or a complete palm up) and kernel meaning (ignorance). "No-speech" shrugs that carry the full communicative burden of an act were more likely to take this recognizable "unreduced" form and convey a predictable meaning of ignorance. Moreover, kernel form and meaning were associated above and beyond the effects of speech presence or absence, suggesting young children use shrugs emblematically when meaning in speech is limited or unavailable.

The analyses that follow build on these findings by describing emblematic and pragmatic uses of shrug gestures by the same cohort of children in early adolescence. While there are clear commonalities in the array of forms and meanings children use in both the pre-school years (for examples, see Figure 5.1) and later in adolescence (Figure 5.2), there may be important qualitative differences in relationships between form and meaning. The first goal of this chapter is to determine whether the patterns of shrug use observed in the early childhood period persist into early adolescence. That is, are older children still more likely to use an unreduced form and

ignorance meaning when gestures are produced without speech? Is there still an emblematic

mapping between kernel form and meaning?

Figure 5.1: Examples of shrug gestures in early childhood sessions. (A) A 14-month-old boy produces a complete palm-up to communicate literal absence. (B) A 46-month-old girl produces a complete palm-up and shoulder raise to communicate investment/exclamation. (C) A 54-month-old girl produces a complete palm-up to communicate ignorance. (D) A 54-month-old boy produces a reduced palm-up to communicate disinvestment/disinterest.
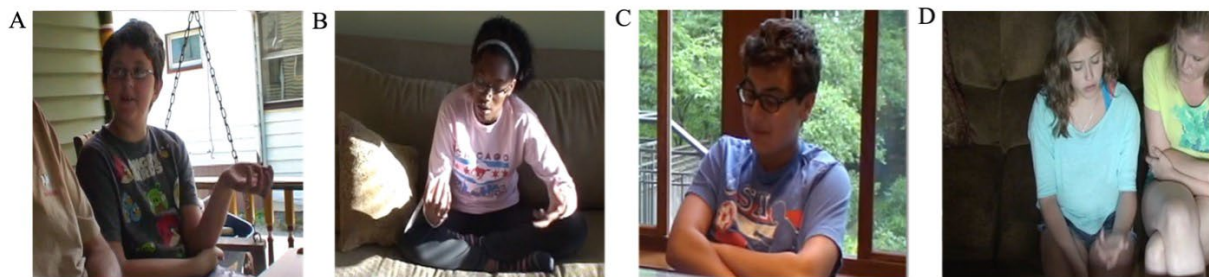
Figure 5.2: Examples of shrug gestures in adolescence sessions. (A) An 11-year-old boy produces a complete palm-up to communicate disaffiliation/disagreement. (B) A 13-year-old girl produces a reduced palm-up (rapid; no frozen peak) to communicate investment/obviousness. (C) A 13-year old boy produces a shoulder raise to communicate affiliation/backchannel acknowledgment. (D) A 12-year-old girl produces a complete palm-up to manage the floor/cede the turn.

Given that older children are much more adept at incorporating situational and discourse

contexts into their contributions to conversation, the second goal of this chapter is to describe

what contextual features adolescents incorporate into their use of shrugs. I consider emergent

properties of both sides of shrugs' form-meaning mapping.

5.2 Methods

5.2.1 Participants

Sixteen of the 18 participants in the early childhood cohort continued to the adolescent observations. Recruitment procedures and subsample selection are provided in Chapter 2. The participants who did not continue included one girl from the high-MLU group (subject 29) and one boy from the median-MLU group (subject 33). Additionally, one boy from the median-MLU group (subject 62) completed only the first of two sessions in this period. Demographic information for individual subjects may be found in Table 2.1.

5.2.2 Data Collection

Children and their families were visited in their homes in the summers following fifth (mean age 11.44 years) and seventh grade (mean age 13.44 years). At each visit, the child and their primary caregiver were presented with a series of 12 tasks designed to elicit higher-order thinking talk (Frausel et al., 2020) in four domains: science, math, social, and art. Dyads were instructed to work together to and were given up to five minutes to complete each activity.

Three tasks from each visit were selected for inclusion in this study, five in total from the social domain and one from the art domain. These tasks were selected for two reasons. First, the tasks did not involve manipulables, leaving participants free to gesture. Second, the prompts successfully elicited high rates of speech from both parents and children. Prompts were open-ended questions about subjective topics which both parents and children typically found engaging. For example, one task asked dyads to compare the morality cheating and stealing. Another asked them to imagine what life would be like without cell phones. Additionally, because the tasks did not have correct answers, conversations tended to prompt more balanced

dialogues between speakers than math and science tasks. More information about the tasks is provided in Appendix B.

5.2.3   Annotation

Methodologies and inter-rater reliability metrics for transcription and annotation are described in Chapter 2. Annotation procedures were identical to those in Chapter 4 with one exception: all communicative acts with head gestures were coded in the adolescence sessions but not the early childhood sessions. In early childhood, children almost never combined head gestures with shoulder raises and palm-ups. The analyses below concerning head gestures without other shrug components would not have been possible with the early childhood data. A complete coding manual is included in Appendix A.

For the first set of analyses concerning kernel form and meaning, categories of meaning and form are unchanged from Chapter 4. Additionally, only gestures that would have been included in the early childhood data based on form were included in the parallel analyses. That is, gestures had to include a shoulder raise or palm-up.

For the analyses that were not performed with the early childhood data (those concerning head gestures and turn design), some variables were simplified:

Imperative request sequences (N=2) are considered non-request sequences. Consequently, non-request sequences are those that lack interrogative syntax or intonation.

Sequential position is simplified to three categories: contingent, non-contingent, and backchannel. Contingent acts are second-pair parts; they directly respond to an initiating act by taking the next turn. Importantly, "contingency" here refers to sequential contingency and not topical contingency. Non-contingent acts take the turn, but do not provide a response. Non-

contingent acts may or may not receive a response. Backchannel acts provide feedback to the speaker without taking the turn.

Shrug meaning is collapsed into four categories: investment, disinvestment, affiliation, and disaffiliation (Table 5.1). This simplification is motivated largely by work from Debras and Cienki (2012), who emphasize the shrug's stance-taking functions and by the stance triangle framework from Du Bois (2007). Theoretical motivation and implications for this approach are discussed further in Chapter 6.

Table 5.1: Collapsed annotation for shrug stance-taking meanings

Meaning	Stance action	Collapsed meanings
Investment	Closing distance between speaker and proposition or topic	Investment
Disinvestment	Widening distance between speaker and proposition or topic	Ignorance, disinvestment
Affiliation	Closing distance between speaker and listener or stance	Affiliation, affirmation, turn negotiation
Disaffiliation	Widening distance between speaker and listener or stance	Disaffiliation, negation

5.3   Results

5.3.1   Kernel Meaning and Form

The descriptive analyses in Chapter 4 explored whether children's earliest shrug gestures suggested a kernel meaning and form for shrugs. The results supported Cooperrider et al.'s (2018) proposal that 'absence of knowledge' is the kernel meaning for the palm-up gesture. The results also suggested this meaning may be a kernel not only for palm-up gestures but for the shrug composite gesture. Ignorance shrugs were associated with an "unreduced" kernel form for shrugs, which included a shoulder raise form and/or a crisp, "complete" palm-up form but did

not include a rapid or minimized "reduced" palm-up form. Between 14 and 58 months, children tended to use this kernel ignorance meaning and kernel unreduced form for shrugs produced without speech. The association between this meaning and form was significant above and beyond the effects of speech presence ("co-speech") or absence ("no-speech") and was not explained by differences in age across this early developmental span. Young children linked this recognizable form and predictable meaning to produce a functional shrug emblem when there was no additional contextual information in speech.

I conducted a parallel set of analyses with the early adolescence corpus to see whether this emblematic association persisted beyond early childhood. I compared the meanings and forms of shrugs produced with and without speech and looked for any association between the proposed kernel ignorance meaning and kernel unreduced form above and beyond effects of speech presence.
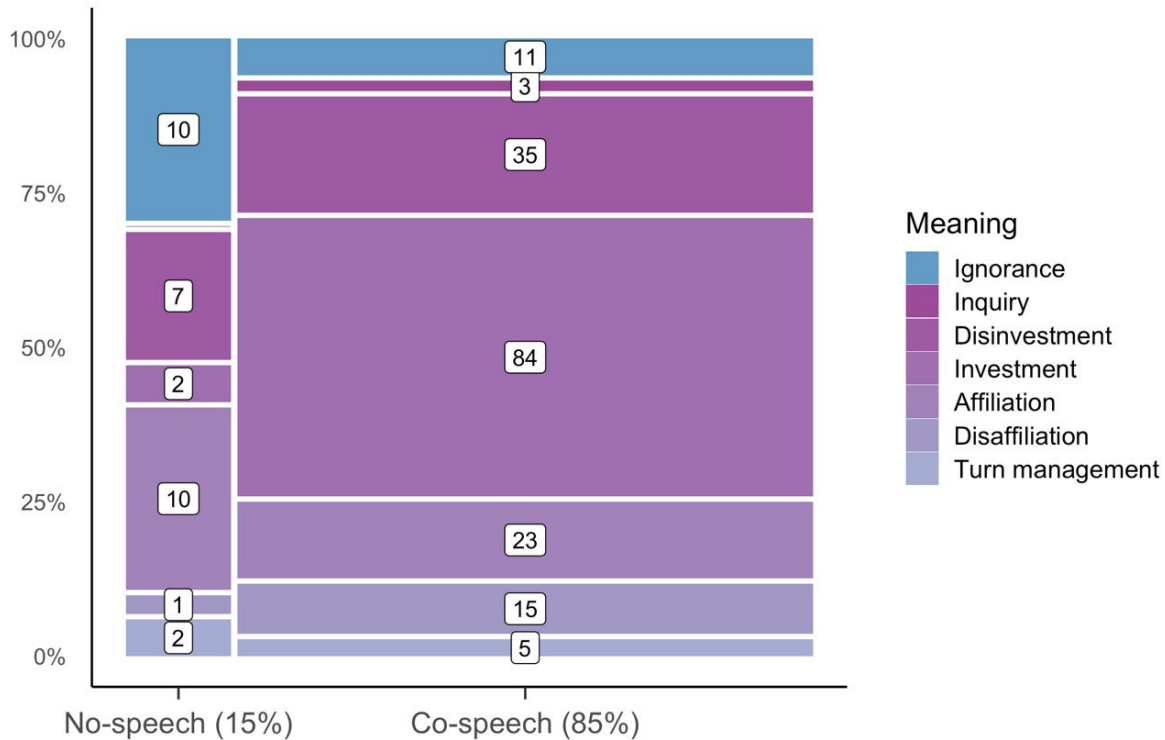
Figure 5.3: Distribution of shrug meanings by speech presence in adolescence. Shrugs are more frequently produced with speech than without. Ignorance is a more common meaning for no-speech shrugs than co-speech shrugs. Boxes label raw frequency (N=208).

The meaning of adolescents' shrugs was related to speech presence (Figure 5.3). A Chi-square Test of Independence revealed a significant relationship between shrug meaning and presence of co-produced speech ($X^2(6, N = 208) = 36.23$, $p < .001$). Post-hoc Chi-square analyses with Bonferroni adjustment confirmed this significant negative association between ignorance and presence of speech (Std. residual = -4.32, $p < .001$). These results are consistent with the early childhood results, though the association between ignorance and absence of speech is considerably weaker in early adolescence (Std. residual for EA = -4.32, for EC = -11.42). Ignorance shrugs were proportionally rarer in adolescence than early childhood, making up only 10% of all shrugs (vs. 23% in early childhood) and 31% of no-speech shrugs (vs. 72% in early childhood).
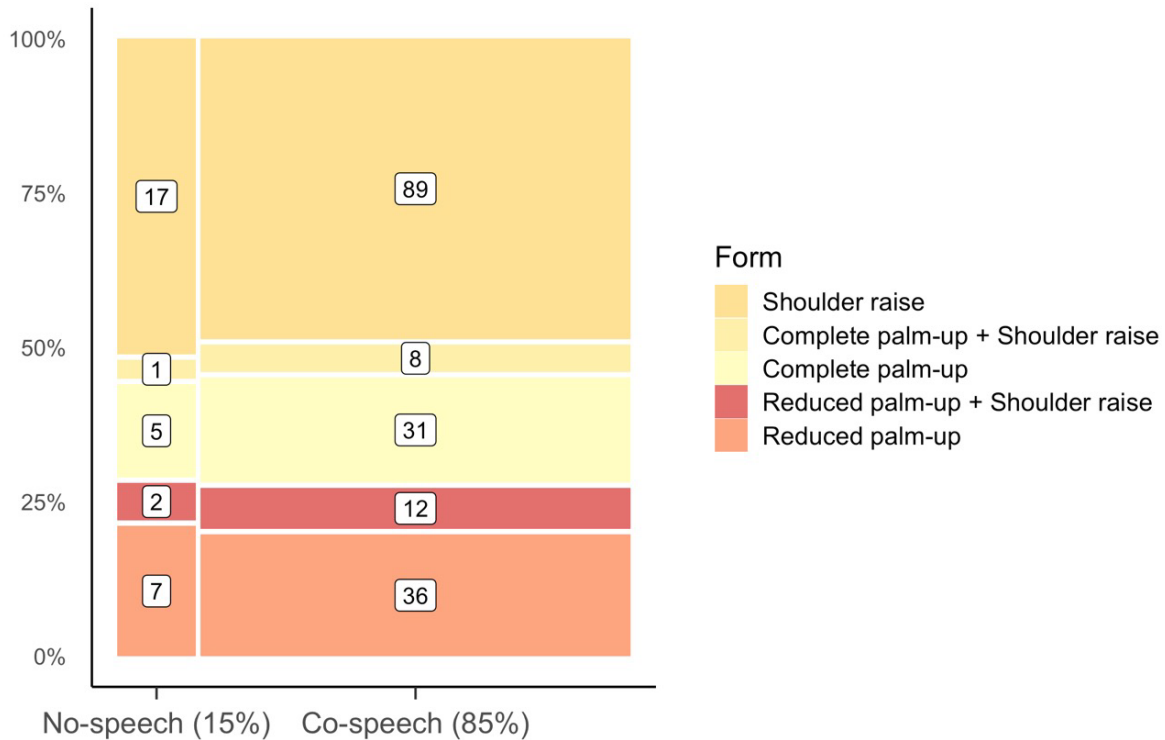
Figure 5.4: Distribution of shrug forms by speech presence in adolescence. Distribution of shrug forms does not differ for shrugs with and without speech.

Unlike in early childhood, there was no relationship between the form of adolescents' shrugs and the presence of speech (X^2(4, N = 208) = 0.26, p = .992). The distribution of forms was strikingly similar for co-speech and no-speech shrugs (Figure 5.4).

The lack of association between form and speech presence does not necessarily mean there is no association between kernel form and kernel meaning. A significant association between unreduced form and ignorance meaning persisting into early adolescence would indicate that shrugs remain rooted in the kernel emblematic mapping even in this later stage of development. The presence of such an association despite the lack of association between form and speech presence could indicate discourse contexts other than presence or absence of speech influencing emblematic use.

To test this possibility I built two mixed-effect logistic regressions corresponding to the models in Chapter 4. Model 1 included shrug meaning as the factorial dependent variable (kernel ignorance meaning/other meaning), gesture form (factorial; kernel unreduced form/other form) and gesture-speech relation (factorial; co-speech/no-speech shrug) as predictors, and a random effect for subject. Unlike the model used in Chapter 4 for the early childhood data, including age did not improve model fit and so was omitted. Model 2 used gesture meaning as a predictor of gesture form but was otherwise identical to Model 1.

These regressions support the findings from the Chi-square analyses above and show no association between form and meaning. There was a significant negative effect of speech presence on kernel meaning (β = -3.00, SE = 0.78, $p < .001$) but no significant effect of speech presence on kernel form (β = 0.14, SE = 0.48, p = .771). There was no significant association between kernel form and meaning in either model. Form did not predict meaning (β = 0.62, SE = 0.68, p = .360) and meaning did not predict form (β = 0.39, SE = 0.60, p = .518).

5.3.1.1 Interim discussion

In early adolescence children map many shrug forms to many shrug meanings, but still tend to use shrugs to convey ignorance more often when there is no supporting speech context. The slight tendency for no-speech shrugs to express ignorance may indicate that ignorance remains a "default" meaning for shrugs. As adolescents move through more advanced stages of pragmatic development, they create meaning with more extralinguistic contextual information and less often default to this meaning.

While the ignorance kernel meaning did seem to persist at least somewhat into adolescence, this was not the case for the kernel unreduced form. There were no differences between the forms of adolescents' shrugs with or without speech. This was the case not only for

the proposed kernel form, but for all five possible constructions of shoulder raises and palm-ups. Despite the ongoing presence of a kernel or default meaning, it was entirely unrelated to differences in form. The emblematic form-meaning mapping prominent in early childhood interaction seems to have dissolved by this point in development. This is not to say that adolescents never use shrugs with an emblematic function, but that they have moved somewhat away from a reliance on the shrugs as a nonverbal symbol and toward a more adult-like use of shrugs with both emblematic and pragmatic functions.

The basic dimension of presence or absence of speech context was no longer a determining factor of shrug form and was minimally influential on shrug meaning. The analyses that follow explore what other features of discourse context adolescents may be incorporating into their use of shrugs in conversation.

First, I explore adolescents' use of head gestures with shrugs. Like shrugs, head gestures serve both emblematic and pragmatic functions (Kendon, 2002; McClave, 2000; Whitehead, 2011), but head gestures were very rarely combined with shoulder raises or palm-up gestures in the early sessions. Are the differences in shrug meaning between early childhood and early adolescence driven by these new combinatory forms?

Second, I analyze features of sequence organization for communicative acts with shrugs to ask how adolescents use gestures as elements of turn design. We have seen that young children tended to use shrugs as a lexical emblem to convey ignorance when other semantic context was limited. How do adolescents use shrugs when context is limited not by the absence of meaning in speech but by the rules of coordinated conversation? I address this question in two parts. First, I explore the roles of sequential positioning (contingent and non-contingent acts) and request type (closed questions, open questions, and non-requests). Contingent acts directly build

on meaning from the initiating act. If adolescent speakers use shrugs to convey ignorance when acts are less contextualized, shrugs in non-contingent positions may be more likely to disinvest than shrugs in highly contextualized contingent positions. By similar logic, shrugs may be more likely to disinvest and express ignorance in non-request sequences than request sequences due to the added context of interrogative syntax. Finally, I narrow focus to contingent acts only to describe shrugs in preferred and dispreferred response formats. Speakers are required to do more interactive work when delivering a dispreferred response than when responding in an expected, preferred format. If speakers use shrugs for different kinds of meaning in preferred and dispreferred responses, this may help explain how shrugs are able to be reliably interpreted despite the wide range of seemingly contradictory meanings.

5.3.2   Head Gestures

There was no relationship between shrug kernel form and gesture-speech in early adolescence, but the question remains whether other aspects of form shape the function of adolescents' gestures. One obvious change across the two age periods was the complexity of composite forms. Children produced shoulder raises and palm-ups together in both early childhood and adolescence, but only adolescents also combined these forms with head gestures. In adolescence, 30.80% of shoulder raises and palm-ups were produced with a head gesture, compared to just 3% of in early childhood. Adolescents' shrugs included head nod, shake, and tilt form components while young children only combined shrugs with head shakes, never tilts or nods (Table 5.2).

Table 5.2: Proportional frequency of head gestures with shrugs, childhood and adolescence

	Shrugs in early childhood	Shrugs in early adolescence
No head	96.96%	69.23%
Nod	0.00%	10.10%
Shake	3.04%	12.98%
Tilt	0.00%	7.69%
Sum	*100.00%*	*100.00%*

Like shrugs, head gestures are early-emerging and function both emblematically and pragmatically in adults' speech (Kendon, 2002; McClave, 2000; Whitehead, 2011). It is possible that the observed patterns of use for shrugs reflect the functions of head nods, shakes, and tilts as gestures in and of themselves rather than those of shrugs characterized by shoulder raises and palm-ups. Adolescents frequently combined head gestures with shrugs[3], but all but one form construction the majority of shrugs were produced without any head gesture[4] (Figure 5.5-A).

A similar distribution was seen across shrug meanings; the majority of shrugs were produced without a head gesture for all meanings except disaffiliation (Figure 5.5-B). Adolescents combined head gestures with shrugs to produce all six meanings, but the frequency of including nods, shakes, and tilts was not evenly distributed across meanings. It is unlikely that the high salience of head gestures fully explains changes in shrug use, but differences in function

---

[3] The shrug composite gesture can be produced with many form components including shoulder raises, palm-up gestures, and head gestures as well as other form features not analyzed here, such as raised brow or tensed lips. In this section I use "shrugs" to mean shoulder raises and palm-up gestures and "head gestures" to mean head nods, shakes, and tilts.
[4] For the fifth form, a reduced palm-up with a shoulder raise, the number of shrugs without a head gesture was the same as those produced with a head shake, both more common than tilts or nods.

may be partially driven by how adolescents are incorporating these head forms into their shrug composite gestures.
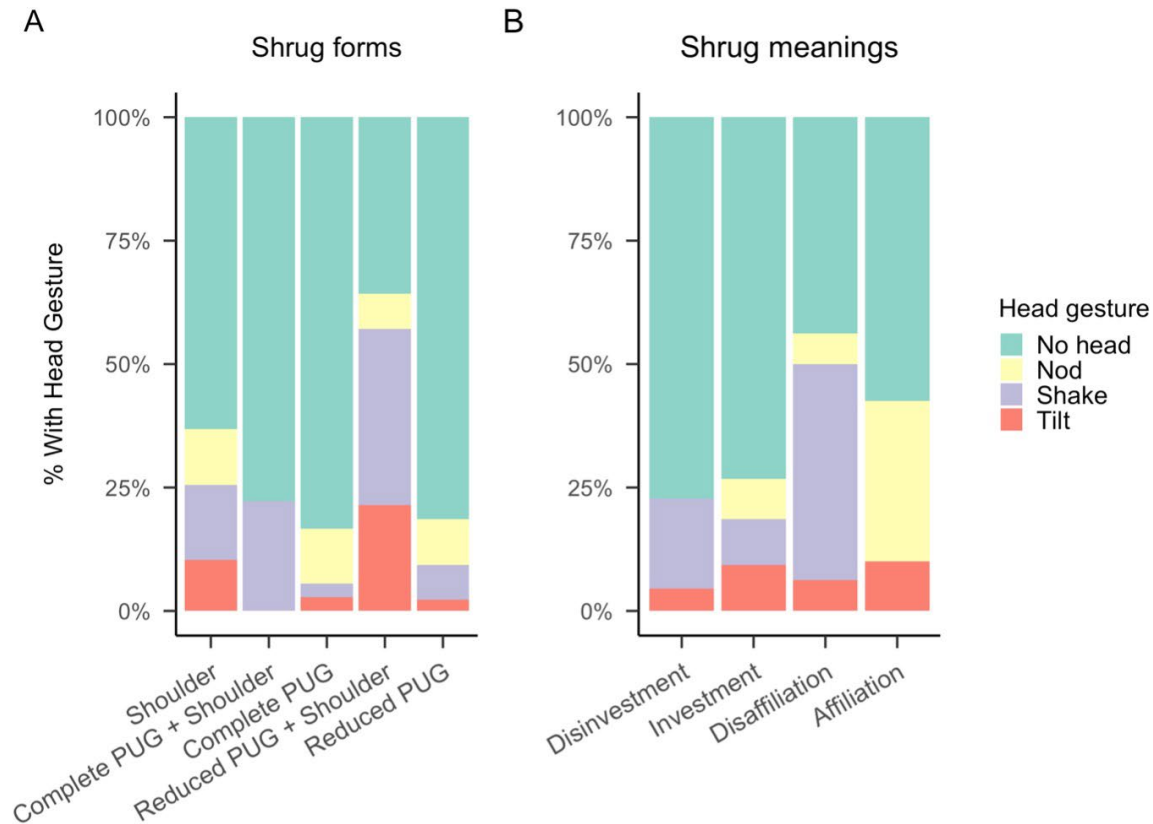
Figure 5.5: Distribution of head gesture forms by shrug form and meaning. Most shrugs do not have head involvement, but head nods, shakes, and tilts are all commonly combined with shoulder raises and palm-up gestures (PUG). Head gestures are not evenly distributed across meanings.

Between early childhood and early adolescence there was a general shift toward using shrugs with more "positive" meanings (Table 5.3). That is, in early childhood more shrugs disinvested (48.10% of all shrugs in EC) than invested (31.50%) and more shrugs disaffiliated (13.50%) than affiliated (6.90%). Just the opposite, in early adolescence more shrugs invested (41.30% of all shrugs in EA) than disinvested (31.70%) and more shrugs affiliated (19.20%) than disaffiliated (7.70%). One interpretation of this shift is that adolescents are using shrugs with

more adult-like versatility, more often using them for meanings distant from the ignorance kernel meaning.

Table 5.3: Proportional frequency of shrug meanings, childhood and adolescence

	Shrugs in early childhood	Shrugs in early adolescence
Disinvestment	48.07%	31.73%
Investment	31.49%	41.35%
Disaffiliation	13.54%	7.69%
Affiliation	6.91%	19.23%
Sum	*100.00%*	*100.00%*

An alternative possibility is that this shift is driven at least in part by the higher rates of head gestures. While it is an oversimplification to simply say "nod means yes, shake means no," there is undoubtedly a general, cross-linguistic tendency for nods to affirm and shakes to negate (Enfield et al., 2019; Kendon, 2002; McClave, 2000; Vincze & Poggi, 2016; Whitehead, 2011). It may be that head gestures, and head nods in particular, communicate more salient meanings than shoulder raises and palm-up gestures, essentially overriding shrugs' meanings. Perhaps it is a mistake to assume *shrugs* are as versatile as they appear.

One way to determine whether the meanings of shrugs are overridden by the meanings of head gestures is to ask whether shrugs that include a head gesture function more like head gestures or more like shrugs. If the meaning distribution for shrugs with head gestures is more similar to head gestures (without shrugs) than to shrugs (without heads), it could be that shrugs are not as flexible in meaning as they appear. Instead, it may be that a head nod or head shake sufficiently communicates meaning and is often combined with movements of the shoulders or hands to reinforce meaning or make the gesture more noticeable. The reverse is possible as well; shrugs with head involvement could be more similar in meaning to shrugs than to head gestures.

Shoulder raises and palm-ups may be sufficiently communicative on their own, so that co-produced head gestures reinforce meaning or call attention to form but do not determine function.

Figure 5.6 gives a visual comparison of the distribution of meanings for head gestures, shrugs (shoulder raises and/or palm-ups), and shrugs produced together with a head gesture. Head gestures are divided into three forms based on implicit assumptions in the literature. Classically, head tilts are always part of the shrug composite gesture and are frequently considered a shrug on their own, without any involvement of the shoulders, hands, or face. Head shakes are often treated as a component form, though not universally. Most definitions of the shrug composite gesture (alternatively compound gesture, gesture family, or gesture complex) do not include head nods as part of the composite, despite commonly co-occurring with other shrug forms.
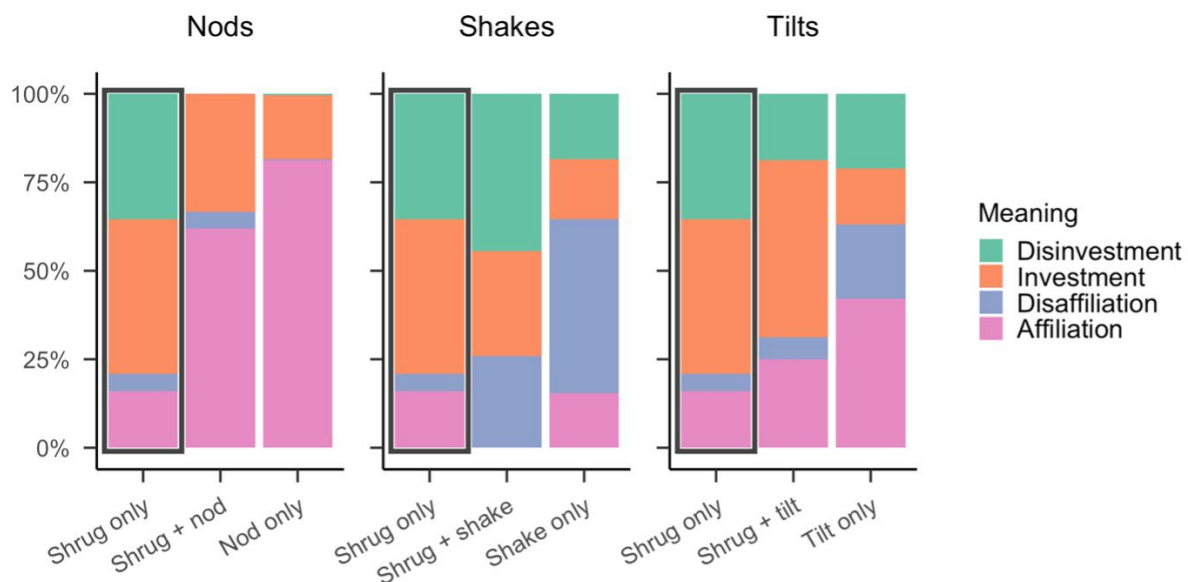
Figure 5.6: Distributions of meanings by head gestures and shrugs. Shrugs without head gestures, outlined in black, are a consistent comparison group for all head forms. The distribution of meaning for shrugs with head nods is more similar to that of head nods alone than shrugs alone. The distribution for shrugs with head tilts is closer to shrugs than to head tilts. The distribution for shrugs with head shakes is between the other distributions.

The proportional distributions of meaning suggest some head gestures carry the meaning of the shrug composite gesture, but not all (Table 5.4). Head nods seem to be more strongly associated with meaning than shrugs. *Shrugs produced with head nods* primarily communicated affiliation and investment (95.2%). This distribution is much more similar to that of nods produced alone (99.1% meaning affiliation or investment) than shrugs produced alone (59.8% meaning affiliation or investment).

Table 5.4: Comparison of meanings across shrugs with and without head gestures

	Shrug only	Nod only	Shrug + nod	Shake only	Shrug + shake	Tilt only	Shrug + tilt
Disinvestment	35.42%	0.43%	0.00%	18.46%	44.44%	21.05%	18.75%
Investment	43.75%	17.95%	33.33%	16.92%	29.63%	15.79%	50.00%
Disaffiliation	4.86%	0.43%	4.76%	49.23%	25.93%	21.05%	6.25%
Affiliation	15.97%	81.20%	61.90%	15.38%	0.00%	42.11%	25.00%
Sum	*100.00%*	*100.00%*	*100.00%*	*100.00%*	*100.00%*	*100.00%*	*100.00%*

In contrast, *shrugs produced with head tilts* were generally much closer in meaning to shrugs produced alone than tilts produced alone. Investment was the most common meaning for both shrugs alone (43.8%) and shrugs with tilts (50.0%), while head tilts alone infrequently expressed investment (15.8%). Affiliative meanings (including both affiliation and disaffiliation) were very common among head tilts produced alone (63.2%) but infrequent among both head tilts produced with shrugs (31.2%) and shrugs produced alone (20.9%).

The meaning distribution for *shrugs produced with head shakes* was not clearly aligned with the distribution of either shakes without shrugs or shrugs alone. About half of head shakes expressed disaffiliation (49.2%) while shrugs very rarely did so (4.9%). Shrugs with head shakes fell neatly in the middle, with about a quarter expressing disaffiliation (25.9%). Similarly, about

one-third of shrugs with shakes communicated investment (29.6%), evenly between shrugs alone (4.9%) and shakes alone (49.2%).

In summary, adolescents' shrugs produced with head nods were similar in meaning to other head nods, shrugs with head tilts were similar to other shrugs, and shrugs with head shakes were somewhere in between. Head nods are mapped tightly to affirmation and agreement, and this emblematic function of nodding seems to outweigh any alternative meanings that might by contributed by shrugs. Although head shakes are largely tied to acts of negation and disagreement, this mapping is not strong enough to entirely overwhelm meaning from shrugs. Shrugs with head shakes are pulled in two directions as both form components contribute meaning, though it is unclear what aspects of discourse context affect this balance. Unlike both nods and shakes, head tilts do not appear to function emblematically. Tilts produced alone can convey a range of meanings rather than mapping to a primary meaning. Without one dominant, conventionalized meaning head tilts do not seem to determine meaning as part of a shrug composite gesture.

5.3.3   Turn Design

Adolescents not only produced new shrug forms, they produced shrugs with a very different distribution of meanings. In early childhood shrugs communicated ignorance more than any other meaning, and a large majority of shrugs produced without speech conveyed ignorance. Although ignorance remained one of the primary meanings for no-speech shrugs, no-speech shrugs communicated affiliation just as often as ignorance. Overall, adolescents used shrugs for a wide range of meanings, both (dis)investing and (dis)affiliative.

Changes in formal features like the addition of head gestures described above may account for some of this difference, but function may play a role as well. By early adolescence,

children are more sensitive to social expectations in conversations and the rules of sequence and preference organization. To explore whether and how adolescents incorporated shrugs into their turn design I analyzed the frequency of shrugs and the distributions of shrug meanings across sequential positioning, types of request sequences, and response preference format.

5.3.3.1   Sequential contingency and request sequence types

Adolescents produced shrugs in contingent, non-contingent, and backchannel positions, but the frequency of shrugs in each position varied by request sequence type (Figure 5.7). For open (e.g., *wh*-questions) and closed (e.g., yes/no questions) request sequences, shrugs overwhelmingly occurred in a contingent position. On the other hand, in non-request sequences there were nearly twice as many shrugs in non-contingent positions ($N = 72$) as contingent positions ($N = 40$).
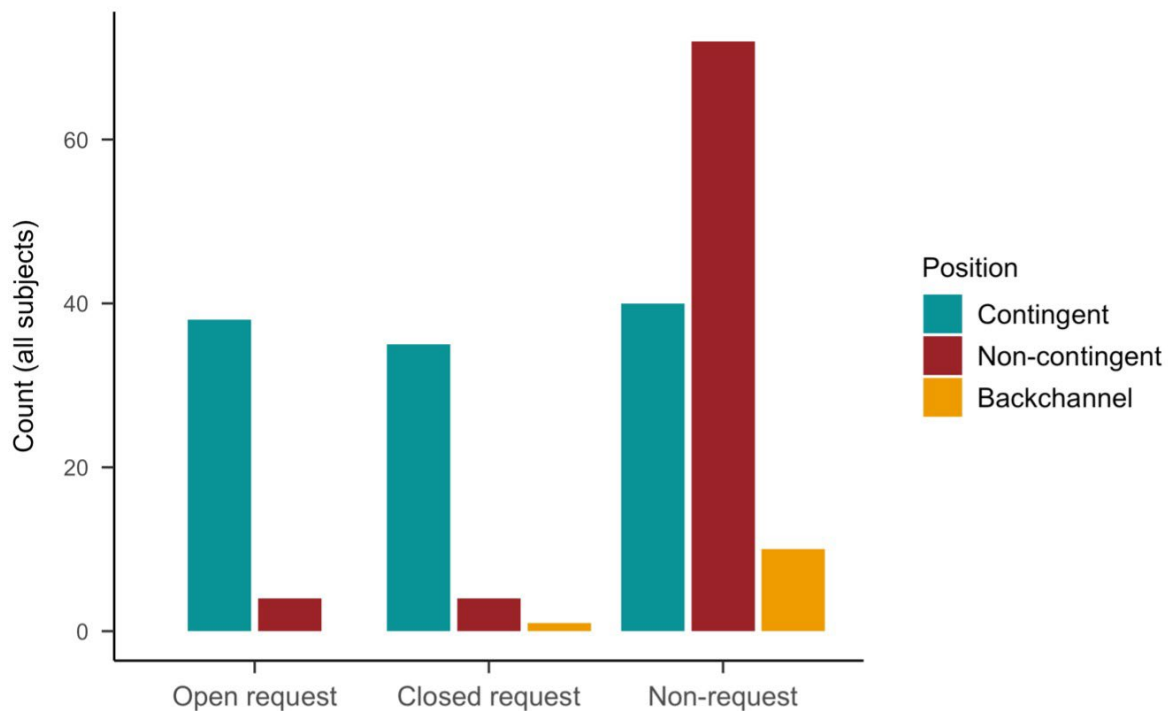
Figure 5.7: Distribution of shrugs by request type and position. Adolescents shrug with contingent acts for closed, open, and non-request sequences, but only shrug with non-contingent acts in non-request sequences.

In other words, adolescents frequently shrugged when offering responses to questions and when building on parents' assertions and assessments. Adolescents also used shrugs for their own turn-initial statements but almost never shrugged while asking a question. These results are in line with the framework of gradient response mobilization from Stivers and Rossano (2010). Questions that attempt to initiate a turn sequence are structurally designed to elicit a response. Assertions, assessments, and other tellings that attempt to initiate a turn sequence do not expect a response in the same way questions do, but marking them with a shrug may help mobilize a response.

Overall, shrugging in contingent positions was more common than shrugging in non-contingent turns. This difference may be explained by shrug meanings across types of request sequences. Contingent acts typically lack linguistic structures that actively solicit engagement like interrogative syntax. If speakers use shrugs to indicate continued engagement, contingent shrugs might more often use affiliative and disaffiliative shrugs. These gestures establish a relationship between speaker and listener in the interaction rather than commenting on the topic of talk.

Figure 5.8 charts the frequencies of adolescents' shrugs with each meaning for each position. This figure shows that the higher rate of shrugs in contingent turns is largely explained by affiliative and disaffiliative shrugs. Contingent (dis)affiliative shrugs made up 16.20% of all shrugs while non-contingent (dis)affiliative shrugs made up just 3%. Production of other shrugs was comparable across contingent and non-contingent turns, accounting for 39.30% and 36.30% respectively of all shrugs.
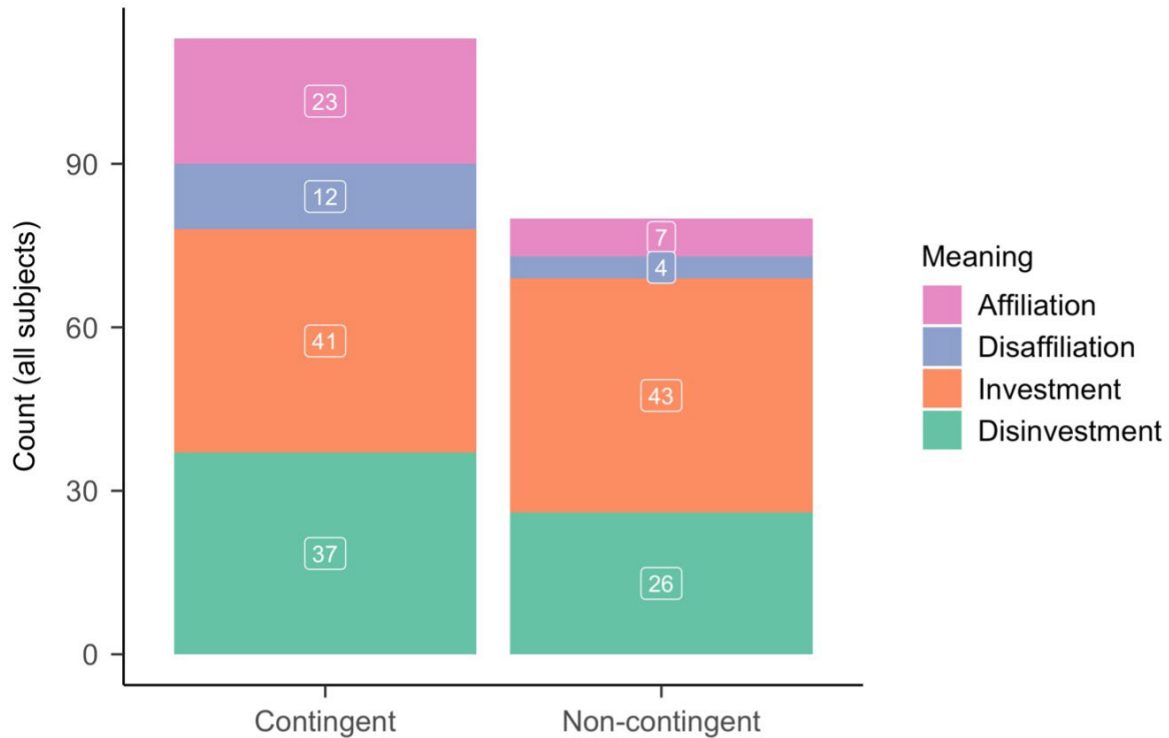
Figure 5.8: Distribution of shrugs by position and meaning. The higher rate of shrugs in contingent turns is largely explained by affiliative and disaffiliative shrugs.

5.3.3.2   Preference

Shrugs have a peculiar property of not only expressing a wide range of meanings but expressing contradictory meanings. They communicate agreement and disagreement, certainty and uncertainty, interest and disinterest, etc. These polar pairs of meaning are an opportunity for shrugs to have interesting interactions with preference organization. This range of meanings might reflect preference in speech, where shrugs communicate generally preferred meanings (e.g., affiliation, investment) in complement to preferred speech acts and dispreferred meanings with dispreferred speech. Alternatively, if shrugs are capable of communicating these meanings independently of speech, speakers may use shrugs to complement or contradict speech. For example, a speaker may give the preferred response "yes" to a polar question but use a shrug to

indicate uncertainty. While the speech is unequivocally preferred, the gesture acts as a hedge by mitigating the strength of that preference.

In the final analysis of this chapter I look within contingent responses to examine how shrugs are used in responses with preferred and dispreferred formats. First, I compared overall use of shrugs in preferred and dispreferred responses across different types of request sequences without considering shrug meaning (Figure 5.9). This revealed a clear division between shrugs are used in response to open requests vs closed requests and non-requests. For open requests sequences, shrugs overwhelmingly occurred with preferred responses (84.20%). The rates of shrugs with preferred and dispreferred responses were much more balanced for both closed request sequences (57.10% preferred) and non-request sequences (52.50% preferred).
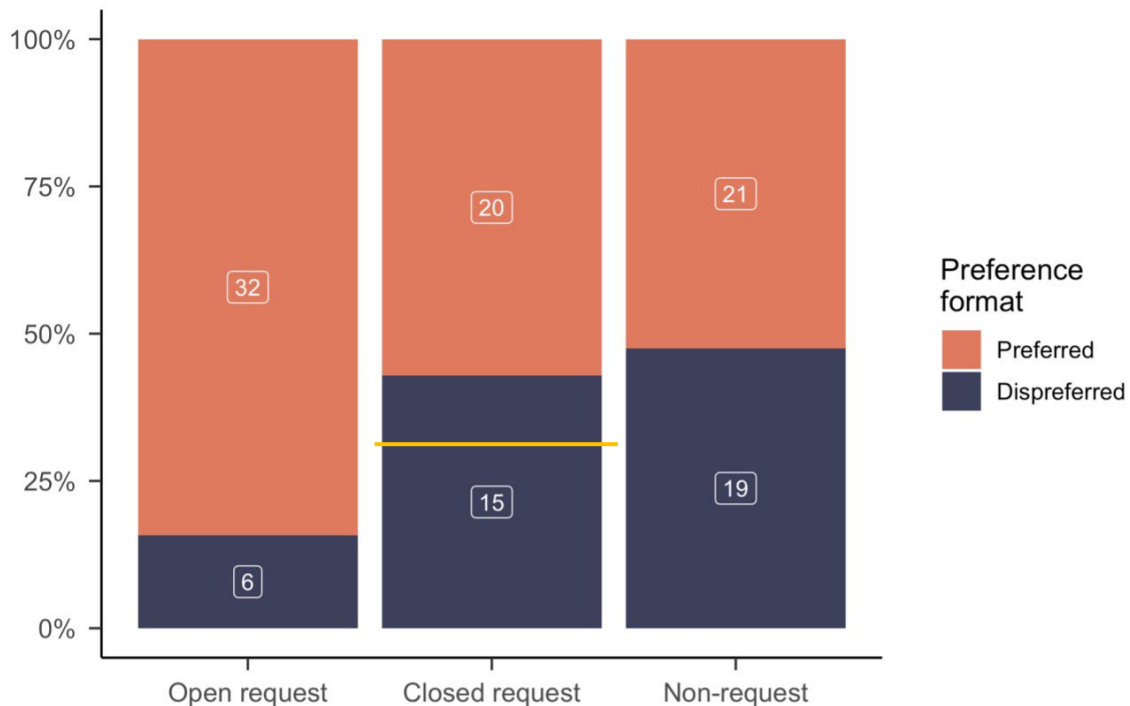
Figure 5.9: Distribution of shrugs by request type and response preference (contingent acts only). Adolescents produce shrugs in preferred responses more often than dispreferred. This difference is considerably larger for responses to open requests than closed or non-requests. The yellow bar indicates expected rates of preference for American English-speaking adults' polar answers (Stivers et al., 2009).

Table 5.5: Distribution of shrugs by preference format and request type, adolescence

	Open request	Closed request	Non-request
Dispreferred	15.79%	42.86%	47.50%
Preferred	84.21%	57.14%	52.50%
Sum	*100.00%*	*100.00%*	*100.00%*

Although the majority of responses with shrugs for all request sequences were in a preferred format, the differences in frequency for closed requests and non-requests are slimmer than we might expect to see for all responses regardless of gesture use. The expected base rate for preferred responses to closed requests in American English has been reported to be around 72% for adults (Stivers, 2010; Stivers et al., 2009) and 64% for children (Stivers et al., 2018), both above the 57.14% of preferred responses with shrugs in these adolescent data (Table 5.5).

Without a direct comparison between preference rates of responses with and without shrugs it is not possible to say conclusively that shrugs occur disproportionately more often with dispreferred responses. If this was the case however, one possible explanation is that shrugs in contingent positions act as hedges. They may mitigate the undesirability of dispreferred responses or the speaker's commitment to a preferred response.

In a final comparison of shrugs' role in turn design, I compared the proportional distributions of stance-taking meanings for contingent acts with shrugs. If adolescents use shrugs to mitigate dispreferred responses, we should expect shrugs to disinvest or disaffiliate in dispreferred responses more often than they invest.[5]

---

[5] Although it is possible to affiliate with the listener in dispreferred responses and disaffiliate in preferred responses, these cases are, predictably, extremely uncommon.
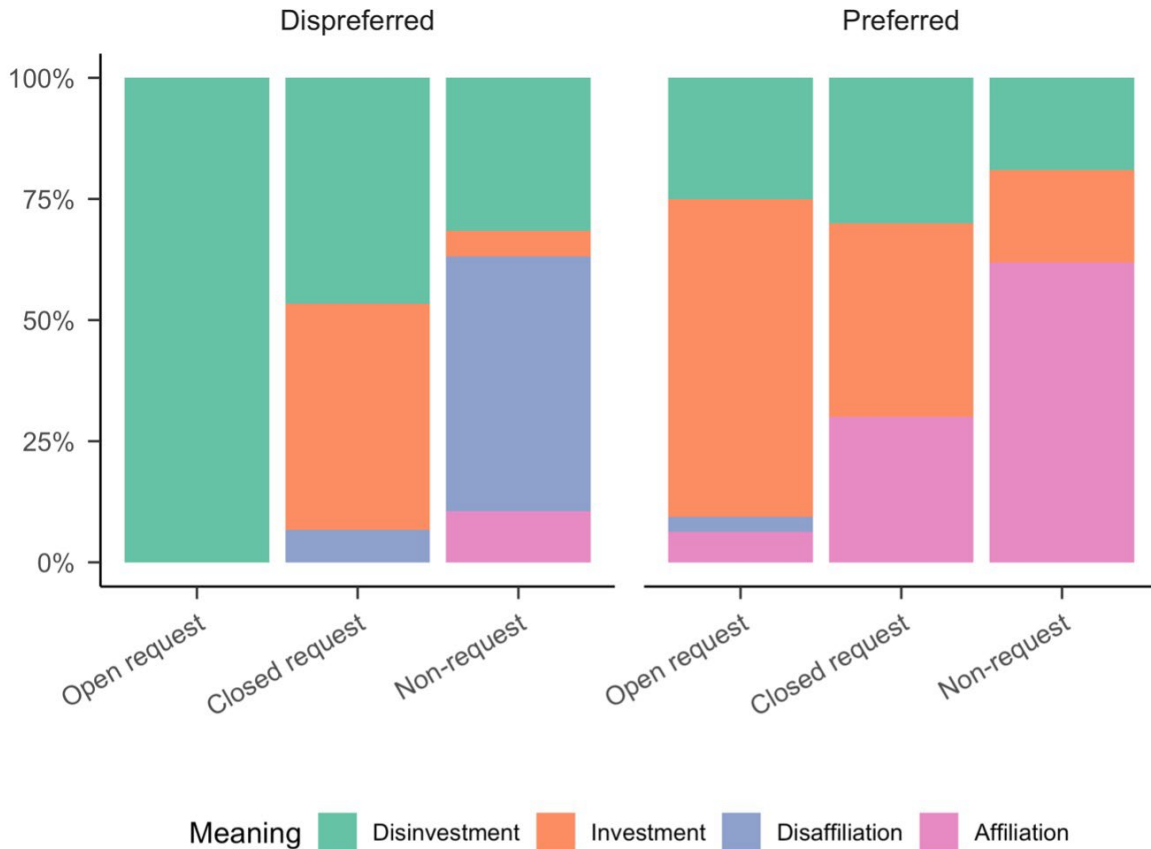
Figure 5.10: Distributions of shrug meanings by request type and preference format(contingent acts only). Meaning distributions differ by both request sequence type and response preference.

Figure 5.10 presents the proportional distributions of shrug meaning in responses to open requests, closed requests, and non-requests for both preferred and dispreferred formats. There is some evidence adolescents are using shrugs to mitigate dispreferred responses for open requests and non-requests, but not for closed requests (Table 5.6). Although there were very few dispreferred responses to open requests (N=6), all of these shrugs expressed disinvestment, typically communicating ignorance. For responses to non-requests a large majority of shrugs either disinvested from the topic (31.60%) or disaffiliated with the listener (52.60%), while only 5.30% invested in the proposition. On the other hand, investment was very common in dispreferred responses to closed requests, occurring at exactly the same rate as disinvestment (46.70%).

141

Table 5.6: Distribution of dispreferred shrug meanings by request type, adolescence

	Open request	Closed request	Non-request
Disinvestment	100.00%	46.67%	31.58%
Investment	0.00%	46.67%	5.26%
Disaffiliation	0.00%	6.67%	52.63%
Affiliation	0.00%	0.00%	10.53%
Sum	*100.00%*	*100.00%*	*100.00%*

In other words, while these shrugs may sometimes hedge negative responses, just as often they do just the opposite. If it is indeed that case that adolescents' rate of dispreferred responses with shrugs is disproportionately higher than the rate of dispreferred responses without shrugs, this difference is *not* explained by shrugs' "default" meaning of ignorance or disinvestment.

Table 5.7: Distribution of preferred shrug meanings by request type, adolescence

	Open request	Closed request	Non-request
Disinvestment	25.00%	30.00%	19.05%
Investment	65.62%	40.00%	19.05%
Disaffiliation	3.12%	0.00%	0.00%
Affiliation	6.25%	30.00%	61.90%
Sum	*100.00%*	*100.00%*	*100.00%*

Shrugs defaulting to disinvestment could also serve to distance the speaker from preferred responses ("I'll answer you to remain cooperative, but I am uncertain and/or disinterested."). As with dispreferred responses, if adolescents use shrugs to distance themselves from preferred responses due to uncertainty or disinterest, we would expect shrugs to disinvest in preferred responses more often than they invest, but this was not the case. For all three types of request sequences, adolescents were much more likely to shrug for investment and affiliation than disinvestment (Table 5.7).

5.4   Discussion

The descriptive analyses in this chapter illustrated that children's pragmatic gestures continue to evolve in terms of both form and function beyond early childhood. Like young children, adolescents tended to use shrugs to mean ignorance when produced without speech. Unlike young children, ignorance was not the primary meaning for adolescents' shrugs. Ignorance was the most common meaning for shrugs without speech, but only accounted for about one-third of no-speech shrugs, roughly equivalent to the share of no-speech shrugs that expressed affiliation.

It may be that the kernel meaning of ignorance persists into adolescence, so that shrugs "default" to an ignorance meaning without sufficient context. However, by 11-13 years of age children are more proficient conversationalists. They are more capable of incorporating discourse contexts into their communicative acts and do not default to the kernel meaning as much as they did in the pre-school years.

The distribution of meanings for adolescents' shrugs somewhat paralleled meanings in early childhood, but the distribution of shrug forms tells quite a different story. Unlike young children, adolescents did not vary shrug form by speech presence. They were equally likely to use any shrug component form or combination of forms with co-speech and no-speech acts. Predictably, there was no relationship between kernel ignorance meaning and kernel unreduced form.

These first analyses were conducted in parallel to those in Chapter 4. The results from early childhood showed that early in pragmatic development the form and meaning of shrugs is strongly associated with the presence or absence of speech. The parallel results from this chapter showed that adolescents may still default back to a predictable kernel meaning of ignorance

when speech context is unavailable, but they do not default to a recognizable form. While shrug gestures may still map to a primary meaning, shrugs no longer function as ignorance emblems with a stable, primary form-meaning mapping.

The co-speech vs. no-speech division is a crude measure of the relationship between gesture and speech. It makes sense that this simple property or speech presence or absence does not play as significant a role in early adolescence. However, just because adolescents seem to use shrugs emblematically less often does not mean that other aspects of discourse context do not systematically affect the forms and functions of shrugs in their dialogues.

5.4.1   Head Gestures

Adults commonly combine head gestures with shoulder raises and palm-ups, and head tilts and shakes are often considered part of the shrug composite gesture rather than being treated as separate, co-produced gestures. In early childhood, children frequently produced shoulder raises and palm-ups in isolation and combination, but only 3% of their shrugs included a head gesture form. In adolescence, nearly one-third of shrugs included a head gesture. The five form constructions used in the early childhood analyses (using shoulder raises, complete palm-ups, and reduced palm-ups) were not associated with shrug meanings, but this major change in combining forms may be. Adults use both head shakes and head nods to add emphasis to their speech, express attitude, and offer backchannel feedback (Enfield et al., 2019; Kendon, 2002; McClave, 2000; Vincze & Poggi, 2016; Whitehead, 2011). The high rates at which adolescents combine shoulder shrugs and palm-ups with head gestures may be driving the lower rates of ignorance and higher rates of investment and affiliation.

Like adults and unlike young children, adolescents frequently combined shoulder raises and palm-ups with head gestures, including tilts, shakes, and nods. Consistent with how the

144

literature treats head tilts as a shrug form (e.g., Beaupoil-Hourdel & Debras, 2017; Debras, 2017; Debras & Cienki, 2012; Jehoul et al., 2017), shrugs with head tilts had a similar distribution of meanings to shrugs alone than to head tilts alone. In other words, adolescents used head tilts as shrug gestures rather than distinct emblems. While there is little discussion in the literature about the high frequency of producing head nods together with other shrug forms, there is an implicit assumption that head nods are not part of the shrug gesture. Consistent with this assumption, adolescents' shrugs with head nods had a distribution of meaning much more similar to head nods alone than shrugs alone.

It seems that the head nod may essentially overwhelm the meaning of a shrug. That is, while the head nod is itself a flexible pragmatic gesture capable of conveying a wide array of meanings, much like the shrug, the meanings of assent and agreement are so dominant it is difficult to say whether co-produced shoulder raises and palm-ups are themselves informative.

On the other hand, the frequency with which adolescents produced head gestures alongside shoulder raises and palm-ups raises concerns about discounting head nods. If tilts are always considered part of a shrug gesture and nods never are, we might expect tilts to be most commonly produced with shoulder raises and nods least commonly produced with other shrug components. In fact tilts were the least common head gesture, both with other shrug forms and overall.

These findings cannot definitively say whether head nods should be considered one component form in the shrug composite gesture or simply another gesture frequently produced together with shrugs, but there are important implications nonetheless. This does not mean the way we think about what "counts" as a shrug form is wrong, but it does call into question

145

whether some of the most ubiquitous assumptions are safe to make without empirical justification.

The distribution of meanings for shrugs with head shakes fell somewhere between the meaning distributions for head tilts and nods. They did not express meaning interchangeably with shoulder raises and palm-ups, but they also did not primarily express disaffiliation like head shakes without shrugs. The meanings of shrugs with head shakes are pulled in two directions by two salient gesture forms. We cannot safely assume the head shake is just another form component of the shrug. We also cannot assume that it is not.

Of course, these results only describe how *adolescents* used shrugs with head gestures in conversation. At age 13, when the final sessions took place, children are far from adult-like in their use of linguistic devices for coordinating conversation (Larson & McKinley, 1998; Nippold, 1998). Whether or not adults show similar patterns has implications for both nonverbal pragmatic development as well as our understanding of pragmatic gesture.

Adults may use shrugs with head nods for a wider variety of meanings and not overwhelmingly for affiliative meanings. If so, this would highlight how long pragmatic development lasts and offer new insight into what kinds of conversational skills are late-developing. Between early childhood and early adolescence children used shrugs less emblematically. It may be that head nods behave similarly, moving from primarily acting as a nonverbal lexical symbol to adopting more pragmatic and interactive functions, with this shift happening even later in development than for shrugs.

On the other hand, if the affiliative meaning of adults' head nods outweighs other possible meanings for co-produced shrugs, we may need to reevaluate how we assess shrug meanings in gesture research. Perhaps the meanings of shrugs that are so strikingly different from ignorance

146

(e.g., certainty, comprehension, obviousness) are more incidental. These meanings may depend more on other salient verbal and nonverbal information, like head nods.

5.4.2   Turn Design

Adolescents shrugged with both sequentially non-contingent turns and sequentially contingent turns and with open requests, closed requests, and non-requests. Overall, most shrugs occurred in contingent positions and in non-request sequences, but there was an interaction between position and request type. Shrugs were rarely used in the first position of a request sequence. In fact, shrugs were more than five times as likely to appear in the first position of a non-request sequence ($N=36$) than a request sequence (open $N=4$; closed $N=3$). These results support the idea that speakers use shrugs to mobilize response and communicate an intent to continue cooperation in the interaction. When an action in conversation is already marked with other features of response mobilization like interrogative syntax and rising intonation (Stivers & Rossano, 2010), there is little need to add a shrug. The interrogative construction already sufficiently communicates the desire for the interaction to continue by explicitly requesting an interlocutor's contribution. Without interrogative syntax or prosody, non-requests like assessments and assertions may benefit from additional cues like sustained gaze, epistemic asymmetry, or gestures.

Along the same lines, non-contingent acts may have more "built-in" intent to continue the interaction than contingent acts. They may be structured as requests or other first-pair parts of adjacency pairs that not only expect a response but expect a particular kind of response (e.g., greetings expect greetings, questions expect answers, offers expect acceptances or rejections). Contingent acts, on the other hand, have the potential to deliver dispreferred or socially undesirable responses, close a turn sequence, or end an interaction entirely. It is an open question

what features of a second-position turn make a third-position response normatively required (see Stivers, 2012), but the principles of response mobilization apply here as well. Contingent acts require more interactional work to solicit uptake than turn-initial non-contingent acts. Adding a shrug to a contingent act may help to signal an intent to continue with the cooperative interaction when there is relatively little else in the act that does so.

The distribution of shrug meanings by sequential positioning supports the idea that shrugs act on responses as signals of engagement. Overall, adolescents shrugged with contingent acts more often than non-contingent acts, but this difference is largely explained by affiliative and disaffiliative shrugs. These gestures function interactively, establishing a relationship of some kind between speaker and listener. Non-contingent acts can and do take affiliative and disaffiliative meanings, they may even explicitly request it, but acts in the first position are more likely to mark an intent to continue the conversation in other ways, like using a question frame.

Another aspect of turn design where a speaker's cooperative intention may be easier or harder to convey is response preference. Very generally, by offering a preferred response a speaker has implicitly indicated their cooperative intent because they have delivered the expected cooperative action. Offering a dispreferred response by no means indicates an intent to be *un*cooperative, but cooperation is not baked into the turn unit in the same way. Shrugs may function differently in preferred and dispreferred format responses depending on implicit intent.

The rates at which adolescents produced shrugs in preferred and dispreferred positions seem to support this idea. Although the majority of shrugs in contingent positions were preferred, shrugs were produced in far more dispreferred responses to closed questions than we might expect based on baseline preference responses to polar questions (Stivers, 2010; Stivers et al.,

2009; Stivers et al., 2018).[6] This study focused exclusively on communicative acts with shrug gestures and cannot claim with certainty that rates of dispreferred acts with shrugs are indeed higher than dispreferred acts without shrugs in these interactions. Speculatively, higher than expected rates of dispreferred responses to closed requests might be due to the shrug's kernel meaning of ignorance.

If shrugs tend to default to disinvestment without additional contextual meaning, shrugs may be frequently hedging dispreferred acts. Using a shrug to disinvest by expressing ignorance or uncertainty might hedge a communicative act, but disinvestment is not the only way to do so. For example, a speaker can deliver a dispreferred act of disagreement by saying "I disagree" either with or without a shrug. Although the shrug is disaffiliative, it may actually serve to "soften" the dispreferred response and communicate a less argumentative stance. However, in this very particular turn position – dispreferred contingent acts following closed requests – fully half of adolescents' shrugs invested in their responses. A shrug that invests in the act definitively does not serve as a hedge. If adolescents shrug with dispreferred responses more than we would expect, this is not because adolescents are using shrugs to disinvest in their answers.

Instead, the high rates of investment in dispreferred responses to closed questions as compared to non-requests may reflect how permissible dispreferred responses are in a turn sequence as relevant, if not preferred, responses. Polar questions are structured to allow for a negative/dispreferred response as easily as an affirmative/preferred response. Affirmation would be a better response, but the way to respond with negation is built-in to the request. Delivering a

---

[6] Rates of dispreferred responses to non-requests was even higher than for closed requests, but to the best of my knowledge there is no appropriate baseline comparison for this broadly defined category.

dispreferred response to a non-request, on the other hand, often has no predetermined form. The preference for possible responses may be just as clear (e.g., invitations should be accepted, not rejected) or somewhat opaque (e.g., acknowledgment, commiseration, and assistance are all preferred formats for responding to complaints, though in a specific context only one may be socially preferable). In either case, the speaker must do additional interactional work to negate, disagree, refuse, or otherwise offer a dispreferred response.

In theory, if the shrug kernel meaning is ignorance, disinvesting in preferred responses should be a common usage. A mitigated response is better than no response at all. Suggesting an answer while indicating uncertainty is preferred to expressing ignorance or changing the topic. In fact, disinvestment was a much less common meaning for shrugs among preferred responses than either investment or affiliation. However, there were small differences in the rates at which adolescents used shrugs to disinvest in their preferred responses to different request types. Open and closed requests are produced with interrogative syntax or intonation; unlike non-requests, they are structured to demand a response. If questions increase pressure for a response, speakers might disinvest in responses to questions more often than to non-requests. Disinvestment did make up a smaller proportion of preferred responses to non-requests (19.0%) than of responses to open requests (25.0%) or closed requests (30.0%).

While this is the pattern we should expect, differences in rates of disinvestment were dwarfed by differences in rates of investment and affiliation. Among preferred responses, the majority of shrugs in open request sequences invested in the response (65.6%) and very few affiliated with the listener (6.2%). The reverse was true for non-request sequences, with the majority affiliating (61.9%) and a smaller portion investing (19.0%). This may be explained with similar reasoning as above, regarding why shrugs might invest or disinvest in dispreferred acts.

For both open and closed requests, there are clear expectations for relevant responses. Polar questions should receive a yes or a no. Information-seeking *wh*-questions should receive information, or else express ignorance. By offering a preferred response to an open request, a speaker abides by obvious cooperative rules of conversation. The rules for responding to non-requests are not so obvious. For requests, the organization of the turn sequence has already established cooperative intent. For non-requests, the speaker must use some other communicative resources to maintain that intent. Shrugs in responses to non-requests may serve that purpose of closing distance between speaker and listener, whereas shrugs in responses to requests are free to comment on the distance between speaker and topic.

This chapter has explored the ways young adolescents use shrugs pragmatically in dialogue by directly comparing forms and meanings to shrugs in early childhood. Additionally, I described new and evolving forms and functions of shrugs within the early adolescence corpus. The results offer insight into how interactional functions of shrugs arise from an initial emblematic ignorance kernel. This study connects multiple bodies of prior research that have so far given some important information about gesture's role in pragmatic development but have also left a significant gap. We know that pragmatic gestures are important for early epistemic expression but not how older children use pragmatic gestures to take epistemic stance. We know that the new social contexts children are exposed to at school-entry allow older children to master complex strategies for managing interaction in speech but not how these contexts may allow for new interactional strategies in gesture. We know that some kinds of pragmatic gestures develop new functions for organizing monologic narrative in late childhood but not how pragmatic gestures develop new functions for organizing dialogic interactions. Finally, we know that in middle and late childhood children express epistemic certainty and uncertainty

multimodally but not how they do so with "uncertainty" or epistemic gestures specifically. This study exists at the intersection of these questions and lays groundwork for how future research can continue to put these literatures in conversation to fill this gap.

In the final chapter of this dissertation I discuss what we can learn from these results in light of early childhood findings from Chapters 3 and 4. I outline some important limitations of this study both on its own and considered together with the parallel analyses of Chapter 4. Finally, I explore concepts of distancing in more depth. Adolescents use shrugs together with the organizational structures of their speech to position themselves in relationship to their interlocutor and to their talk. Thinking critically about what this might look like for adults can help us situate shrugs not simply as ignorance emblems or even symbols for obviousness, agreement, or anything else, but instead as tools for dynamic stance-taking.

6   General Discussion

6.1   Emblematic and Pragmatic Shrugs in Early Childhood

The first two studies in this dissertation explored the forms and functions of pragmatic gestures in early childhood. Together, these studies demonstrate the necessity of accounting for gesture variation with the broad functional category of "pragmatic gestures" in communicative development. Further, they provide some insight into the ontology of shrug gestures by illustrating how young children use shrugs as emblems.

Chapter 3 showed the potential pitfalls of grouping together "pragmatic" or "non-referential" gestures without attention to differences in the dimensions of form and speech presence. Specifically, we saw that onset of production varied by gesture form, frequency of use over time varied by speech presence, and pragmatic function was associated with both. Breaking apart the larger category of pragmatic gestures even by these two basic dimensions revealed multiple distinct roles these gestures can fill in early communicative development. If these crude divisions revealed meaningful differences within this class of gestures, it seems a safe bet that more nuanced divisions would be even more informative.

The analyses in Chapter 4 again looked at dimensions of gestures' form and the presence or absence of speech. These results provide another example of why such divisions are important for understanding the function of children's earliest pragmatic gestures. This study complements the previous chapter but differs in perspectives on both form and function. Rather than comparing beat and palm-up forms, Chapter 4 explored early use of shoulder raises and palm-up gestures as component forms of the shrug composite gesture (following Debras, 2017; Givens, 1977; Morris, 1994/2015; and Streeck, 2009). Additionally, this study broke down palm-ups into

complete (full rotation with clear freeze in motion) and reduced forms (lacking full rotation or freeze). This attention to palm-up form builds on a sizable literature on palm-up gestures as a discrete phenomenon (e.g., Cooperrider et al., 2018; Ferré, 2012; Müller, 2004). Shoulder raises and palm-ups are frequently produced together as one shrug gesture by both children and adults. This chapter acknowledges that these forms are far from mutually exclusive and allows for the possibility that multi-form shrugs are functionally different from shoulder raises or palm-ups produced in isolation.

Alongside this alternative perspective on form, this study used a different approach to the determination of function. Chapter 3 discussed pragmatic function as a property of each communicative act that included a beat or palm-up. Chapter 4 centered on the function of individual gestures as they are contextualized at the discourse level rather than utterance level. Co-speech gestures do not necessarily communicate the same meaning that is encoded verbally. Many gestures do reinforce meaning, but many others supplement speech with new meaning or even contradict meaning in speech. In this study a gesture's "function" refers to the gesture's communicative contribution to the interaction, recognizing that some gestures may simultaneously serve both lexical and pragmatic functions. With this complementary approach, we again saw that even in very young children's interaction, these are meaningful distinctions. In these analyses we saw that form and meaning are related above and beyond the effect of presence or absence or speech. It is clear that both of these dimensions *together* tell us something about the meaning of children's shrugs.

One notable difference between the findings in these two early childhood studies was the effect of age across this span (14-58 months). In Chapter 3 children's co-speech gestures changed over time in terms of both function and frequency. In contrast, there was no effect of

age on the relationship between form and meaning in Chapter 4. These results may seem contradictory at first glance, but this apparent contradiction may be attributable to the two different approaches to function. The absence of an age effect in one study can help explain the presence of an age effect in the other.

In Chapter 3 both form and speech presence interacted with changes in function across early development, resulting in the functional distribution for acts with co-speech palm-ups being pulled in two directions with increasing age. The findings in Chapter 4 show us what might be driving that "pull." By looking at the associations between meaning and form in different speech contexts, we examine this interaction in a deeper way. This study used differences in patterns of use for shrug gestures with and without speech to suggest what a kernel shrug emblem might look like, then looked for changes in the use of that proposed emblem over time. Critically, even though this study used differences no-speech shrugs to propose a kernel meaning and form, there is no presumption that children of any age never use co-speech shrugs for emblematic functions or no-speech shrugs for pragmatic functions. In other words, the distribution of shrug meanings may change considerably over this early span, but that is different from changes in functional mapping between form and meaning.

As discussed in Chapter 1, in referring to emblematic and pragmatic functions of children's gestures I follow Bates (1976) and Peirce (1991). Function concerns the gesture's relationship to the context of the interaction; meaning concerns the listener's interpretation of the gesture, regardless of how they came to that interpretation. Emblems are nonverbal symbols, equivalent to verbal word units. They are an arbitrary mapping of some form to some meaning, conventionally agreed upon by a linguistic or cultural community. Context may allow for a more nuanced interpretation, but the essential meaning of the emblem is interpretable without any

contextual support. Gestures that serve a pragmatic function index something contextual and lose their meaning when removed from that context. Moreover, pragmatic gestures index some aspect of the interaction itself rather than an object or entity referent (like deictic gestures). In order to interpret these gestures, the listener needs information about social relationships, the emotions and knowledge of the interlocutors, and what contributions have already been made to the dialogue and by whom. Further, they must understand the social and linguistic rules that guide successful indexing within the interaction.

Function may be related to meaning, but it is not the same. A shrug that communicates ignorance may be emblematic or pragmatic depending on how much contextual information is necessary to interpret that meaning. Children may use shrugs with a widening range of meanings across the pre-school years, producing more gestures with non-kernel meanings that require some contextual information to interpret. What does not change in these early years is how children use shrugs emblematically. When contextual information is limited, children use a predictable form-meaning mapping and they do so throughout early childhood.

Taking together these two studies of pragmatic gestures in early childhood, the principal lesson learned has to do with how we approach this topic going forward. Research on multimodal interaction – whether focusing on adults or children, taking a cognitive or developmental approach, or originating in psychology or linguistics – should first and foremost recognize that pragmatic gestures are not throw-away hand-waving. Following that recognition, we should appreciate the formal and functional complexities of pragmatic gestures and afford them the same attention to detail we give deictic and representational gestures.

6.2   Connecting Early Childhood and Early Adolescence

Although the relationship between kernel form and kernel meaning did not change over the 14–58-month period, the first analyses of Chapter 5 show that by early adolescence children's functional use of shrugs had changed considerably. Obviously infants do not gesture like adults, so it should not be surprising that children use shrugs differently at age 2 and age 12. Predictable or not, this drives home the point that communicative development – and pragmatic development in particular – extends beyond early childhood.

The field of language development is dominated by research with infants and toddlers. Studies of older children and adolescents are often motivated by an application to academic outcomes, and consequently center around literacy and oral language skills that directly translate to the classroom, like metaphoric language, narrative, and persuasive discourse. Although there is a growing interest in pragmatic deficits in atypically developing children (Eigsti et al., 2011; Joffe & Nippold, 2012; Zaidman-Zait & Most, 2020), there remains relatively little research on the extended development of interactional resources as part of typical first language development. Further, the work that does exist often presents a description of typical language development as a means of better understanding language delays and deficits rather than a worthwhile avenue of inquiry on its own (for example Brice, 1992; Matthews et al., 2018; Russell & Grizzle, 2008; but see Nippold, 2016 for discussion of typical later language development from a speech-language pathology lens).

Unfortunately, despite a growing body of research on later language development near the end of the twentieth century, it seems interest has waned significantly in the last several decades (see Table 1.2 for a selection of this literature). Also unfortunate is the fact that the majority of

this literature concerns the development of pragmatics in speech with minimal attention to nonverbal or non-linguistic interaction in adolescence.

The descriptive analyses in Chapter 5 begin to fill some of these gaps. The observation that adolescents frequently used head nods with shrugs again speaks to the primary conclusion of Chapter 3, the necessity of a fine-grained approach to studying pragmatic gestures. Adolescents' use of shrugs with head gestures supports the presumption that nods are functionally distinct from head shakes and tilts and are not necessarily part of the shrug composite gesture. At the same time, adolescents combined shoulder raises and palm-ups with nods as often as or more often than other head gestures. While it may be inappropriate to consider head nods a form feature of shrugs on par with shoulder raises, palm-ups, or head tilts, it is also inappropriate to ignore the influence of nods on shrug meaning.

The relationships between forms and functions in all three studies support treating the shrug as a composite or compound gesture rather than considering shoulder raises, palm-ups, head tilts, and other component forms as individual but frequently co-produced gestures. It is clear that even 2- and 3-year-old children "mix and match" shrug forms and meanings to some extent. It is equally clear, however, that children and adolescents do not use shrug component forms interchangeably. Creating overgeneralized categories of gesture can be counterproductive to understanding how these gestures are actually used in conversation.

The division of the pragmatic gesture space by dimensions of form and speech presence in Chapter 3 revealed important differences between how they emerge and develop across early childhood and illustrated the problematic nature of oversimplification in this regard. The dramatic differences in meaning we see when adolescents combine shoulder raises and palm-up gestures with head nods, shakes, and tilts point to a similar problem. The gesture literature takes

for granted that head nods are not part of the shrug composite in the same way as head shakes may be or head tilts definitively are, but it is not enough to intuitively presume this to be the case. We need to explicitly recognize and account for differences in how head gestures interact with other formal elements of shrugs or else risk conflating effects and misattributing meaning. We recognize that shrugs often convey obviousness and certainty. Are these meanings as prevalent or strong as they appear, or might they be explained by co-occurring head gestures that go ignored or discounted in function analyses? If ignorance is the kernel or default meaning of shrugs, perhaps an understanding how shrugs are used together with head nods can offer clearer picture of why shrugs can mean not only absence of knowledge or certainty but also *abundance* of knowledge or certainty.

Head nods are also another opportunity to think about how gestures enter a child's nonverbal lexicon as emblems and grow outward to more flexible pragmatic, interactive functions. This question is beyond the scope of this dissertation, but it is worth recognizing that what I propose for the shrug gesture is not necessarily idiosyncratic.

6.2.1 Outstanding Questions About Adolescents' Gesture

The Language Development Project corpus is truly unique, and the middle school visits in early adolescence are a wealth of unexplored data we can use to begin to address the lack of research in this area. There is much more to discover in these interactions than simple descriptions of middle-schoolers' shrugs, but even this narrow look at pragmatic gestures points to some avenues of exploration.

One of the most striking differences between the way children interacted with their parents in the early childhood and adolescent sessions was the use of backchannel feedback. Although it is not possible say with certainty that children did not make any kinds of

backchannel contributions to interaction in early childhood, they never used shrug gestures to backchannel. In contrast, all the children used gestures to offer backchannel signals of agreement and understanding in the adolescent sessions. These backchannel comments nearly always included a head nod, but shoulder raises and palm-ups were not infrequent, with about 5% of shrugs in the adolescent data in the backchannel position.

It is interesting that backchannel shrugs were – like nods – used to signal comprehension ("following") and agreement. While these are the most typical functions of backchannels, they are not the only ones. Backchannel feedback can mean the converse as well, confusion or disagreement, letting the speaker know they need to repair, expand on, or defend their position. If the shrug's kernel meaning is ignorance, then shrugs are well-suited to send these signals without taking the turn. There is a clearer extension from absence of knowledge to *absence* of understanding or agreement than *presence* of understanding or agreement.

One possible explanation for why adolescents do not use shrugs as negative backchannels is that adults simply might not use them this way. We know that adults use shoulder raises, head tilts, and palm-ups both to distance themselves from topics and interlocutors and to express confusion and indicate need for conversational repair (Bavelas et al., 1992; Debras & Cienki, 2012). However, adults may primarily use shrugs to communicate these meanings when they themselves hold or take the floor and not as backchannel devices when their interlocutor has the floor. This is one possibility, but it would be surprising in light of the evidence for a kernel meaning. If shrugs default to an ignorance meaning when supporting context is limited, shrugs should be a perfect tool for passively communicating these meanings in a backchannel position without claiming the floor.

An alternative explanation may instead have to do with the kinds of feedback – both verbal and nonverbal – adolescents give their parents in these interactions. The later sessions took place when children were in middle school, in the summers following fifth and seventh grade. In these transitional years between being a child and being a teenager, children's social worlds are changing dramatically. They spend as much time in the classroom as at home, balance identities between multiple peer groups, and distance themselves from their parents. At this age, and in the high school years, adolescents develop new ways of giving cooperative feedback in conversation, like using constructive interruptions and expressing empathy. Adolescents not only increase their use of these behaviors over time, they use them more often in interactions with peers than interactions with familiar adults (Larson & McKinley, 1998). It may be that either the social dynamics of parents and children at this age or the topics of the elicited conversations led to a narrower set of feedback signals.

This leads to a final point about the specific context in which these adolescent data were collected, which is both a limitation and an opportunity. In these later sessions, spontaneous caregiver-child conversation is elicited though a series of "school-like" tasks. The tasks are designed to be familiar to this population of American middle-school students, with topics and formats similar to some classroom assignments. While children should feel comfortable with these kinds of activities, they may not be as comfortable when the tasks are transplanted from the classroom environment into their kitchens and living rooms. They are working through problems from academic domains with parents rather than teachers or peers. As a final complication, the studies in this dissertation use tasks in the social and art domains, prompting discussion about school policies, moral dilemmas, and music preferences. Unlike the math and science tasks they complete, these "problems" have no right answer.

As children enter adolescence as 11- and 13-year-olds, they develop unique identities and see a major upheaval to their relationships with parents. How might new stages of social and cognitive development interact with the familiar-yet-unfamiliar nature of the tasks and affect the ways young adolescents communicate in this context? Do children feel more comfortable in their homes with family, or are the academic tasks uncomfortably out of place? Are they more able to voice dissenting opinions or risk being wrong with uncertain guesses when they are in the physical and social safety of their homes? Or are the stakes of dissent higher in conversation with caregivers who are significant influences on developing values? What kinds of individual differences in pragmatic strategies could be attributable to parenting style and household culture? We can only speculate about how any of these circumstances may have affected children's gestures in these sessions, which is a significant limitation of this research. Even so, a better understanding of the ways children use language and gesture in these contexts can encourage us to see these as important research questions in their own right and to imagine how we can use these data to explore them.

Although one of the most valuable aspects of the Language Development Project corpus is its longitudinal nature, the clear differences in interactional contexts between early childhood and early adolescence complicate longitudinal analyses. Conversations in the adolescent sessions were elicited, centered on particular topics some families might not ordinarily discuss, and took place in one space without interruption. Interactions in the early childhood had as few limitations as possible. Not only was there no guidance for conversation topics, there was no direction for parents to interact with their children at all.

The communicative intent variable ultimately was not included in these analyses, but illustrates how these differences affect the language data. In this annotation, communicative acts

were coded as objective if they concerned facts – assertions about the world that are provable or disprovable by a third party. Acts were subjective if they concerned opinions, beliefs, or topics without objective truth. A third party might agree or disagree with a subjective act, but could not prove or disprove it. The third and final category of communicative intent was behavior or scenario management. These acts functioned to direct the immediate context, for example announcing an intention to change activities, instructing an interlocutor take an action, or narrating pretend play. More detail about communicative intent can be found in Appendix A.

Scenario intent was extremely common in the early childhood interactions, making up 49% of all children's communicative acts with shrugs. This should be unsurprising given that behavior was unrestricted and that caring for toddlers generally requires a good deal of behavior management. In contrast, scenario intent was extremely rare in the later visits with more restricted interactions, only 6% of coded communicative acts. For non-scenario acts, the balance between subjective and objective acts was also dramatically different across the two periods. In early childhood objective intent was three times as common as subjective intent (75% vs 25% of non-scenario acts). In adolescence the reverse was true; objective intent was about half as common as subjective (33% vs 67%).

These differences in communicative intent between early childhood and adolescence are worth investigating. Could rates of scenario talk speak to parenting style or family dynamics? Could relative rates of objective and subjective talk reflect children's persuasion strategies? Are these rates for the intent of acts with shrugs comparable to those without? If not, what does this say about how children use shrugs to express epistemics and attitude? But these vast difference in the kinds of talk children are producing means a longitudinal comparison of sequence

organization properties (i.e., preference, request sequence types, and positioning) would be inappropriate before we have a better understanding of what drives this shift.

Although longitudinal analyses with these data may difficult, it is worth recognizing that this difficulty is due to contextual differences rather than the annotation scheme. The coding systems developed for this dissertation were designed to be applicable to dyadic interactions regardless of age or data collection method. Coders successfully used a single, comprehensive annotation scheme for all transcripts and achieved high inter-rater reliability for both developmental periods. That is, this annotation is relevant and applicable to infants' first words, preschoolers' pretend play, and middle schoolers' problem solving and persuasion strategies. Although parent language was not coded for the analyses in this dissertation, the scheme was designed to apply to adults' speech as well. Further, although these analyses only include communicative acts with gestures, the conversation analysis-inspired sequence organization coding may also be used for speech alone.

These two coding systems for pragmatic gesture and sequence organization open the door to comparisons across age, speaker, gesture presence, and discourse contexts. Many of the questions that emerge from these studies may be answered with matched coding. For example, in order to determine whether adolescents use shrugs to mobilize response we can compare sequential positioning and request type of acts with shrugs to those without. One step further, we could do the same analysis with parents or elicited conversations between adults to ask when children develop adult-like patterns of response mobilization.

6.3   Stance-taking Functions of Shrugs

The pragmatic gesture coding scheme used in Chapters 4 and 5 has one critical difference from most comparable schemes in existing literature. There is considerable variation in how

researchers label gesture meanings, but some of the most common glosses for shrugs

(un)certainty, (dis)interest, (dis)agreement, incapacity, and obviousness. My annotation scheme

uses neither these terms nor terms that can be directly mapped onto these typical meanings.

Instead, my coding uses the terms (dis)investment and (dis)affiliation. *Investing* shrugs

may be expressions of epistemic state (e.g., certainty, obviousness) or attitude (e.g. interest).

Affiliation shrugs express an alignment between speaker and listener, which may take the form of

agreement, acknowledgment, or affirmation. The original coding distinguishes some

submeanings, but there is no distinction between knowledge and affect.[1]

Why deviate from the literature in this way? There are two reasons, one practical and one

theoretical. Practically, it was often impossible to say with certainty that a shrug was

communicating something about knowledge but not attitude or vice versa. There were many

occasions when a shrug clearly communicated uncertainty rather than disinterest or agreement

rather affirmation, but just as many where these distinctions were highly ambiguous.

Consider a parent and child participating in one of the adolescent tasks. The parent asks,

"Which is worse, cheating or stealing?" The child shrugs their shoulders and responds,

"Stealing." The shrug may be epistemic, indicating uncertainty or indetermination, but may just

as likely be affective, indicating disinterest in the question or lack of commitment to their

opinion. It may very well express both uncertainty and disinterest at the same time.

A similar problem exists for those meanings in the affiliation category. In the same

scenario, the parent asks "Well, stealing is wrong, don't you think?" The child shrugs and

---

[1] Specifically: inquiry and ignorance as forms of disinvestment; directing turn and affirmation as forms of affiliation; negation as a form of disaffiliation.

responds, "Yeah." This shrug expresses agreement but also offers affirmation to the parent's closed question. It may also be that the child is not in agreement with their parent but simply acknowledging an understanding of the proposition to indicate the parent should continue.

One possibility for discerning epistemic and affective meanings might be to differentiate acts by communicative intent. As discussed above, the annotation scheme classifies acts as objective or subjective. Objective acts refer to provable or disprovable information (regardless of veracity). Stance-taking gestures with objective acts might be more epistemic than affective in nature. The speaker has full, partial, or no knowledge about the topic and can similarly assess or ascribe knowledge states to others. Gestures that take stance with subjective acts, those regarding information that may be agreeable or disagreeable but not provable, may generally express affect and attitude. While a speaker may lack access to objective information, in theory they cannot lack access to their own emotional states, beliefs, and judgments.

By this reasoning, shrug gestures could be interpreted with epistemic meanings (e.g., certainty, obviousness, ignorance) for objective acts and affective meanings (e.g., interest, non-responsibility, inability) for subjective acts. But there is an obvious problem with this strategy: people have subjective attitudes toward objective information and make objective reasoning about subjective claims. That is, a speaker may hold and express emotions, interests, and opinions about undisputed truths. They may express subjective assessments contingent on verifiable facts or assess their own or others' objective knowledge about subjective opinions. In short, considering not only the structure of an act but the intent behind it may give a better understanding of stance-taking, but cannot reliably discern knowledge and affect.

This leads to a second, theoretical reason for using these unconventional categories of meaning. Although shrugs are often discussed as stance-taking gestures (e.g., Debras & Cienki,

2012; Jehoul et al., 2017), it is easy to lose sight of how shrugs take stance when we break down function into many discrete categories of meaning. Using concepts of investment and affiliation allows us to map the stance-taking functions of shrugs onto an existing framework for talking about stance-taking, the stance triangle.

Du Bois' (2007) stance triangle is made up of three vertices: two subjects of interaction, typically speaker and listener, and the object of interaction, the proposition or topic of talk. Stance-taking happens when speakers establish a relationship between two of these vertices with acts of evaluation, positioning, and alignment. In Du Bois' words, "I evaluate something, and thereby position myself, and thereby align with you" (163).

Subjects relate to objects through acts of evaluation and positioning. Speakers evaluate topics of talk with assertions and assessments, assigning some property to the object: it is good, gross, true, important, etc. They position themselves in relation to a topic by announcing their own state: I am happy, interested, frustrated, unsure, certain, etc. Critically, both evaluation and positioning can be either epistemic or affective. Whether a shrug expresses uncertainty or disinterest, it takes stance in the same way, relating subject and object. With respect to gesture, evaluation and positioning make up investment and disinvestment meanings. Although the difference between evaluation and positioning is an essential component of the stance triangle, the distinction arguably requires a precision of meaning found in language but not gesture. For the purposes of this discussion, the critical characteristic of evaluation, positioning, and investment is a relationship between speaker and talk.

Subjects relate to other subjects by taking an alignment stance to "[calibrate] the relationship between two stances, and by implication between two stancetakers" (Du Bois, 2007, p. 144). Speakers align themselves in relation to their listener with actions like agreeing,

disagreeing, affirming, and contradicting. Alignment can be expressed explicitly in speech, but is often communicated nonverbally with head nods, shakes, and shrugs. In relation to gestures' functions, shrugs affiliate or disaffiliate to take an aligning stance.

Through this lens, we can think about the many stance-taking meanings of shrug gestures as falling somewhere on a map of two stance axes. In Figure 6.1 below I place stance-taking meanings of shrugs in a space defined by the relationships between speaker and topic and between speaker and listener. In this space, the speaker is centered at the intersection of two axes. Evaluation and positioning are represented on the x-axis. Speakers may move right and toward the topic of talk by investing in their proposition. They move left and away from the topic by disinvesting in their proposition. Clear examples of investing and disinvesting are certainty and ignorance, respectively.
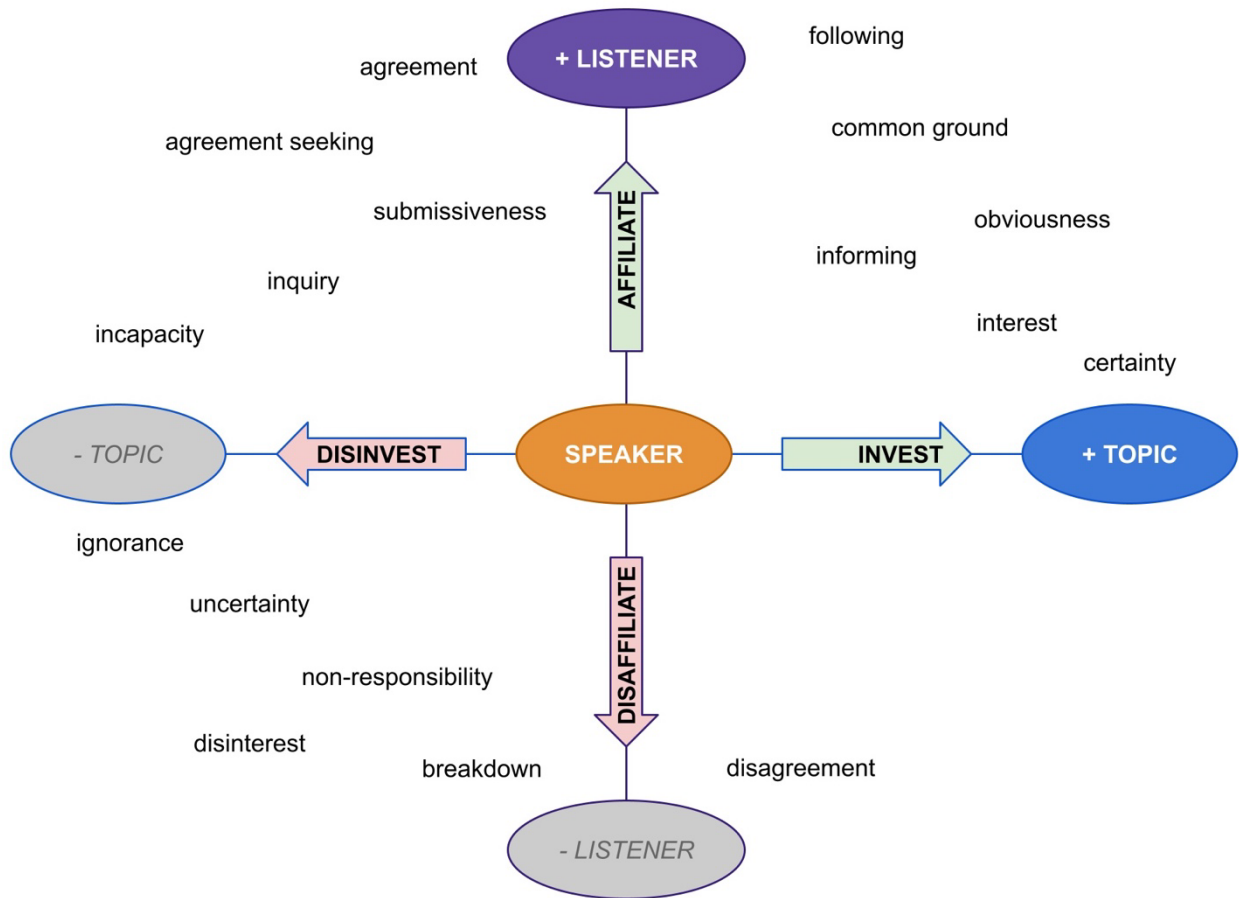
Figure 6.1: Map of shrugs' stance-taking meanings. Speakers establish distance between self and topic or talk by moving along the investment (x) axis. Speakers establish distance between self and listener by moving along the affiliation (y) axis.

Alignment is represented on the y-axis. Speakers may move upwards and toward their listener by affiliating. They move downwards and away from their listener by disaffiliating. Examples of affiliating and disaffiliating are agreeing and disagreeing, respectively.

In the stance triangle model, evaluation, positioning, and alignment are three distinct types of stance, but speakers do not take just one stance at a time. Two museum-goers look at a painting and the first evaluates it, "It's beautiful." The second responds, "I'm in awe." By expressing an internal state of awe, the second speaker has positioned herself in relation to the

169

painting. She has simultaneously implicitly aligned herself in agreement with the first speaker and consequently evaluated the painting as beautiful.

If we conceive of stance-taking meanings in a space like Figure 6.1, we allow shrugs to take multiple stances simultaneously. If a shrug can position the speaker in relation to both listener and talk, it is easier to see connections between the diverse set of meanings. Shrugs can express contradictory meanings when they share an axis. Uncertainty and certainty both relate speaker to talk by expressing some degree of investment. Agreement and disagreement both relate speaker to listener by expressing some degree of affiliation. Shrugs' meanings that are seemingly unrelated can be explained by relative positioning along both axes. The meaning 'submissiveness' may seem unconnected to 'obviousness,' but both meanings affiliate speaker and listener and both express a degree of investment, albeit in opposite directions.

I do not claim that the arrangement of meanings in Figure 6.1 is a hard and fast truth. Relationships between speaker and topic or listener vary in the contexts of an interaction. Arguably some of these meanings operate in entirely different quadrants as often they do as depicted here. For example, in this diagram obviousness is positioned close to both topic and listener. This positioning would be appropriate for taking the stance that a proposition is obvious because the speaker, listener, and third-parties share common ground: "No argument here, we all know this is true." In other cases obviousness might operate close to topic but distant from listener: "How can you say that? Everyone but you knows better!"

The point of situating meanings along these axes is not to definitively say how meanings in either language or gesture take stance. The point is that the meaning labels themselves may not be the most reliable way of discussing these pragmatic functions. Shrugs are stance-taking gestures. They are a speaker's signal to their listener that their relationship to the topic of talk or

the listener is important beyond the talk itself and that it is the listener's task to determine what those relationships are. But it may not be the case that this interpretation involves selecting just the right meaning from more than a dozen possibilities – ignorance, obviousness, submissiveness, incapacity, disinterest, etc. The listener's interpretive work may instead be determining stance along just two axes. In the context of an individual turn sequence, it may not matter that the listener precisely identifies a shrug as expressing uncertainty or disinterest. It may instead matter that the listener recognizes the speaker creating significant distance from the topic and a lesser degree of distance from the listener.

It is important to acknowledge that throughout this work I assign meanings to shrugs. Even in the analyses in Chapter 5 where I collapse meanings into (dis)investment and (dis)affiliation, these stance-taking functions are mutually exclusive. This is unarguably contrary to the claim above that shrugs both invest and affiliate simultaneously. However, my suggestion that these axes of stance may be useful for understanding how speakers and listeners actively coordinate conversation in the moment does not mean that ascribing meaning labels to shrugs is useless. The two axes of investment and affiliation may operate simultaneously, but they are still two separate axes. Not only do they make differentiated contributions to conversation, they do not draw on identical cognitive skills or social inferences. These differences may be particularly informative for the study of pragmatic development, addressing questions about how children build complex systems for taking stance verbally and nonverbally. For example, it is well established that very young children learn to evaluate and express their own internal states before others' (e.g., Bretherton & Beeghly, 1982). We might then expect children to express subject-object relationships before subject-subject relationships, initially use stance-taking resources like shrugs for investment strategies before adding affiliative meanings. Even looking at variations in

meaning along one axis can be informative. Given that children assess their own compete knowledge and ignorance before partial knowledge (e.g., Rohwer et al., 2012; Sodian & Wimmer, 1987), we might find children take stance at the poles of the investment axis before expressing mitigated investment. Children may use shrugs for mitigated investment before language the same way they more accurately express partial knowledge nonverbally (Hübscher et al., 2019; Kim et al., 2016) or later if shrugs take hold in a child's gesture lexicon as an ignorance emblem with relatively little pragmatic flexibility.

6.4 Future Directions

A multimodal model of communicative development includes more than speech and gesture. Characterizing multimodal interaction begins with speech (or the formal systems of a signed language), but includes many signals of the voice and body, both intentionally and unintentionally communicative. It encompasses paralinguistic vocal resources like prosody, pauses, laughter, and voice quality and nonverbal behaviors like co-speech gesture, but also posture, eye gaze, and facial expression. In some perspectives, "multimodal" might also include communicative behaviors like waiting to act or engaging with the physical environment. Although the core questions and findings of this dissertation focus on the systematic integration of speech and co-speech gesture, it is important to remember these are just two facets of multimodality in pragmatic development.

Relatedly, one major limitation of the studies in this dissertation is the restriction of analyses to only communicative acts with gesture. Co-speech gesture is powerfully communicative, but ultimately the impressive flexibility with which adult speakers across the globe use pragmatic gestures like shrugs relies heavily on the way gesture and speech work together to make meaning. Without similar analyses for speech acts *without* gesture, we can only

speculate on what precisely gestures add to conversation that is not already clearly communicated verbally. For example, in Chapter 5 adolescents produced shrugs with dispreferred responses at surprisingly high rates. Stivers et al. (2009) found that in American English 29% of responses to polar questions were dispreferred, compared to 42% for the adolescents in the LDP corpus. Without a comparison to response preference in adolescents' speech-only acts, we cannot rule out explanations for this discrepancy that have nothing to do with gesture. Perhaps adolescents simply give higher rates of dispreferred responses than adults, or perhaps qualities of parent-adolescent dynamics or the academic nature of the task encourage more dispreferred responses.

Although these studies do not compare communicative acts with and without gesture, one strength of the annotation system is its applicability beyond analysis of pragmatic gesture. Future work should leverage the annotation's potential by coding sequence organization properties of speech. This would allow for direct comparisons of speech-only acts, acts with gesture, and gesture-only acts. Such comparisons could rule out alternative possibilities for high rates of dispreferred responses. Additional coding could also explore shrugs' role in response mobilization and answer questions about relative developmental onset of verbal and nonverbal stance-taking strategies.

6.5   Final Thoughts

At the outset of this work my intention was to operationalize forms and functions of children's shrug gestures to provide insight into both the ontology of shrug gestures and their place in ontogeny. As I worked toward these aims it became clear that a straightforward typology of form and meanings based on standard methods from the field of gesture studies would not be sufficient to capture the complex pragmatic and interactional functions these gestures served. It

was easy to see how the literatures on this topic appear inconsistent with one another on so many counts. If "interactive" is a strictly functional classification, so that an interactive gesture may also be a conduit or beat gesture, why does it include the formal criterion of being addressee-indicating? Why are shrugs classically treated as emblems by researchers who do not question that they serve pragmatic and interactive functions? Why do we draw clear-cut divisions of meaning for these "stance-taking gestures" that do not align with existing schemas of discourse stance? I believe these are not methodological flaws or oversights, but rather a reflection of the fact that no single approach can provide a complete and cohesive understanding of how these gestures contribute to conversation or the role they play in multimodal pragmatic development.

These critical questions have reasonable answers, but there is also space to challenge individual approaches to put them in conversation with one another. Why do we impose formal limitations on functional classifications? Why do we break palm-up gestures into presentational, addressee-indicating forms and lateral, outward-directed forms to draw conclusions about interactive and epistemic functions? Because some constraints of form are necessary, at least initially, to recognize generalizable patterns of function. But it need not be the case that these constraints are necessary to move forward once we find a functional system. Once we recognize that addressee-indicating points and palm-up gestures serve specific interactive functions (e.g., Bavelas et al., 1992; 1995) we must ask what gestures fit within this functional framework once formal constraints are lifted. Similarly, once we propose relationships between the epistemic meanings of lateral palm-up gestures (Cooperrider et al., 2018) we must consider whether this map of meanings applies only to the lateral-palm up or to established inter-related forms, like other components of the shrug composite gesture.

Why do we treat shrugs as ignorance emblems when they unequivocally convey a host of other meanings? Because just as unequivocally they also frequently function as non-verbal symbols of ignorance. Once we acknowledge both these points are true (a claim which is not, to my knowledge, in dispute regardless of differences in theoretical or methodological approaches) the next necessary step is to move beyond calling them one thing or another for the purposes of individual studies. Instead we must recognize that these functions are not mutually exclusive, even for a single gesture. If a shrug functions as both a recognizable symbol of ignorance and simultaneously as a tool for managing interaction, we need to consider that these may not be two co-occurring meanings but rather one stance-taking meaning that does not easily conform to our conceptions of one clear-cut meaning or another.

When we disregard the notion that shrugs are simply emblems and instead treat them as stance-taking gestures, why do we then return to assigning meanings rather than functions and to drawing functional divisions that do not reflect current models of stance? This is partially explained by one basic problem: current models of stance are not designed for gesture. Although Du Bois' stance triangle (2007) offers a relatively tidy framework for operationalizing dialogical stance, it is fundamentally about stance in language, leaving non-verbal stance markers as ways of supporting or augmenting stance but not independently taking stance. We cannot smoothly incorporate stance-taking functions of pragmatic gestures like shrugs into this model (or other models, e.g., Cook, 2012; Kockelman, 2004; Ochs, 1996), but that does not mean we need to start from scratch or return to more familiar methods of analysis. Instead we should trust that if these gestures do in fact take stance, there is a model of stance they can operate within. The stance triangle is not irrelevant, just incomplete.

Through my work in this dissertation I have found that these contradictions are not problems to be fixed but puzzles to be solved. Describing the form and meaning of shrugs cannot capture the interactional and relational implementation of these gestures in conversation and consequently cannot wholly describe how they are integrated into processes of discourse-pragmatic development. Pragmatics is the study of language in context, but context is in constant flux across development. The lexical items, phonological rules, and syntactic structures acquired in early childhood exist without context and will remain equally applicable across social situations, across discourse topics, and across cognitive states into adulthood. The "rules" of pragmatics are ever-changing and inherently dependent on not only the more stable properties of language but on each other. The pragmatic functions of gestures in development cannot be treated like stable, context-independent communicative rules. In the first study of this dissertation I argued that pragmatic gestures must be afforded the same attention to detail given to topic gestures. Reflecting on this work as a whole, I argue further that pragmatic gestures must be afforded the same potential for dynamism and interrelatedness given to the linguistic pragmatic resources we see continue to evolve in early childhood, adolescence, and into adulthood.

While the findings from the three studies in this dissertation contribute to a description of gesture's role in pragmatic development, perhaps ultimately the larger contribution is a more general comment on theoretical and methodological approaches to the study of gesture in pragmatic development. Adhering to standard practices – for example assigning meaning glosses – allowed me to show that complex pragmatic functions of shrugs blossom from a more simplistic semantic kernel representation. This adherence also highlighted the limitations of these standards and allowed for initial, minor deviations that begin to give a more comprehensive

picture. For example, although I initially attempted to follow precedent by using separate meanings for epistemic and affective functions, I found it impossible to reliably distinguish between, for example, uncertainty and disinterest, obviousness and commitment, or agreement and understanding. Combining these meanings (into disinvestment, investment, and affiliation, respectively) proved more useful for appreciating the interrelatedness of pragmatic functions. Collapsing further into four stance-taking functions in the final study hinted at what kind of analyses might be even more productive for the integration of these gestures into models of conversational stance. These results illustrated the potential of a modified stance triangle framework, understanding the meaning-making of pragmatic gestures to operate along the two axes of investment and affiliation.

It is intriguing to imagine what more could be gleaned about children's stance-taking were I to use this framework from the start. There was evidence that the shrug enters an American English-speaking child's gesture lexicon as an ignorance emblem, but expressing ignorance is a form of stance-taking. Expressing ignorance is certainly epistemic, but it is connected to concepts of affective and affiliative concepts inability, openness, and submissiveness. Re-examining these data as operating within systems of stance-taking functions rather than as mutually exclusive – though clearly interrelated – meaning glosses might show that non-ignorance meanings do not simply grow vaguely outwards from a kernel. Perhaps these meanings emerge in a recognizable pattern along these axes. From a perspective of early cognitive development and theory of mind, children may first acquire meanings moving from low to high investment and then moving along the affiliative axis before integrating the two together. Alternatively, a perspective beginning from the contributions of pragmatic inference to early word learning might suggest children begin at the axis poles with meanings that take

relatively unambiguous stance – ignorance, obviousness, agreement, disagreement – and move inward toward more nuanced and context-dependent stance meanings.

It often feels ironic to take shrugs so seriously when shrugs are so often a caricature of apathy. But the very fact that they do not always symbolize apathy is what makes them worth taking seriously. How is it that this gesture can be so emblematic as to have its own emoji and at the same time be so functionally flexible that even the emoji is often only interpretable in context? Shrugs are gestures of contradiction. They agree; they disagree. They say, "I don't know"; they say "of course!" They take the turn; they pass the turn. They are highly salient; they are produced without the speaker even noticing. They are a nonverbal, conventionalized lexical symbol; they are a flexible and universal index of interactional context. They are one of the most commonly produced gestures for both infants and adults; they function so differently in preverbal speech and adult dialogue that they might as well be entirely different gestures.

Shrugging may feel trivial, but it simply is not. It is not trivial that infants can map multiple component forms to a single meaning while they struggle with synonymy in speech. It is not trivial that children's gestures give a more accurate impression of their cognitive capacities for assessing epistemic states. It is not trivial that the same gesture forms used emblematically in early childhood are used with astonishing pragmatic and interactive flexibility in adulthood, and it is not trivial that it takes at least a decade for this transformation to occur. I argue that the role of shrugs in pragmatic development is not trivial, but truly remarkable. I mean obviously! This is interesting, don't you agree? Well whatever, it doesn't matter. I have nothing else to say. 🤷

Appendix A: Interaction Coding Manual

Additional annotation notes, clarifications, and examples are available upon request. The codes *align* and *disalign* were originally used following Du Bois' terminology but were replaced with *affiliate* and *disaffiliate* once coding was complete for internal consistency.

A.1 Pragmatic Gesture

A.1.1 Inclusion

1. Code gestures that serve interactive functions.
2. Do not code deictic, emblematic, or representational (iconic, metaphoric) gestures unless they take interactive forms.
3. Code gestures produced during spontaneous interaction. Do not code gestures produced while a speaker is reading directly from a prompt.

Code	Description	Other remarks
FORM		
nod	head up/down	
shake	head side to side	
tilt	head tilt to shoulder	does not include nods or shakes tilted slightly
pug	palm-up gesture	also code canonicity; one or both palms
shrug	shoulder raise	one or both shoulders
point	extended index finger	for points with interactive functions only; addressee-indicating
beat	punctuated or rhythmic movement	
other	other forms with interactive meanings	usually forms similar to palm-ups
GLOSS		
invest	commitment to proposition, obviousness, certainty	I'm telling you, obviously, clearly, I'm sure, I feel strongly about this, as we all know, of course
disinvest_0	ignorance, inability	I don't know, I can't answer that, how should I know?, can you help me?
disinvest_1	uncertainty, unwillingness; distance, lack of commitment, non-responsibility, disinterest	I'm not sure, I think so, I guess so, maybe, I'm no expert, if I had to guess, I don't care, whatever, it doesn't matter
disinvest_2	absence, completion	all gone, all done
align	agreement, acknowledgment, following	ok, me too, I understand, I agree, you're right, I'm following, keep going, we're on the same page
disalign	disagreement, confusion, not following	I disagree, I don't like what you're saying, I'm confused, you're wrong, I'm not following you
affirm	affirmation, confirmation	yes, correct
negate	negation, disconfirmation	no, incorrect

Code	Description	Other remarks
turn_0	claiming turn; negotiate turn taking	I'll talk next, I'm not done yet, don't interrupt, let me interrupt
turn_1	ceding turn; negotiate turn taking	go ahead, what do you think?, I'm done talking, feel free to interrupt me
exclude	pragmatic forms taking non-pragmatic meaning	e.g., pugs metaphorically representing contrast ("on the one hand..."); pugs and head nods that are unconventional points/deictics

RELATION

add	add meaning not present in text or audio	aka supplement
rf	reinforce meaning found in either text or audio	
da	disambiguate unclear or incomplete meaning in speech	aka complement; typically deictics DA pronouns, rare but possible for shrugs
sub	substitute gesture for speech	communicative act is gesture only or "no-speech"
con	contradict meaning in speech	e.g. adding investment to an utterance that contains clear hedging or words indicating uncertainty

RM
l	lexico-syntactic	uncertainty/ignorance epistemic state words (think, believe, bet); hedges and fillers (um, well); question syntax; directing turn transition "you go now"
p	prosodic	try-marking
f	facial	mouth/eyebrow shrugs

PALM-UP FORM

np	non-canonical presentational	forward, addressee-indicating
cp	canonical presentational	
nl	non-canonical lateral	
cl	canonical lateral	
no	non-canonical other; unclear	may point to something other than addressee; may be angled between addressee and lateral
co	canonical other	theoretically possible but not expected to appear

A.1.2 Gesture Form

These gesture forms may serve interactive functions. Do not code deictic, emblematic, or representational (iconic, metaphoric) gestures unless they take these forms *or* they have a clear meaning from among the interaction gloss codes.

- nod: vertical head movement
    - sharp single downward movement of the chin OR both an upward and a downward movement
- shake: horizontal head movement
- tilt: head tilts toward one shoulder
    - code head tilts only when they have communicative intent; do not code head tilts with a functional purpose, e.g., comfort, looking at interlocutor, fidgeting, etc.
    - must have clear start and stop

- PUG: palm-up gesture
  - canonically, an outward rotation of the wrist so the open palm faces upward and fingers extend away from the body either laterally in neutral space or forward toward the addressee
  - more commonly produced non-canonically as "reduced" or "incomplete" rotation/extension
- shrug: one or two shoulders raised toward the ears
- point: one finger indicating the addressee when serving an interactive function
  - conceptually similar to an extremely reduced PUG
  - most pointing gestures are *not* interactive!
- other: gestures that seem to perform interactive functions but do not take any of the above forms
  - unlike referential gestures (points, iconics, emblems, etc.), the meanings of these gestures clearly match the interaction gloss codes
  - often may be similar to a PUG, but missing an essential element; for example there is not wrist rotation, the palm faces down, the palm/fingers do not open, or the hand is holding an object
  - includes gestures that take the form of a point to the addressee when serving an interactive (rather than deictic) function; similar to extremely reduced presentational PUG
- beat: punctuated or rhythmic movement
  - must be combined with one of the above forms
  - produced to emphasize a word or syllable, with at least *some* prosodic prominence on that word/syllable as well
- unclear: gesture is generally recognizable as one of the above forms but not clear enough to make informed judgments about
  - camera angle makes it impossible to see the whole gesture
  - an object or person blocks the camera's view temporarily

A.1.2.1 Form notes
1. Forms can be combined with "+", but do not include more than one head movement in a gesture form. Code the most prominent head movement.
2. Do not code beat-like gestures produced while reading that serve to follow text, such as rhythmically pointing to each word on the prompt.

A.1.3 Gloss
- invest: commitment to the proposition
  - "here's what I'm telling you"
  - "obviously"
  - "clearly"
  - "I'm sure"
  - "I feel strongly about this"
  - "as we all know"
  - "of course"
  - "here's a clear example"
  - "well duh"

- ▪ "have to" / "must be"
- disinvest: distancing from the proposition
  - ▪ disinvest_0: ignorance, non-response, inability
    - ○ "I don't know"
    - ○ "I can't answer that"
    - ○ "I have nothing to say"
    - ○ "can you help me?" (inability)
  - ▪ disinvest_1: uncertainty, non-responsibility
    - ○ "I'm not sure"
    - ○ "I think/guess so"
    - ○ "maybe"
    - ○ "I'm not an expert"
    - ○ "if I had to guess"
    - ○ "I don't care"
    - ○ "whatever"
    - ○ "it doesn't matter"
    - ○ "not my fault"
    - ○ "I can't help it"
    - ○ "oh well"
  - ▪ disinvest_2: literal absence
    - ○ "all gone"
    - ○ "all done"
- align: agreeing or affiliating with someone (typically the addressee, can be a non-present topic of talk)
  - ▪ "I agree"
  - ▪ "me too"
  - ▪ "I understand"
  - ▪ "you're right"
  - ▪ "I'm following"
  - ▪ "good idea"
  - ▪ "fine with me"
  - ▪ "we're on the same page"
  - ▪ "you know the rest"
  - ▪ "that's a good point"
  - ▪ "can I?" (ask permission)
  - ▪ "you can" (grant permission)
  - ▪ "know what?", "guess what!" (seek engagement)
    - ○ "what?" (offer engagement when sought)
- disalign: disagreeing or disaffiliating with someone (typically the addressee, can be a non-present topic of talk)
  - ▪ "I disagree"
  - ▪ "I don't like what you're saying"
  - ▪ "I'm confused by what you're saying/asking"
  - ▪ "you're wrong"
  - ▪ "I'm not following you"
  - ▪ "I misunderstood you"

- "bad idea"
- "on the other hand" (i.e. disaligning with self)
- "you can't" (refuse permission)
- affirm: affirming/confirming <u>the proposition</u>
  - "yes"
  - "correct"
  - can be "no" when confirming a negative proposition, e.g. "you didn't notice that?" "no."
- negate: negating/denying <u>the proposition</u>
  - "no"
  - "incorrect"
  - "not so"
  - can be "yes" when negative a negative proposition, e.g. "you didn't notice that?" "yeah I did notice."
- turn: negotiating turn taking
  - turn_0: claiming the floor
    - "I'll talk next"
    - "I'm not done talking yet"
    - "don't interrupt"
    - "let me interrupt"
  - turn_1: ceding the floor
    - "go ahead"
    - "what do you think?"
    - "I'm done talking"
    - "feel free to interrupt"
- other: non-interactional meaning for a target form

A.1.4 Gesture-Speech Relation

How does this gesture contribute meaning to the utterance? Consider the utterance to be the full speech signal, including syntax, word choice, and prosody. Judge relation based on the gesture gloss.

- add: provides new, additional meaning not present in the utterance; aka "supplement"
  - includes mid-act or act-final gestures that occur during pauses
  - gesture is supplemental; act is comprehensible without gesture
- rf: reinforces meaning already present in the utterance
- da: disambiguates meaning in the utterance; aka "complement"
  - gesture is necessary for full comprehension
  - this is very rarely the case for interactional gestures!
    - occasionally occurs when the gesture serves as a "word" mid-utterance, e.g. "she asked and I was like -- [shake head] -- because that's gross."
    - can occur when giving directives, disambiguating the intended recipient (the spoken or implied "you")
- con: contradicts meaning present in the speech

- e.g. nod adding *investment* to an utterance containing hedges or cognitive state words that suggest *disinvestment* "I mean, yeah, because they probably want them to learn --"
- sub: substitutes for speech *entirely*; the gesture is the entire "utterance"
  - Examples of what is <u>NOT</u> substitution:
    - gestures produced with "mmhm" or "mmMMmm" (prosodic "I don't know") as speech
    - gestures that *immediately* precede or introduce speech without a pause
    - gestures that hold the floor mid-utterance
    - gestures that *immediately* follow an utterance, including when the speech trails off and the gesture concludes the speech act

A.1.5 Palm-up Form Completeness

For gestures with a PUG component, determine whether the gesture is canonical ("complete") or non-canonical ("reduced") and presentational or lateral. All gestures should take the form of a two-letter code. Letter 1 indicates canonical/non-canonical PUG form. Letter 2 indicates presentational/lateral direction.

- PUG form (c/n)
  - c: canonical/complete PUG
    - one or both wrists rotate outward
    - palm faces upward
    - at least two fingers fully extended exposing palm
    - movement "freezes" (may be extremely brief, but perceivable at full speed video)
  - n: non-canonical/reduced PUG
    - outward rotation is small or incomplete
    - palm may face outwards rather than upwards
    - one or more fingers stretch outward but may not fully extend
    - handshape may be more similar to a point than a palm
- Direction: (p/l/o)
  - p: presentational
    - extended fingers point toward addressee
  - l: lateral
    - extended fingers point outward in neutral space
    - may coincidentally point toward an object or person in the space but is not a deictic reference to it
    - may be more "neutral" than lateral, especially for very small movements
  - o: other (any of the following)
    - direction is unclear
    - points to *and indicates* an object or person other than the addressee
    - movement is so small the direction cannot be determined
    - hand is resting on the body or another surface and does not take a clear direction
    - extended fingers point directly forward but not towards the addressee

A.1.6 Response Mobilization

What response mobilizing features are present in the communicative act? This does not include all RM features identified by Stivers and Rossano (2010). Use multiple codes if applicable, e.g. *lpf* where all three features are present.

- l: lexico-syntactic
  - interrogative syntax
  - epistemic uncertainty words: think, believe, guess, don't know, etc.
    - *except* when explicitly discussing opinions, beliefs, cognitive states; e.g., "I think Joe is worse for stealing" in response to "who do you think is worse?"
  - hedges, fillers, & modals: maybe, possibly, might be, could be, well, um, uh
- p: prosody
  - try-marking, rising intonation with declarative statements
  - elongation, e.g. "welllll"
  - prosodic emphasis on epistemics, hedges, modals
  - extended pauses
- f: facial expression
  - eyes rolling or wide
  - mouth, nose, and brow scrunching/shrugging
  - tight or thin lips

A.2 Sequence Organization

Code	Label	Description
POSITION		
1	first-pair part	receives a response
2	second-pair part	includes responding to questions in the prompt that are read silently or read aloud by the same speaker
3	no-transition	act neither offers nor receives a response
4	backchannel	feedback of acknowledgment, understanding, or agreement that does not take the turn
0	unclear	
9	scripted exception	experimenter not free to respond naturalistically; addressee is not transcribed; speaker is reading from task prompt
REQUEST		
1	closed request	turn-initial act uses interrogative syntax or rising intonation and has given or implied set of responses; "which of these" questions; affirming questions
2	open request	turn-initial act uses interrogative syntax; most wh-questions
3	imperative	turn-initial act issues a command or directive with imperative syntax
NR	non-request sequence	sequences that do not meet any of the above criteria are not coded for request type
INTENT		
1	scenario	comments on behavior in the immediate scenario, announces next action, narrates pretend play, does not present fact or opinion
2	objective	presents (dis)provable claim, objective facts, may be correct or incorrect
3	subjective	presents (dis)agreeable claim, subjective opinions, assessment, announcements of cognitive or emotional state
4	rote	ritualized interaction (rare)
PREFERENCE		
1	preferred	SPP or backchannel in preferred format; affirms, acknowledges, agrees, accepts
2	dispreferred	SPP or backchannel in dispreferred format (other than ignorance); negates, ignores, disagrees, rejects
3	ignorance	SPP or backchannel giving dispreferred response, ignorance as non-answer but not answers with marked uncertainty
ACT FEATURES		
repair	sequence attempts to correct a problem with hearing or understanding	code 1 if present, else blank; self- or other-initiated
interact	sequence directs interaction, manages turn, offers backchannel feedback	code 1 if present, else blank; includes all acts coded as position 4
report	utterance contains reported speech, direct or indirect quotes	code 1 if present, else blank; communicative intent and format of speaker's utts may not match intention of the quoted speech (e.g. "she said 'are you going?' - reported speech is a request but C's utterance is not); character embodiment

A.2.1 Inclusion

1. Early childhood: Annotate sequence organization for children's shrug and palm-up ("flip") gestures already, with or without speech. Annotation is at the level of the communicative act; utterances with multiple gestures are considered one communicative act.
2. Early adolescence: Annotate sequence organization for all children's acts that were coded for pragmatic gesture.

A.2.2 Position

All target gestures are coded for sequential positioning. Use numeric codes rather than descriptive names. An utterance may be in the first (1) or second (2) position of an adjacency pair, prompting a response from the listener or providing a response. Utterances that are not responses and do not receive a response are coded outside of a turn transition sequence (3). Addressee backchannel feedback does not interrupt the turn sequence (4). Unintelligible speech and other utterances where meaning is indecipherable from the transcribed context are coded as unclear (0). In experimenter-child sessions, it is impossible to determine the "true" positioning of an act due to the artificial limitations of the experimenter's script (9).

- 1: first-pair part (FPP)
    - non-contingent
    - initiates a turn sequence AND receives a response
    - may or may not directly request a response
- 2: second-pair part (SPP)
    - contingent
    - act closes a turn sequence AND offers a response
- 3: no transition
    - non-contingent
    - does not receive or offer a response
    - may be a failure to mobilize response
- 4: backchannel feedback
    - offer contingent feedback but do not take the conversational form
    - gives listener feedback of (mis)understanding or (dis)agreement without taking the turn

- often accompanied by "yeah" or communicative sounds like "mmhm" or "uhhuh"
- 9: scripted exception (must not meet any of above coding criteria)
  - followed by experimenter's scripted response
  - impossible to tell if the act would have received a response in natural conversation without the experimenter's artificial constraints
  - the child is reading the prompt aloud *AND* does not receive a response from the addressee
- 0: unclear (any of the following)
  - unintelligible speech
  - camera or other technical issues

A.2.2.1 Position notes

1. Scripted exceptions are acts which should not be coded due to format constraints of visit or transcript protocol. For example, when data collectors are instructed to remain silent (as in early childhood visits) or follow a script (as in adolescence visits) they may not respond naturalistically to children's acts directed toward them. In such cases it is "impossible to know" whether the child's utterance would have elicited a response in unrestricted interactions, and the act should be excluded from analyses
2. Acts that provide a satisfactory response to a turn-initial request are second-pair parts whether or not the act then receives a follow-up/third-position response.
3. Prioritize first-pair parts for serial requests. Responding to a question with another question is not a relevant or satisfactory response.
4. Turn transitions may span more than two utterances.
5. In cases where speakers initiate a turn sequence but do not clearly provide the opportunity for a response, the distinction between FPP and no-turn-transition is dependent on whether the addressee attempts a response, indicating an attempt to complete the turn.

A.2.3 Request Form

Code adjacency pairs that request information, confirmation, or action; do not code non-request sequences that do not meet the criteria below. Code both the first pair parts (the requests/questions) and the second pair parts (the responses/answers). These categories are mutually exclusive. Requests that are ambiguous in form (can be interpreted as either open or closed) are determined by the response.

- 1: closed
  - FPP initiates a closed question; interrogative syntax or intonation
  - yes/no questions
  - multiple choice questions
  - does not require word search to answer

- 2: open
    - FPP initiates an open question; interrogative syntax
    - *wh*-questions
    - requires word search to answer
- 3: imperative
    - FPP issues a directive; imperative syntax
    - commands, offers, instructions

A.2.3.1 Request form notes

1. Always code based on request form if possible. When form is ambiguous (e.g., "Do you remember what that's called?" takes closed syntax but seeks an open-ended response) code based on addressee's response.
2. For the phrases "how about" and "what about", use the listener's response to determine whether the request was looking for some kind of confirmation (closed) or description (open).
3. Do not treat acts that follow requests but do not provide a contingent response as part of the request sequence.
4. Request-response pairs may span more than two utterances.

A.2.4 Intended Effect

Intended communicative effect of an act is coded as commenting on the current behavior or scenario, objective/observable information, or subjective/unobservable information.

- 1: scenario
    - comments on the immediate scenario
        - does not include comments about the physical environment (e.g. locating or describing present objects) - these are *objective* propositions that can be disputed/disproven
    - announce actions, decisions, wants, and needs regarding immediate context (e.g., "I want to go upstairs")
    - ask or grant permission
    - make, grant, or refuse requests and offers
    - narrate ongoing events
    - (meta)comments on the immediate interaction
- 2: objective
    - assertions of objective or observable information
    - discussion of locations, visual descriptions, facts about the world
    - open to epistemic positioning
    - disputable, disprovable
    - responses to test questions
    - reasoning about objective, observable, or mechanical causes
- 3: subjective
    - assessments and evaluations

- statements of opinion or belief
        - open to (dis)agreement but not (dis)provable
        - open to (dis)affiliative alignment
        - wants and needs outside immediate context (e.g., "I want to be an astronaut")
        - reasoning about subjective, unobservable causes; personal motivations
- 4: rote
    - spelling, counting
    - labeling
    - social scripts (e.g. apologies, greetings, gratitude)
    - sound effects during play (e.g. "choo choo", "meow")
        - exception: sounds/onomatopoeia produced in response to objective questions (e.g. "what sound does a cat make?" "meow."

A.2.4.1 Intended effect notes

1. Discussing wants and needs can have either scenario or subjective intent, depending on whether fulfilling the want/need would affect the immediate context. Rarely, wants/needs may have objective intent if they describe something inanimate or incapable of subjectivity.
2. Simple exclamations (e.g. wow, uhoh, oh no, yay) may have subjective or scenario intent, depending on whether they comment on something in the immediate context or a decontextualized/abstract topic.
3. Speakers may use objective statements to support larger subjective arguments. Code at the level of the turn construction unit, taking into consideration how the listener interprets the intent and responds.

A.2.5 Preference

Only SPP and backchannel acts are coded for preference.

- 1: preferred (all of the following):
    - response takes relevant form
    - confirms, agrees, accepts
- 2: dispreferred (any of the following):
    - response does not take relevant form
    - non-response (except ignorance)
    - disconfirms, disagrees, refuses
    - inability to provide a preferred response *not due to lack of knowledge*
- 3: ignorance:
    - special case of dispreferred response
    - inability to provide a preferred response *due to lack of knowledge*

A.2.5.1 Preference notes

1. Code for preferred *format* rather than preferred *action*. For example, although a speaker may hope their addressee declines their polite offer, an acceptance is still the preferred response format.
2. Basic preference principles for American English (Pomerantz & Heritage, 2012) are listed here in a plausible order of importance, with rules that are strictly about *form* being most important and rules that take into account preferred *action* being less important.
   a. Provide a relevant response, i.e. appropriate response type/form.
   b. Confirm, agree, and accept.
   c. Be truthful.
   d. Avoid self-praise.
   e. Avoid burdening others.
   f. The selected next speaker should respond.

A.2.6 Other Codes

These codes are not mutually exclusive. Code each with a 1 if the feature is present, otherwise leave empty. These codes are intended to be potential post-hoc filters to catch likely coding disagreements and complicated or interesting cases to come back to in a later analysis.

- REPAIR: indicates problems with hearing or understanding
  - includes upgraded backchannels; speaker did not have a problem with hearing or understanding but initiated a "repair" to show interest or encourage the interlocutor to continue
  - does not include self-repair
- INTERACT: serves *only* to direct the conversation
  - attention-getters: "see?" "Mom!"
  - pre-expansions and standardized prompts: "guess what./what?" "oh hey -- / yeah?"
  - backchannel feedback (including upgraded backchannels also coded as repairs)
  - "meaningless" and affective exclamations: "oh man!" "wow!" "oh no!"
  - sequence closings and minimal expansions: "well ok then"
  - performative acts: "sorry" "bless you" "thanks"
    - note: performative acts or social scripts are usually coded as intent 4 (rote speech) in addition to being marked as interact
- REPORT: reported speech
  - directly or indirectly quotes another speaker
  - quoted speaker may be present, absent, or imaginary/hypothetical
  - may be the entire utterance or contained within a larger utterance

A.3 Early Childhood Session Notes

Code using context. These codes are based on *interaction* rather than singular communicative acts. In most cases, you should use the surrounding rows of transcript to make coding decisions rather than sticking to hard and fast rules based on just the gesture or utterance.

A.3.1 Exclusion criteria

1. *Insufficient context.* Neither the co-produced utterance or the surrounding contextual utterances provide enough information to make reasonable assumptions about the gesture, typically due to either unintelligible speech or conversation with an interlocutor who is not transcribed and/or unable to freely participate in the interaction.
2. *Single gestures represented in multiple utterances.* Transcription conventions dictate that single gestures formed or held over multiple utterances be coded for all relevant utterances and noted as continuations. Code only the first appearance of these gestures and treat the following utterances as particularly relevant context for coding decisions.
3. *Non-interactive pointing.* All *potential* interactive gestures are pulled for coding, including all deictic (points, palms) gestures indicating the addressee (PCG). In most cases points to PCG are not interactive. Only code these pointing gestures if they serve clear interactive functions (Bavelas et al., 1992). Add a 1 to the *skip* column to indicate the act has been intentionally excluded.
4. *Response mobilization coding is limited.* It is not possible to code facial expressions without access to the video, so the RM column should never contain an "f" code. Prosodic RM should be based on the transcriber's decision to use a question mark.
5. *Ignorance of location vs absence.* "Where" questions primarily for missing information. When co-produced speech is 1) present and 2) structured as an interrogative, this communicates ignorance and should be coded as disinvest_0. When a gesture 1) responds to a "where" question and 2) does not explicitly contain "all gone" speech, this is an ignorance response. Only utterances/gestures that implicitly or explicitly declare "all gone" should be coded as disinvest_2 (i.e. the speech is not a question or there is no co-produced speech). Acts coded as literal absence are necessarily commenting on the here-and-now and should be coded as scenario communicative intent.

A.4 Non-referential Gesture Pragmatic Function

Code pragmatic function of both speech and gesture (separately) for all child gestures with the initial form codes flip, beat, iconic, shake, and nod. Coding occurs at the level of the gesture, so that single utterances containing multiple co-speech gestures have multiple pragmatic function codes. Use surrounding transcript context to make judgments about communicative

intent. In cases in which there is insufficient context to code the pragmatic function of an utterance or gestures; for example, when the child is talking to the experimenter or someone else who is not transcribed, we will use the "unclear" code.

Appendix B: Session Format and Exclusion Criteria

B.1 Early Adolescence Session Format

During each of two home visits, following 5th and 7th grade, adolescents and caregivers were video-recorded completing 12 domain-based tasks. Tasks were limited to a maximum of 5 minutes each so that maximum session length was one hour. The experimenter sat nearby to issue directions, distribute and collect prompts and manipulables, and ensure tasks did not exceed the time limit. Before beginning the session, dyads were given the following written general instructions which the experimenter also read aloud:

> *"This project is about how you and your child interact during activities and puzzles. We're not at all concerned about how much you know about the activities or the answers you give. We are really interested in how you talk about the activities and how you collaborate to work on them. Please don't worry if you don't finish any particular activity or puzzle or if you have difficulty with any of them."*

B.1.2 Tasks Annotated For Gesture And Sequence Organization

Tasks were designed to elicit higher-order thinking talk in four semi-academic domains: math, natural world, social world, and art. At the beginning of each task, dyads were presented with written task-specific instructions. Instructions were not read aloud by the experimenter. Dyads were free to end each task when they felt they had completed it, with a 5-minute time limit that was not explicitly stated to dyads. Upon reaching the 5-minute mark, the experimenter would collect materials and written instructions and say "I'm sorry, we have to move on. You can come back to this at the end if you want." Experimenters generally did not answer questions about each task but did provide definitions of unfamiliar words and occasionally quoted the

written prompt as clarification, otherwise they responded with a variation of "I'm not sure about that."

For each session, three tasks were selected to be annotated for pragmatic gesture and sequence organization. The six tasks included five from the social world domain and one from the art domain. None of these tasks included manipulable objects, leaving participants free to gesture. Written instructions for these tasks are as follows:

1. People often debate how we can best educate kids in this country. Some people think that all high school students should take exactly the same classes. Other people think that high school students should be allowed to choose all their own classes. Discuss with each other what you think. Should high school students be allowed to choose their own classes? What are the pros and cons to each of these positions?

2. Listen to the following two clips of music by Billy Joel and Beethoven. Once you have heard both songs, please discuss any similarities and differences you noticed between the two pieces of music.

3. Two young men, brothers, had gotten into serious trouble. They were secretly leaving town in a hurry and needed money. Alex broke into a store and stole $500. Joe went to an old man who was known to help people in town. Joe told the man that he was very sick and he needed $500 to pay for an operation. Really Joe wasn't sick at all, and he had no intention of paying the man back. Although the man didn't know Joe very well, he loaned him the money. So Joe and Alex skipped town, each with $500. Everyone knows that what both Alex and Joe did was wrong, but if you had to make a choice, would you say Alex was worse to break in the store and steal the $500 or Joe was worse to borrow the $500 with no intention of paying it back? Discuss which one you think is worse and why (you don't have to agree). Which would make you feel worse, stealing like Alex or cheating like Joe?

4. Every morning, Mike takes the public bus to school. It is always very crowded with other people on their way to work and school. Today, Mike is very relieved because he found a seat before the bus got too full. This was especially nice because in addition to his heavy backpack, Mike is carrying a large science fair project poster. Mike notices an old man with a cane and a pregnant woman both standing. Do you think someone on the bus should give up their seat to let these people sit down? Should Mike give up his seat on the bus? If there is only one seat, who do you think should be allowed to sit down?

195

5. Everyone knows that inventions change the way we live. For example, cell phones have become an important part of our lives. However, just 20 years ago, most people did not have cell phones. Instead, they could only use phones in certain places, and people could not send each other text messages. Talk to each other about how the invention of cell phones has changed people's lives. How was life different before cell phones? How would your life be different if cell phones didn't exist?

6. There are two 7th grade boys who go to the same school, Jerome and Dominic. Jerome's parents give him a weekly allowance regardless of his behavior, because his parents believe an allowance teaches Jerome to be good with money. Dominic's parents only give him an allowance as a reward for doing chores or earning good grades. Which plan do you agree with? How should an allowance be given?

B.1.3 Home Visit Exclusions

Home visits were excluded from the analyses in Chapter 4 if there was very low parent speech. In these sessions (N=9) the primary caregiver typically was not present for most or all of the session, leaving the child to talk with other familiar adults whose speech was not transcribed or to the silent experimenter. In these cases transcripts provided only the child's half of the interaction, leaving insufficient context to reliably annotate pragmatic gesture or sequence organization. Sessions were excluded if a log transformation of the ratio of parent utterances to child utterances was less than -2.
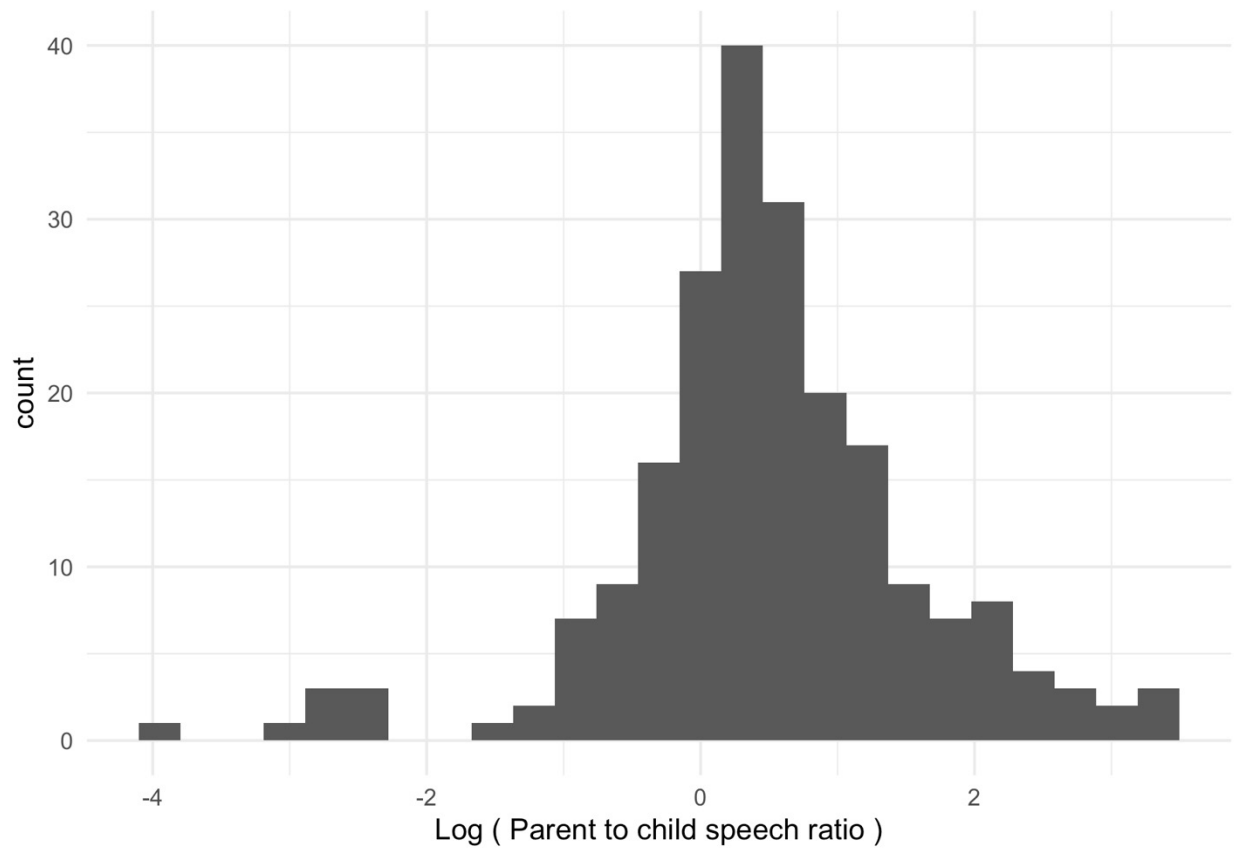
Appendix Figure 1: Log transformation of parent-to-child speech ratio for early childhood visits. Outlier sessions with a log-ratio less than -2 were excluded from analyses.

Appendix C: Reliability Metrics

C.1 General Reliability Procedures

For each round of annotation, a selection of transcripts from each session were randomly chosen to be double-coded in full. Disagreements were resolved by coder consensus. Percent agreement was established for all annotation. Cohen's kappa was calculated when possible using the *irr* R-package. Reliability metrics are not reported for annotation that was completed but not used in any analyses in this dissertation.

Appendix Table 1: Reliability metrics

Annotation system	Double-coded transcripts	Early childhood		Early adolescence	
		% Agreement	Cohen's κ	% Agreement	Cohen's κ
Pragmatic function, speech acts	20%	89.0%	0.846	*n/a*	*n/a*
Pragmatic function, gesture (including no-speech acts)	20%	87.1%	0.819	*n/a*	*n/a*
Palm-up completeness	25% (EC); 33% (EA)	85.0%	*n/a*	92.1%	0.751
Shrug meaning	33%	82.1%	0.771	80.0%	0.721
Sequential position	33%	88.9%	0.820	83.3%	0.739
Request type	33%	91.3%	0.873	87.0%	0.760
Preference format	33%	84.1%	0.758	77.8%	0.679

Appendix D: Statistical Models and Tables

D.1 Chapter 3 Statistics

Appendix Table 2: ANOVA, gesture construction onset

Group 1	Group 2	n1	n2	Statistic	df	*p*	*p*-adj	Adj. sig.
Co-speech BEAT	Co-speech PALM-UP	18	18	3.817	17	0.001	0.004	**
Co-speech BEAT	No-speech PALM-UP	18	18	2.899	17	0.010	0.030	*
Co-speech PALM-UP	No-speech PALM-UP	18	18	0.170	17	0.867	1.000	*ns*

Appendix Table 3: GLMM Model 1 (without age), Poisson; gesture frequencies predicted by gesture construction and communicative acts

	Estimate	Std. error	z value	Pr(>|Z|)
(Intercept)	-0.870	0.290	-3.001	0.003
gtypeCo-speech palm-up	0.974	0.146	6.655	0.000
gtypeNo-speech palm-up	-0.047	0.169	-0.277	0.781
gtypeCo-speech beat:n_acts_c	0.965	0.108	8.942	0.000
gtypeCo-speech palm-up:n_acts_c	0.686	0.078	8.821	0.000
gtypeNo-speech palm-up:n_acts_c	-0.173	0.123	-1.403	0.161

n_gestures ~ gtype + gtype:n_acts_c + (1|subject)

Appendix Table 4: GLMM Model 2 (with age), Poisson; gesture frequencies predicted by gesture construction, communicative acts, and child age

	Estimate	Std. error	z value	Pr(>|Z|)
(Intercept)	-4.123	0.614	-6.720	0.000
gtypeCo-speech palm-up	2.170	0.450	4.818	0.000
gtypeNo-speech palm-up	2.149	0.534	4.023	0.000
gtypeCo-speech beat:months	0.068	0.012	5.882	0.000
gtypeCo-speech palm-up:months	0.042	0.010	4.279	0.000
gtypeNo-speech palm-up:months	0.016	0.013	1.250	0.211
gtypeCo-speech beat:n_acts_c	0.958	0.115	8.358	0.000
gtypeCo-speech palm-up:n_acts_c	0.680	0.078	8.765	0.000
gtypeNo-speech palm-up:n_acts_c	-0.007	0.131	-0.054	0.957

n_gestures ~ gtype + gtype:months + gtype:n_acts_c + (months|subject)

Appendix Table 5: ANOVA model comparison, with and without age predictor

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
Model 1 (-age)	9	1523.317	1563.540	-752.6583	1505.317			
Model 2 (+age)	12	1495.685	1549.316	-735.8426	1471.685	33.63147	3	<.0001

Appendix Table 6: Post-hoc Chi-square residuals, construction and pragmatic function

Dimension	Epistemic agreement	Epistemic uncertainty	Expressive speech act	Negation	Requesting speech act	Unbiased assertion
Co-speech beat	-0.435	-7.562***	-2.872	0.663	1.039	7.881***
Co-speech palm-up	0.640	-0.136	0.459	1.423	1.158	-1.806
No-speech palm-up	-0.343	9.809***	3.011*	-2.835	-2.941	-7.499***

* p<0.05, **p<0.01, ***p<0.001

Appendix Table 7: Multinomial logistic regression, form/function over time

	Term	Estimate	Std.error	Statistic	P.value
Epistemic uncertainty	Co-speech beat	3.407	3.189	1.068	0.285
	Co-speech palm-up	3.621	1.025	3.533	0.000
	No-speech palm-up	7.510	3.773	1.990	0.047
	months_0	-0.077	0.094	-0.814	0.416
	Co-speech palm-up:months_0	0.024	0.100	0.238	0.812
	No-speech palm-up:months_0	-0.052	0.136	-0.385	0.700
Expressive speech act	Co-speech beat	5.667	3.303	1.716	0.086
	Co-speech palm-up	2.061	1.114	1.850	0.064
	No-speech palm-up	5.735	3.801	1.509	0.131
	months_0	-0.192	0.105	-1.828	0.068
	Co-speech palm-up:months_0	0.150	0.111	1.354	0.176
	No-speech palm-up:months_0	0.073	0.145	0.503	0.615
Negation	Co-speech beat	3.118	3.118	1.000	0.317
	Co-speech palm-up	1.531	1.129	1.355	0.175
	No-speech palm-up	-1.829	0.826	-2.213	0.027
	months_0	-0.057	0.091	-0.621	0.535
	Co-speech palm-up:months_0	0.038	0.098	0.388	0.698
	No-speech palm-up:months_0	-0.284	7.136	-0.040	0.968
Requesting speech act	Co-speech beat	-0.318	3.228	-0.099	0.922
	Co-speech palm-up	1.079	1.150	0.939	0.348
	No-speech palm-up	-1.812	2.471	-0.734	0.463
	months_0	0.049	0.092	0.529	0.597
	Co-speech palm-up:months_0	-0.051	0.099	-0.518	0.604
	No-speech palm-up:months_0	-0.504	12.870	-0.039	0.969
Unbiased assertion	Co-speech beat	4.161	2.872	1.449	0.147
	Co-speech palm-up	1.475	1.051	1.404	0.160
	No-speech palm-up	5.183	3.879	1.336	0.181
	months_0	-0.028	0.083	-0.331	0.741
	Co-speech palm-up:months_0	0.055	0.089	0.612	0.540
	No-speech palm-up:months_0	-0.179	0.146	-1.230	0.219

pragmatics6 ~ gtype + months_0 + months_0:gtype + 0; multinomial logistic regression conducted using the nnet *R-package*

D.2 Chapter 4 Statistics

Appendix Table 8: Post-hoc Chi-square residuals, meaning and speech, early childhood

Meaning	Speech presence, resid.
Ignorance	-11.423***
Investment	5.260***
Disaffiliation	1.617
Disinvestment	0.957
Inquiry	3.695**
Affiliation	1.144
* p<0.05, **p<0.01, ***p<0.001	

Appendix Table 9: Post-hoc Chi-square residuals, form and speech, early childhood

Form	Speech presence, resid.
Reduced palm-up + Shoulder raise	0.596
Reduced palm-up	5.718***
Complete palm-up + Shoulder raise	-2.054
Complete palm-up	-0.476
Shoulder raise	-5.400***
* p<0.05, **p<0.01, ***p<0.001	

Appendix Table 10: GLMM, kernel form and meaning, early childhood

Dependent		Estimate	Std. error	z value	Pr(>|z|)
Kernel meaning	(Intercept)	-0.039	0.003	-14.034	<0.0001
	kernel_form_YNTRUE	1.059	0.003	383.136	<0.0001
	cospeechCo-speech shrug	-2.863	0.003	-1035.279	<0.0001
	age_months_c	-0.222	0.003	-80.312	<0.0001
kernel_form ~ kernel_meaning + cospeech + age_months + (1 + age_months|subject)					
Kernel form	(Intercept)	0.961	0.396	2.423	0.015
	kernel_meaning_YNTRUE	1.188	0.378	3.141	0.002
	cospeechCo-speech shrug	-1.253	0.398	-3.151	0.002
	age_months_c	-0.242	0.203	-1.192	0.233
kernel_form ~ kernel_meaning + cospeech + age_months + (1 + age_months|subject)					

D.3 Chapter 5 Statistics

Appendix Table 11: Post-hoc Chi-square residuals, meaning and speech, adolescence

Meaning	Speech presence, resid.
Ignorance	-4.317***
Investment	4.382***
Disaffiliation	1.054
Disinvestment	-0.257
Inquiry	0.743
Affiliation	-2.589
Turn management	-0.983
* p<0.05, **p<0.01, ***p<0.001	

*Appendix Table 12: Post-hoc Chi-square residuals, form and speech, adolescence*

Form	Speech presence, resid.
Reduced palm-up + Shoulder raise	0.118
Reduced palm-up	-0.182
Complete palm-up + Shoulder raise	0.363
Complete palm-up	0.273
Shoulder raise	-0.266
* p<0.05, **p<0.01, ***p<0.001	

Appendix Table 13: GLMM, kernel form and meaning associations, adolescence

Dependent		Estimate	Std. error	z value	Pr(>|z|)
Kernel meaning	(Intercept)	-5.422	1.462	-3.709	0.000
	kernel_form_YNTRUE	0.622	0.679	0.915	0.360
	cospeechCo-speech shrug	-3.000	0.782	-3.835	0.000
kernel_form ~ kernel_meaning + cospeech + (1 |subject)					
Kernel form	(Intercept)	1.118	0.270	4.147	0.000
	kernel_meaning_YNTRUE	0.387	0.599	0.646	0.518
	cospeechCo-speech shrug	0.138	0.475	0.291	0.771
kernel_form ~ kernel_meaning + cospeech + (1 |subject)					
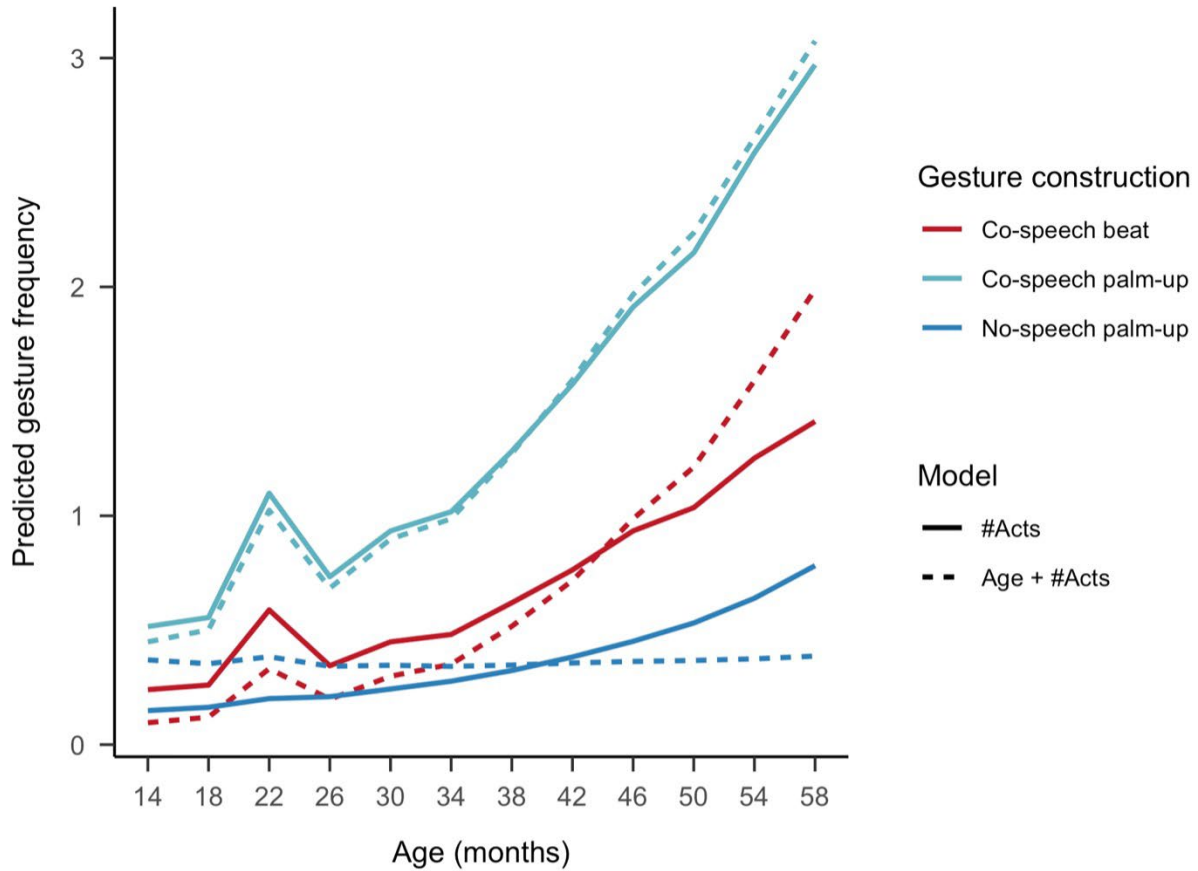
Appendix Figure 2: Comparison of predictive models for gestures frequency . The model fit of GLMM predicting gesture frequency by gesture construction and communicative act count is improved by including child age as a predictor. Children increase use of co-speech gestures, particularly beat gestures, above and beyond expected increases from increasing amount of talk. A model including age predicts no increase in frequency for no-speech palm-ups despite increases in amount of talk.
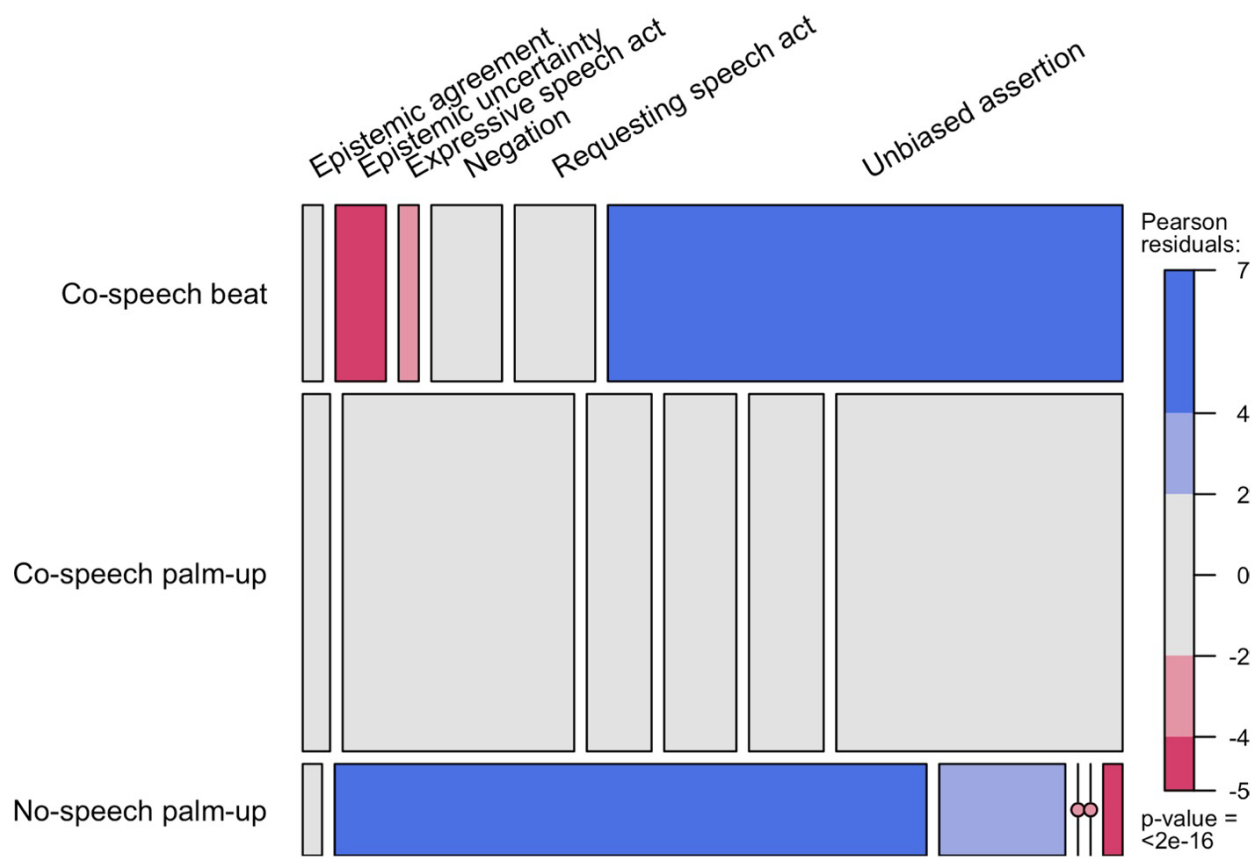
Appendix Figure 3: Chi-square residuals of form-function associations
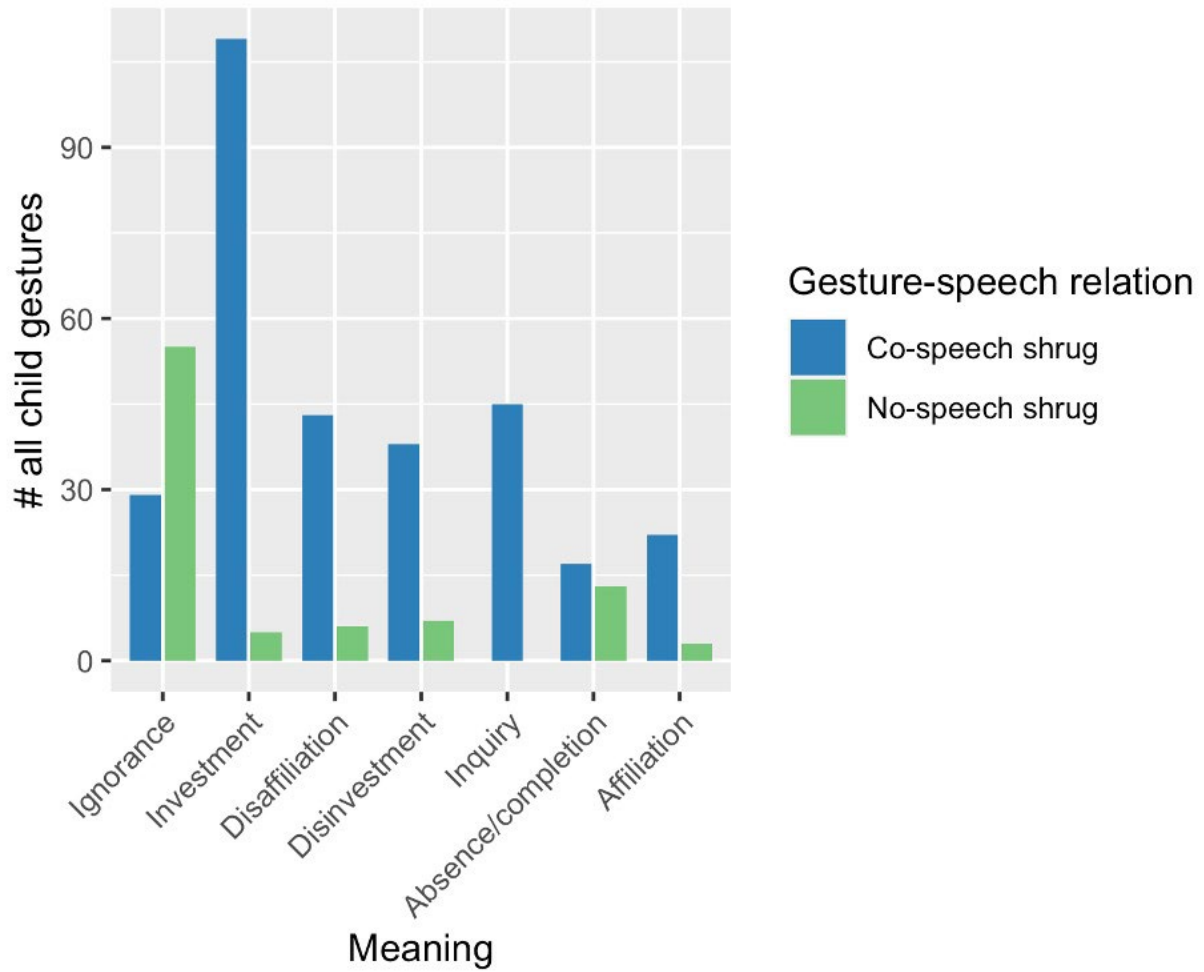
Appendix Figure 4: Raw frequencies of shrug meanings, early childhood. Ignorance is the only meaning produced more often with no-speech shrugs than co-speech shrugs.

References

Abbeduto, L., & Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. *Journal of Child Language*, *12*(3), 621–641.

Acredolo, L. P., & Goodwyn, S. W. (1985). Symbolic gesturing in language development. *Human Development*, *28*(1), 40–49.

Alamillo, A. R., Colletta, J.-M., & Guidetti, M. (2013). Gesture and language in narratives and explanations: The effects of age and communicative activity on late multimodal discourse development. *Journal of Child Language*, *40*(03), 511–538.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-Speech Mismatch and Mechanisms of Learning: What the Hands Reveal about a Child′s State of Mind. *Cognitive Psychology*, *25*(4), 468–523.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, *15*(6), 593–613.

Andersen, E. (2014). *Speaking with style: The sociolinguistics skills of children*. Routledge.

Anderson, A. H., Clark, A., & Mullin, J. (1994). Interactive communication between children: Learning how to make language work in dialogue. *Journal of Child Language*, *21*(2), 439–463.

Asher, S. R. (1976). Children's ability to appraise their own and another person's communication performance. *Developmental Psychology*, *12*(1), 24–32.

Axia, G. (1996). How to persuade mum to buy a toy. *First Language*, *16*(48), 301–317.

Axia, G., & Baroni, M. R. (1985). Linguistic politeness at different age levels. *Child Development*, *56*(4), 918–927.

Balcomb, F. K., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science*, *11*(5), 750–760.

Balog, H. L., & Brentari, D. (2008). The relationship between early gestures and intonation. *First Language*, *28*(2), 141–163.

Bara, B., Bosco, F., & Bucciarelli, M. (1999). Developmental pragmatics in normal and abnormal children. *Brain and Language*, *68*, 507–528.

Bara, B. G. (2010). *Cognitive pragmatics: The mental processes of communication*. MIT Press.

Barak, L., Fazly, A., & Stevenson, S. (2012). Modeling the Acquisition of Mental State Verbs. *Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2012)*, 1–10.

Barakat, R. A. (1973). Arabic gestures. *The Journal of Popular Culture*, *6*(4), 749–793.

Barnes, D., & Todd, F. (1977). *Communication and learning in small groups*. Routledge and Kegan Paul.

Baroni, M. R., & Axia, G. (1989). Children's meta-pragmatic abilities and the identification of polite and impolite requests. *First Language*, *9*(27), 285–297.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford University Press.

Bartz, D. T. (2017). *Young children's meta-ignorance*. [Doctoral dissertation, Harvard University].

Bassano, D. (1996). Functional and formal constraints on the emergence of epistemic modality: A longitudinal study on French. *First Language*, *16*(46), 077–113.

Bassano, D., Hickmann, M., & Champaud, C. (1992). Epistemic modality in french children's discourse: To be sure or not to be sure? *Journal of Child Language*, *19*(2), 389–413.

Bateman, A., & Church, A. (Eds.). (2017). *Children's Knowledge-in-Interaction*. Springer Singapore.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

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

Bates, E. (1979). *The emergence of symbols: Cognition and communication in infancy*. Academic Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development*, *21*(3), 205–226.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, 19(2), 163-194

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures Specialized for Dialogue. *Personality and Social Psychology Bulletin*, *21*(4), 394–405.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*(4), 469–489.

Beattie, G., & Shovelton, H. (2002). An experimental investigation of the role of different types of iconic gesture in communication: A semantic feature approach. *Gesture*, *1*(2), 129–149.

Beaupoil-Hourdel, P., & Debras, C. (2017). Developing communicative postures: The emergence of shrugging in child communication. *Language, Interaction and Acquisition*, *8*(1), 89–116.

Beaupoil-Hourdel, P., Morgenstern, A., & Boutet, D. (2015). Beaupoil-Hourdel, P., Morgenstern, A., & Boutet, D. (2016). A child's multimodal negations from 1 to 4: The interplay between modalities. In *Negation and polarity: Experimental perspectives* (pp. 95-123). Springer, Cham.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology*, *41*(2), 328–337.

Benazzo, S., & Morgenstern, A. (2014). A bilingual child's multimodal path into negation. *Gesture*, *14*(2), 171–202.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. L. Erlbaum Associates.

Bernard, S., Proust, J., & Clément, F. (2015). Four- to six-year-old children's sensitivity to reliability versus consensus in the endorsement of object labels. *Child Development*, *86*(4), 1112–1124.

Billow, R. M. (1975). A cognitive developmental study of metaphor comprehension. *Developmental Psychology*, *11*(4), 415–423.

Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford University Press.

Borràs-Comes, J., Roseano, P., Vanrell Bosch, M., Chen, A., & Prieto, P. (2011). Perceiving uncertainty: facial gestures, intonation, and lexical choice. *Proceedings of GESPIN*.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, *51*(3), 437–457.

Boutet, D. (2008). Une morphologie de la gestualité: Structuration articulaire. *Cahiers de Linguistique Analogique*, *5*, 81–115.

Boutet, D. (2018). The dual origin of gesture: Semiotic preconditions for the emergence of two kinds of symbolic gestures. *Language, Interaction and Acquisition. Language, Interaction et Acquisition*, *8*(2), 288–310.

Boutet, D., Blondel, M., Beaupoil-Hourdel, P., & Morgenstern, A. (2021). A multimodal and kinesiological approach to the development of negation in signing and non-signing children. *Languages and Modalities*, *1*, 31–47.

Bressem, J., & Müller, C. (2017). The "negative-assessment-construction" – a multimodal pattern based on a recurrent gesture? *Linguistics Vanguard*, *3*(1).

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*, 906–921.

Brice, A. E. (1992). The adolescent pragmatics screening scale: Rationale and development. *Howard Journal of Communications*, *3*(3–4), 177–193.

Brinton, B., & Fujiki, M. (1984). Development of Topic Manipulation Skills in Discourse. *Journal of Speech, Language, and Hearing Research*, *27*(3), 350–358.

Brodzinsky, D. M. (1977). Children's Comprehension and Appreciation of Verbal Jokes in Relation to Conceptual Tempo. *Child Development*, *48*(3), 960–967.

Brookes, H. (2004). A repertoire of South African quotable gestures. *Journal of Linguistic Anthropology*, *14*(2), 186–224.

Brown, A. R., Pouw, W., Brentari, D., & Goldin-Meadow, S. (2021). People Are Less Susceptible to Illusion When They Use Their Hands to Communicate Rather Than Estimate. *Psychological Science*, *32*(8), 1227–1237.

Brown, L., & Winter, B. (2019). Multimodal indexicality in Korean: "Doing deference" and "performing intimacy" through nonverbal behavior. *Journal of Politeness Research*, *15*(1), 25–54.

Brown, L., Winter, B., Idemaru, K., & Grawunder, S. (2014). Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues. *Journal of Pragmatics*, *66*, 45–60.

Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–258). New York: Cambridge University Press.

Cairns, H. S., Waltzman, D., & Schlisselberg, G. (2004). Detecting the Ambiguity of Sentences: Relationship to Early Reading Skill. *Communication Disorders Quarterly*, *25*(2), 68–78.

Calbris, G. (1990). *The semiotics of french gestures*. Indiana University Press.

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, *3*(4), 207–220.

Capelli, C. A., Nakagawa, N., & Madden, C. M. (1990). How children understand sarcasm: The role of context and intonation. *Child Development*, *61*(6), 1824–1841.

Capirci, O., Cristilli, C., de Angelis, V., & Graziano, M. (2011). Learning to use gesture in narratives. In *Integrating Gestures: The interdisciplinary nature of gesture* (Vol. 4). John Benjamins Publishing Company.

Cartmill, E. A., Hunsicker, D., & Goldin-Meadow, S. (2014). Pointing and naming are not redundant: Children use gesture to modify nouns before they modify nouns in speech. *Developmental Psychology*, *50*(6), 1660–1666.

Caselli, M. C. (1983). Communication to language: Deaf children's and hearing children's development compared. *Sign Language Studies*, *1039*(1), 113–144.

Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn-taking, timing, and planning in early language acquisition. *Journal of Child Language*, *43*(06), 1310–1337.

Casillas, M., & Hilbrink, E. (2020). Communicative act development. In K. P. Schneider & E. Ifantidou (Eds.), *Developmental and Clinical Pragmatics* (pp. 61–88). De Gruyter.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, *143*(2), 694–709.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*(1), 43–71.

Clark, E. V. (2009). *First language acquisition*. Cambridge University Press.

Clark, R. A., & Delia, J. G. (1976). The development of functional persuasive skills in childhood and early adolescence. *Child Development*, *47*(4), 1008–1014.

Coates, J. (1988). The acquisition of the meanings of modality in children aged eight and twelve. *Journal of Child Language*, *15*(2), 425–434.

Cochet, H., & Vauclair, J. (2014). Deictic gestures and symbolic gestures produced by adults in an experimental context: Hand shapes and hand preferences. *Laterality: Asymmetries of Body, Brain and Cognition*, *19*(3), 278–301.

Colletta, J.-M., Guidetti, M., Capirci, O., Cristilli, C., Demir, O. E., Kunene-Nicolas, R. N., & Levine, S. (2015). Effects of age and language on co-speech gesture production: An investigation of French, American, and Italian children's narratives. *Journal of Child Language*, *42*(01), 122–145.

Colletta, J.-M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from french children and adults. *Speech Communication*, *52*(6), 565–576.

Cook, H. M. (2012). Language Socialization and Stance-Taking Practices. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 296–321). Wiley-Blackwell.

Cooperrider, K., Abner, N., & Goldin-Meadow, S. (2018). The palm-up puzzle: Meanings and origins of a widespread form in gesture and sign. *Frontiers in Communication*, *3*, 23.

Corsaro, W. A. (1979). "We're Friends, Right?": Children's Use of Access Rituals in a Nursery School. *Language in Society*, *8*(3), 315–336.

Couper-Kuhlen, E. (2010). Commentary on Stivers and Rossano: "Mobilizing Response." *Research on Language and Social Interaction*, *43*(1), 32–37.

Creider, C. A. (1977). Towards a description of east African gestures. *Sign Language Studies*, *14*, 1–20.

Crowder, E. M., & Newman, D. (1993). Telling what they know: The role of gesture and language in children's science explanations. *Pragmatics & Cognition*, *1*(2), 341–376.

Dannequin, C. (1987). Les enfants baillonnes (gagged children): The teaching of French as mother tongue in elementary school. *Language and Education*, *1*(1), 15–31.

Darwin, C. (1998). *The expression of the emotions in man and animals*. Oxford University Press. (Original work published 1872).

Debras, C. (2015). Stance-taking functions of multimodal constructed dialogue during spoken interaction. *Actes Du Colloque Gesture and Speech in Interaction*, *4*, 6.

Debras, C. (2017). The shrug: Forms and meanings of a compound enactment. *Gesture*, *16*(1), 1–34.

Debras, C., & Cienki, A. (2012). Some uses of head tilts and shoulder shrugs during human interaction, and their relation to stancetaking. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, 932–937.

Demir, Ö. E., Fisher, J. A., Goldin-Meadow, S., & Levine, S. C. (2014). Narrative processing in typically developing children and children with early unilateral brain injury: Seeing gesture matters. *Developmental Psychology*, *50*(3), 815–828.

Demir, Ö. E., Levine, S. C., & Goldin-Meadow, S. (2015). A tale of two hands: Children's early gesture use in narrative production predicts later narrative structure in speech. *Journal of Child Language*, *42*(3), 662–681.

Dent, C. H. (1984). The Developmental Importance of Motion Information in Perceiving and Describing Metaphoric Similarity. *Child Development*, *55*(4), 1607–1613.

Dijkstra, C., Krahmer, E., & Swerts, M. (2006). Manipulating uncertainty. The contribution of different audiovisual prosodic cues to the perception of confidence. *Proceedings of the Speech Prosody Conference*.

Dimitrova, D., Chu, M., Wang, L., Özyürek, A., & Hagoort, P. (2016). Beat that word: How listeners integrate beat gesture and focus in multimodal speech discourse. *Journal of Cognitive Neuroscience*, *28*(9), 1255–1269.

Dodane, C., Boutet, D., Didirkova, I., Hirsch, F., Ouni, S., & Morgenstern, A. (2019). An integrative platform to capture the orchestration of gesture and speech. *Proceedings of the 6th Gesture and Speech in Interaction Conference*.

Dohen, M., & Lœvenbruck, H. (2009). Interaction of audition and vision for the perception of prosodic contrastive focus. *Language and Speech*, *52*(2–3), 177–206.

Dorval, B., & Eckerman, C. O. (1984). Developmental Trends in the Quality of Conversation Achieved by Small Groups of Acquainted Peers. *Monographs of the Society for Research in Child Development*, *49*(2), 1–91.

Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Pragmatics & Beyond New Series* (Vol. 164, pp. 139–182). John Benjamins Publishing Company.

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*(2), 331–338.

Eigsti, I.M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681–691.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. In A. Kendon, T. A. Sebeok, & J. Umiker-Sebeok (Eds.), *Nonverbal Communication, Interaction, and Gesture* (Vol. 41, pp. 57–106). De Gruyter Mouton.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication*, *22*(4), 353–374.

Emihovich, C. A. (1981). The intimacy of address: Friendship markers in children's social play. *Language in Society*, *10*(2), 189–199.

Enfield, N. J., Stivers, T., Brown, P., Englert, C., Harjunpää, K., Hayashi, M., Heinemann, T., Hoymann, G., Keisanen, T., Rauniomaa, M., Raymond, C. W., Rossano, F., Yoon, K.-E., Zwitserlood, I., & Levinson, S. C. (2019). Polar answers. *Journal of Linguistics*, *55*(2), 277–304.

Enfield, N. J., Stivers, T., & Levinson, S. C. (2010). Question–response sequences in conversation across ten languages: An introduction. *Journal of Pragmatics*, *42*(10), 2615–2619.

Ervin-Tripp, S., Guo, J., & Lampert, M. (1990). Politeness and persuasion in children's control acts. *Journal of Pragmatics*, *14*(2), 307–331.

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

Esteve-Gibert, N., & Prieto, P. (2018). Early development of the prosody-meaning interface. In *Trends in Language Acquisition Research* (Vol. 23, pp. 227–246). John Benjamins Publishing Company.

Evans, M. A. (1987). Discourse characteristics of reticent children. *Applied Psycholinguistics*, *8*(2), 171–184.

Ferré, G. (2012). Functions of three open-palm hand gestures. *Journal Multimodal Communication*, *1*(1), 5–20.

Floyd, S., Giovanni, R., & Enfield, N. J. (2020). *Getting others to do things: A pragmatic typology of recruitments*. Zenodo.

Frausel, R. R., Silvey, C., Freeman, C., Dowling, N., Richland, L. E., Levine, S. C., Raudenbush, S., & Goldin-Meadow, S. (2020). The origins of higher-order thinking lie in children's spontaneous talk across the pre-school years. *Cognition*, *200*, 104274.

Gardner, H., & Forrester, M. A. (Eds.). (2010). *Analysing interactions in childhood: Insights from conversation analysis*. Wiley-Blackwell.

Gawne, L. (2018). Contexts of use of a rotated palms gesture among Syuba (Kagate) speakers in Nepal. *Gesture*, *17*(1), 37–64.

Gerwing, J., & Bavelas, J. B. (2013). The social interactive nature of gestures: Theory, assumptions, methods, and findings. In *Body—Language—Communication* (pp. 1–17). Walter de Gruyter.

Givens, D. B. (1977). Shoulder shrugging: A densely communicative expressive behavior. *Semiotica*, *19*(1–2), 13–28.

Gleason, J. B., & Weintraub, S. (1976). The acquisition of routines in child language. *Language in Society*, *5*(2), 129–136.

Gnepp, J., & Hess, D. L. (1986). Children's understanding of verbal and facial display rules. *Developmental Psychology*, *22*(1), 103–108.

Goldin-Meadow, S. (1998). The development of gesture and speech as an integrated system. *New Directions for Child and Adolescent Development*, *1998*(79), 29–42.

Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Harvard University Press.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review*, *100*(2), 279–297.

Goldin-Meadow, S., & Butcher, C. (2008). Pointing Toward Two-Word Speech in Young Children. In *Pointing: Where Language, Culture, and Cognition Meet*.

Gombert, J. É. (1992). *Metalinguistic Development*. University of Chicago Press.

Goodwin, M. H. (1980). he-said-she-said: Formal cultural procedures for the construction of a gossip dispute activity. *American Ethnologist*, *7*(4), 674–695.

Goodwin, M. H. (2006). *The hidden life of girls: Games of stance, status, and exclusion*. Blackwell Pub.

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492–3496.

Graziano, M. (2014a). The development of two pragmatic gestures of the so-called open hand supine family in Italian children. In M. Seyfeddinipur & M. Gullberg (Eds.), *From Gesture*

in Conversation to Visible Action as Utterance (pp. 311–330). John Benjamins Publishing Company.

Graziano, M. (2014b). Gestures in Southern Europe: Children's pragmatic gestures in Italy. In *Body—Language—Communication* (pp. 1253–1258). De Gruyter Mouton.

Greif, E. B., & Gleason, J. B. (1980). Hi, thanks, and goodbye: More routine information. *Language in Society*, *9*(2), 159–166.

Guidetti, M. (2002). The emergence of pragmatics: Forms and functions of conventional gestures in young French children. *First Language*, *22*(3), 265–285.

Guidetti, M. (2005). Yes or no? How young French children combine gestures and speech to agree and refuse. *Journal of Child Language*, *32*(4), 911–924.

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, *56*(1), 155–196.

Hadar, U., & Butterworth, B. (1997). *Iconic gestures, imagery, and word retrieval in speech*. *115*(1–2), 147–172.

Harris, P. L. (2020). The Point, the Shrug, and the Question of Clarification. In L. P. Butler, S. Ronfard, & K. H. Corriveau (Eds.), *The Questioning Child* (1st ed., pp. 29–50). Cambridge University Press.

Harris, P. L., Bartz, D. T., & Rowe, M. L. (2017). Young children communicate their ignorance and ask questions. *Proceedings of the National Academy of Sciences*, *114*(30), 7884–7891.

Harris, P. L., Yang, B., & Cui, Y. (2017). 'I don't know': Children's early talk about knowledge. *Mind & Language*, *32*(3), 283–307.

Hemphill, L., & Siperstein, G. N. (1990). Conversational competence and peer response to mildly retarded children. *Journal of Educational Psychology*, *82*(1), 128–134.

Hickmann, M., & Bassano, D. (2016). Modality and mood in first language acquisition. In *The Oxford Handbook of Modality and Mood*.

Hickmann, M., Champaud, C., & Bassano, D. (1993). Pragmatics and metapragmatics in the development of epistemic modality: Evidence from French children's reports of think-statements. *First Language*, *13*(39), 359–388.

Hollos, M. (1975). *Comprehension and use of social rules in pronoun selection by Hungarian children. Working papers in sociolinguistics, no. 24*.

Holtgraves, T. (2000). Preference Organization and Reply Comprehension. *Discourse Processes*, *30*(2), 87–106.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315.

Hübscher, I., Borràs-Comes, J., & Prieto, P. (2017). Prosodic mitigation characterizes Catalan formal speech: The frequency code reassessed. *Journal of Phonetics*, *65*, 145–159.

Hübscher, I., Esteve-Gibert, N., Igualada, A., & Prieto, P. (2017). Intonation and gesture as bootstrapping devices in speaker uncertainty. *First Language*, *37*(1), 24–41.

Hübscher, I., Garufi, M., & Prieto, P. (2019). The development of polite stance in preschoolers: How prosody, gesture, and body cues pave the way. *Journal of Child Language*, *46*(05), 825–862.

Hübscher, I., & Prieto, P. (2019). Gestural and prosodic development act as sister systems and jointly pave the way for children's sociopragmatic development. *Frontiers in Psychology*, *10*.

Hübscher, I., Vincze, L., & Prieto, P. (2019). Children's signaling of their uncertain knowledge state: Prosody, face, and body cues come first. *Language Learning and Development*, *15*(4), 366–389.

Hundertmark, A. (2016). *Form and Function: A descriptive study of palm-up gestures in preschool children* [Master's thesis]. University of Chicago.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365.

Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language*, *28*(2), 164–181.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367–371.

Jehoul, A., Brône, G., & Feyaerts, K. (2017). The shrug as marker of obviousness. *Linguistics Vanguard*, *3*(s1).

Joffe, V. L., & Nippold, M. A. (2012). Progress in understanding adolescent language disorders. *Language, Speech, and Hearing Services in Schools*, *43*(4), 438–444.

Johnson, H. G., Ekman, P., & Friesen, W. V. (1975). *Communicative body movements: American emblems*. 335–354.

Jokinen, K., & Allwood, J. (2010). Hesitation in intercultural communication: Some observations and analyses on interpreting shoulder shrugging. In T. Ishida (Ed.), *Culture and Computing* (Vol. 6259, pp. 55–70). Springer Berlin Heidelberg.

Kassambara, A. (2021). *Rstatix: Pipe-friendly framework for basic statistical tests*. Retrieved from https://CRAN.R-project.org/package=rstatix

Keenan, T. R., & Quigley, K. (1999). Do young children use echoic information in their comprehension of sarcastic speech? A test of echoic mention theory. *British Journal of Developmental Psychology*, *17*(1), 83–96.

Kendon, A. (2002). Some uses of the head shake. *Gesture*, *2*(2), 147–182.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kendon, A. (2017). Reflections on the "gesture-first" hypothesis of language origins. *Psychonomic Bulletin & Review*, *24*(1), 163–170.

Kendon, A., & Versante, L. (2003). Pointing by hand in "Neapolitan." In *Pointing: Where Language, Culture, and Cognition Meet* (pp. 109–137).

Kernan, K. T. (1977). Semantic and expressive elaboration in children's narratives. In S. Ervin-Tripp & C. Mitchell-Keraan (Eds.), *Child Discourse* (pp. 91–102). Academic Press.

Kessel, F. S. (1970). The Role of Syntax in Children's Comprehension from Ages Six to Twelve. *Monographs of the Society for Research in Child Development*, *35*(6), iii–95.

Kim, J., Cvejic, E., & Davis, C. (2014). Tracking eyebrows and head gestures associated with spoken prosody. *Speech Communication*, *57*, 317–330.

Kim, S., Paulus, M., Sodian, B., & Proust, J. (2016). Young children's sensitivity to their own ignorance in informing others. *PLOS ONE*, *11*(3), e0152595.

Kita, S. (Ed.). (2003). *Pointing: Where language, culture, and cognition meet*. L. Erlbaum Associates.

Kita, S. (2009). *How representational gestures help speaking*. Cambridge University Press.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245–266.

Kleeck, A. van. (1984). Metalinguistic Skills: Cutting across Spoken and Written Language and Problem-Solving Abilities. In G. P. Wallach & K. G. Butler (Eds.), *Language Learning Disabilities in School-Age Children* (Vol. 1–xiv, 378, pp. 128–153). Williams & Wilkins.

Kockelman, P. (2004). Stance and Subjectivity. *Journal of Linguistic Anthropology*, *14*(2), 127–150.

Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and speech, 48*(1), 29-53

Krahmer, E., & Swerts, M. (2009). Audiovisual prosody—Introduction to the special issue. *Language and Speech*, *52*(2–3), 129–133.

Krauss, R. M. (1998). Why Do We Gesture When We Speak? *Current Directions in Psychological Science*, *7*(2), 54–54.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 28, pp. 389–450). Academic Press.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and Gesture* (pp. 261–283). Cambridge University Press.

Krifka, M. (2015). Bias in commitment space semantics: Declarative questions, negated questions, and question tags. *Semantics and Linguistic Theory*, *25*, 328–345.

Kunene, R. N. (2010). *A comparative study of the development of multimodal narratives in French and Zulu children and adults* [Grenoble].

Larson, V. L., & McKinley, N. L. (1998). Characteristics of adolescents' conversations: A longitudinal study. *Clinical Linguistics & Phonetics*, *12*(3), 183–203.

Laval, V., & Bert-Erboul, A. (2018, 49-12 10:49:46). *French-speaking children's understanding of sarcasm* (world) [Research-article]. ASHA Wire; American Speech-Language-Hearing Association.

Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences*, *24*(1), 65–78.

Levy, E. T., & McNeill, D. (2013). Narrative development as symbol formation: Gestures, imagery and the emergence of cohesion. *Culture & Psychology*, *19*(4), 548–569.

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787.

MacLure, M., & French, P. (1981). A comparison of talk at home and at school. In G. Wells (Ed.), *Learning Through Interaction: Volume 1: The Study of Language Development*. Cambridge University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (Third). Lawrence Erlbaum Associates.

Maltz, D., & Borker, R. (1996). A cultural approach to male-female miscommunication. In D. L. Brenneis & R. K. S. Macaulay (Eds.), *The matrix of language: Contemporary linguistic anthropology*. Westview Press.

Martin, J. R. (1983). The development of register. In J. Fine & R. O. Freedle (Eds.), *374186* (Vol. 1–xi, 315, pp. 1–39). Ablex Publishing.

Mathew, M., Yuen, I., & Demuth, K. (2018). Talking to the beat: Six-year-olds' use of stroke-defined non-referential gestures. *First Language*, *38*(2), 111–128.

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: A review of associations with formal language, social cognition, and executive functions. *Language Learning and Development*, *14*(3), 186–223.

McClave, E. Z. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, *32*(7), 855–878.

McClements, R. J. (1976). Aspects of the structure of narrative speech in children aged 5, 7, 9 & 11 years. *Occasional Papers in Linguistics and Language Learning*, *2*.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350–371.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

McNeill, D., Cassell, J., & McCullough, K.E. (1994). Communicative Effects of Speech-Mismatched Gestures. *Research on Language and Social Interaction*, *27*(3), 223–237.

McTear, M. (1985). *Children's conversation*. B. Blackwell.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*.

Michaels, S. (1981). "Sharing Time": Children's Narrative Styles and Differential Access to Literacy. *Language in Society*, *10*(3), 423–442.

Milosky, L. (1994). Nonliteral language abilities: Seeing the forest for the trees. In G. P. Wallach & K. G. Butler (Eds.), *Language learning disabilities in school-age children and adolescents: Some principles and applications*. Merrill; Maxwell Macmillan Canada; Maxwell Macmillan International.

Mondada, L. (2013). 13. Conversation analysis: Talk and bodily resources for the organization of social interaction. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/1* (pp. 218–226). De Gruyter.

Morris, C. (1946). *Signs, language, and behavior*. Prentice Hall.

Morris, D. (2015). *Bodytalk: A World Guide to Gestures*. Random House. (Original work published 1994)

Müller, C. (2004). Forms and uses of the palm up open hand: A case of a gesture family? *The Semantics and Pragmatics of Everyday Gestures*, *9*, 233–256.

Müller, C. (2017). How recurrent gestures mean: Conventionalized contexts-of-use and embodied motivation. *Gesture*, *16*(2), 277–304.

Nicoladis, E., Mayberry, R. I., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology*, *35*(2), 514–526.

Ninio, A., & Snow, C. E. (1996). *Pragmatic development*. Westview Press.

Ninio, A., Snow, C. E., Pan, B. A., & Rollins, P. R. (1994). Classifying communicative acts in children's interactions. *Journal of Communication Disorders*, *27*(2), 157–187.

Ninio, A., & Wheeler, P. (1986). *A manual for classifying verbal communicative acts in mother-infant interaction*.

Nippold, M. A. (1998). *Language development in school-age children, adolescents, and adults*. 368–373.

Nippold, M. A. (2000). Language development during the adolescent years. *Topics in Language Disorders*, *20*(2), 15–28.

Nippold, M. A. (2016). *Later language development: School-age children, adolescents, and young adults* (Fourth Edition).

Nippold, M. A., Leonard, L. B., & Anastopoulos, A. (1982). Development in the use and understanding of polite forms in children. *Journal of Speech, Language, and Hearing Research*, *25*(2), 193–202.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

Noveck, I. A., Ho, S., & Sera, M. (1996). Children's understanding of epistemic modals. *Journal of Child Language (Print)*, *23*(3), 621–643.

Ochs, E. (1996). Linguistic Resources for Socializing Humanity. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 407–437). Cambridge University Press.

Ochs, E. (1982). Talking to Children in Western Samoa. *Language in Society*, *11*(1), 77–104.

Ochs, E., & Schieffelin, B. B. (2001). Language acquisition and socialization: Three developmental stories and their implications. In Linguistic Anthropology: A Reader (p. 25). Blackwell.

Ochs, E., & Schieffelin, B. B. (2016). *Acquiring conversational competence*.

O'Neill, D. K., & Atance, C. M. (2000). "maybe my daddy give me a big piano": The development of children's use of modals to express uncertainty. *First Language*, *20*(58), 029–052.

Özçalişkan, Ş., Gentner, D., & Goldin-Meadow, S. (2014). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistics*, *35*(6), 1143–1162.

Özçalişkan, Ş., & Goldin-Meadow, S. (2005). Do parents lead their children by the hand? *Journal of Child Language*, *32*(3), 481–505.

Özçalişkan, Ş., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, *96*(3), B101–B113.

Özçalişkan, Ş., Levine, S. C., & Goldin-Meadow, S. (2013). Gesturing with an injured brain: How gesture helps children with early brain injury learn linguistic constructions. *Journal of Child Language*, *40*(1), 69–105.

Papafragou, A. (1998). The acquisition of modality: Implications for theories of semantic representation. *Mind and Language*, *13*(3), 370–399.

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. Language Acquisition, 12(1), 71–82.

Parrill, F. (2008). Form, meaning, and convention: A comparison of a metaphoric gesture with an emblem. In A. Cienki & C. Müller (Eds.), *Gesture Studies* (Vol. 3, pp. 195–217). John Benjamins Publishing Company.

Parrill, F., Lavanty, B., Bennett, A., Klco, A., & Demir-Lira, O. E. (2018). The relationship between character viewpoint gesture and narrative structure in children. *Language and Cognition*, *10*(3), 408–434.

Paulus, M., Tsalas, N., Proust, J., & Sodian, B. (2014). Metacognitive monitoring of oneself and others: Developmental changes during childhood and adolescence. *Journal of Experimental Child Psychology*, *122*, 153–165.

Payrató, L. (1993). A pragmatic view on autonomous gestures: A first repertoire of Catalan emblems. *Journal of Pragmatics*, *20*(3), 193–216.

Pedlow, R., Sanson, A., & Wales, R. (2004). Children's production and comprehension of politeness in requests: Relationships to behavioural adjustment, temperament and empathy. *First Language*, *24*(3), 347–367.

Peirce, C. S. (1991). *Peirce on signs: Writings on semiotic*. University of North Carolina Press.

Perniss, P. (2018). Why we should study multimodal language. *Frontiers in Psychology*, *9*.

Perry, M., Breckinridge Church, R., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*(4), 359–400.

Peterson, C., & Dodsworth, P. (1991). A longitudinal analysis of young children's cohesion and noun specification in narratives. *Journal of Child Language*, *18*(02), 397.

Piché, G. L., Rubin, D. L., & Michlin, M. L. (1978). Age and social class in children's use of persuasive communicative appeals. *Child Development*, *49*(3), 773–780.

Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, *47*(1), 116–129.

Pomerantz, A., & Heritage, J. (2012). Preference. In *The Handbook of Conversation Analysis* (pp. 210–228).

Prieto, P., Cravotta, A., Kushch, O., Rohrer, P., & Vilà-Giménez, I. (2018). Deconstructing beat gestures: A labelling proposal. *Speech Prosody 2018*, 201–205.

Pyers, J. E., Magid, R., Gollan, T. H., & Emmorey, K. (2021). Gesture Helps, Only If You Need It: Inhibiting Gesture Reduces Tip-of-the-Tongue Resolution for Those With Weak Short-Term Memory. *Cognitive Science*, *45*(1), e12914.

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Raffaelli, M., & Duckett, E. (1989). "we were just talking ...": Conversations in early adolescence. *Journal of Youth and Adolescence*, *18*(6), 567–582.

Rawlins, W. K. (1992). *Friendship matters: Communication, dialectics, and the life course*. Aldine de Gruyter.

Raymond, G. (2003). Grammar and Social Organization: Yes/No Interrogatives and the Structure of Responding. *American Sociological Review*, *68*(6), 939–967.

Read, B. K., & Cherry, L. J. (2009). Preschool children's production of directive forms. *Discourse Processes*, *1*, 233–245.

Rector, M. (1986). Emblems in Brazilian culture. In P. Bouissac, M. Herzfeld, & R. Posner (Eds.), *Iconicity: Essays on the Nature of Culture* (pp. 447–462).

Robinson, J. D. (2020). Revisiting Preference Organization in Context: A Qualitative and Quantitative Examination of Responses to Information Seeking. *Research on Language and Social Interaction*, *53*(2), 197–222.

Rohwer, M., Kloo, D., & Perner, J. (2012). Escape from metaignorance: How children develop an understanding of their own lack of knowledge. *Child Development*, *83*(6), 1869–1883.

Romaine, S. (1951-). (1984). *The language of children and adolescents: The acquisition of communicative competence*. Basil Blackwell.

Roseano, P., González, M., Borràs-Comes, J., & Prieto, P. (2016). Communicating epistemic stance: How speech and gesture patterns reflect epistemicity and evidentiality. *Discourse Processes*, *53*(3), 135–174.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(1), 185–205.

Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, *12*(1), 182–187.

Ruffman, T. K., & Olson, D. R. (1989). Children's ascriptions of knowledge to others. *Developmental Psychology*, *25*(4), 601–606.

Russell, R. L., & Grizzle, K. L. (2008). Assessing child and adolescent pragmatic language competencies: Toward evidence-based assessments. *Clinical Child and Family Psychology Review*, *11*(1), 59–73.

Saarni, C. (1984). An Observational Study of Children's Attempts to Monitor Their Expressive Behavior. *Child Development*, *55*(4), 1504–1513.

Sachs, J. (1983). Gordon wells, learning through interaction: The study of language development. Cambridge: Cambridge university press, 1981. pp. viii + 304. *Journal of Child Language*, *10*(3), 676–679.

Sacks, H. (1992a). *Lectures on conversation* (G. Jefferson, Ed.; Repr., Vol. 1). Blackwell.

Sacks, H. (1992b). *Lectures on conversation* (G. Jefferson, Ed.; Repr., Vol. 2). Blackwell.

Schegloff, E. A. (2010). Commentary on Stivers and Rossano: "Mobilizing Response." *Research on Language and Social Interaction*, *43*(1), 38–48.

Schegloff, E. A., & Sacks, H. (1973). Opening up Closings. *Semiotica*, *8*(4).

Schober-Peterson, D., & Johnson, C. J. (1993). The performance of eight- to ten-year-olds on measures of conversational skillfulness. *First Language*, *13*(38), 249–269.

Scott, C. M. (1984). Adverbial connectivity in conversations of children 6 to 12. *Journal of Child Language*, *11*(2), 423–452.

Shattuck-Hufnagel, S., & Ren, A. (2018). The prosodic characteristics of non-referential co-speech gestures in a sample of academic-lecture-style speech. *Frontiers in Psychology*, *9*, 1514.

Shattuck-Hufnagel, S., Ren, A., Mathew, M., Yuen, I., & Demuth, K. (2016). Non-referential gestures in adult and child speech: Are they prosodic? *Proc. Speech Prosody 2016*, 836–839.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, *14*(3), 301–321.

Short-Meyerson, K. J., & Abbeduto, L. J. (1997). Preschoolers' communication during scripted interactions. *Journal of Child Language*, *24*(2), 469–493.

Shultz, T. R., & Pilon, R. (1973). Development of the Ability to Detect Linguistic Ambiguity. *Child Development*, *44*(4), 728–733.

Sidnell, J. (2016). A Conversation Analytic Approach to Research on Early Childhood. In *The SAGE Handbook of Early Childhood Research* (pp. 255–276).

Snow, D. P. (2017). Gesture and intonation are "sister systems" of infant communication: Evidence from regression patterns of language development. *Language Sciences*, *59*, 180–191.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, *58*(2), 424–433.

Sparhawk, C. M. (1978). Contrastive-identificational features of Persian gesture. *Semiotica*, *24*(1–2).

Spector, C. C. (1996). Children's Comprehension of Idioms in the Context of Humor. *Language, Speech, and Hearing Services in Schools*, *27*(4), 307–313.

Stivers, T. (2010). An overview of the question–response system in American English conversation. *Journal of Pragmatics*, *42*(10), 2772–2781.

Stivers, T. (2012). Sequence Organization. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 191–209). John Wiley & Sons, Ltd.

Stivers, T. (2015). Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction*, *48*(1), 1–19.

Stivers, T., & Enfield, N. J. (2010). A coding scheme for question–response sequences in conversation. *Journal of Pragmatics*, *42*(10), 2620–2626.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592.

Stivers, T., & Rossano, F. (2010). Mobilizing Response. *Research on Language and Social Interaction*, *43*(1), 3–31.

Stivers, T., Sidnell, J., & Bergen, C. (2018). Children's responses to questions in peer interaction: A window into the ontogenesis of interactional competence. *Journal of Pragmatics*, *124*, 14–30.

Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning* (Vol. 2). John Benjamins Publishing Company.

Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in mandarin- and Cantonese-speaking children. *Developmental Psychology*, *36*(1), 25–43.

Tardif, T., Wellman, H. M., Fung, K. Y. F., Liu, D., & Fang, F. (2005). Preschoolers' understanding of knowing-that and knowing-how in the united states and Hong Kong. *Developmental Psychology*, *41*(3), 562–573.

Teßendorf, S. (2013). Emblems, quotable gestures, or conventionalized body movements. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body—Language—Communication* (pp. 82–100). de Gruyter.

Tomasello, M. (2003). On the different origins of symbols and grammar. *Studies in the Evolution of Language*, *3*, 94–110.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691.

Turkstra, L. S. (2001). Partner effects in adolescent conversations. *Journal of Communication Disorders*, *34*(1–2), 151–162.

Van Kleeck, A. (1982). The Emergence of Linguistic Awareness: A Cognitive Framework. *Merrill-Palmer Quarterly*, *28*(2), 237–265.

Venables, W. N., & Ripley, B. D. (2002b). *Modern applied statistics with s* (Fourth). New York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/

Vilà-Giménez, I., Demir-Lira, Ö. E., & Prieto, P. (2020, September 9). The role of referential iconic and non-referential beat gestures in children's narrative production: Iconics signal oncoming changes in speech. *Proceedings of the 7th Gesture and Speech in Interaction (GESPIN), KTH Speech, Music & Hearing and Språkbanken Tal*. Gesture and Speech in Interaction (GESPIN), KTH Speech, Music & Hearing and Språkbanken Tal, Stockholm, Sweden.

Vilà-Giménez, I., Dowling, N., Demir-Lira, Ö. E., Prieto, P., & Goldin-Meadow, S. (2021). The predictive value of non-referential beat gestures: Early use in parent–child interactions predicts narrative abilities at 5 years of age. *Child Development*, cdev.13583.

Vilà-Giménez, I., Igualada, A., & Prieto, P. (2019). Observing storytellers who use rhythmic beat gestures improves children's narrative discourse performance. *Developmental Psychology*, *55*(2), 250–262.

Vilà-Giménez, I., & Prieto, P. (2020). Encouraging kids to beat: Children's beat gesture production boosts their narrative performance. *Developmental Science*, *23*(6).

Vilà-Giménez, I., & Prieto, P. (2021). The value of non-referential gestures: A systematic review of their cognitive and linguistic effects in children's language development. *Children*, *8*(2), 148.

Vincze, L., & Poggi, I. (2016). I am definitely certain of this! Towards a multimodal repertoire of signals communicating a high degree of certainty. *Proceedings of the 4th European and 7th Nordic Symposium on Multimodal Communication*.

Visser, M., Krahmer, E., & Swerts, M. (2011). Children's expression of uncertainty in collaborative and competitive contexts. *Auditory-Visual Speech Processing*, 25–30.

Volterra, V., & Antinucci, F. (1979). Negation in child language: A pragmatic study. *E. Ochs Keenan*.

Wallach, G. P., & Butler, K. G. (1994). *Language learning disabilities in school-age children and adolescents*. Macmillan College Publishers.

Wellman, H. M., & Lempers, J. D. (1977). The naturalistic communicative abilities of two-year-olds. *Child Development*, *48*(3), 1052–1057.

Wells, G., & Bridges, A. (1981). *Learning through interaction: Volume 1: the study of language development*. Cambridge University Press.

Whitehead, K. A. (2011). Some uses of head nods in "third position" in talk-in-interaction. *Gesture*, *11*(2), 103–122.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Wilkens, D. (2003). Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Ed.), *Pointing: Where Language, Culture, and Cognition Meet* (pp. 171–215). Psychology Press.

Willes, M. (1981). Learning to take part in classroom interaction. In P. French & M. MacLure (Eds.), *Adult-Child Conversation* (Vol. 1–310, pp. 73–90). St. Martin's Press.

Wimmer, H., Hogrefe, G.-J., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development*, *59*(2), 386–396.

Winter, B., & Grawunder, S. (2012). The phonetic profile of Korean formal and informal speech registers. *Journal of Phonetics*, *40*(6), 808–815.

Yap, D. F., Brookshire, G., & Casasanto, D. (2018). Beat gestures encode spatial semantics. *40th Annual Conference of the Cognitive Science Society*, 6.

Yoshioka, K. (2008). Gesture and information structure in first and second language. *Gesture*, *8*(2), 236–255.

Yu, Y., Bonawitz, E., & Shafto, P. (2019). Pedagogical Questions in Parent–Child Conversations. *Child Development*, *90*(1), 147–161.

Zaidman-Zait, A., & Most, T. (2020). Pragmatics and peer relationships among deaf, hard of hearing, and hearing adolescents. *Pediatrics*, *146*(Supplement 3), S298–S303.