Children's Signaling of Their Uncertain Knowledge State: Prosody, Face, and Body Cues Come First

Iris Hübscher, Laura Vincze & Pilar Prieto

To cite this article: Iris Hübscher, Laura Vincze & Pilar Prieto (2019) Children’s Signaling of Their Uncertain Knowledge State: Prosody, Face, and Body Cues Come First, Language Learning and Development, 15:4, 366-389, DOI: 10.1080/15475441.2019.1645669

To link to this article: https://doi.org/10.1080/15475441.2019.1645669
Children’s Signaling of Their Uncertain Knowledge State: Prosody, Face, and Body Cues Come First

Iris Hübscher, Laura Vincze, and Pilar Prieto

ABSTRACT
Children achieve their first language milestones initially in gesture and prosody before they do so in speech. However, little is known about the potential precursor role of those features later in development when children start using more complex linguistic skills. In this study, we explore how children’s ability to reflect on their degree of uncertainty develops. A total of 40 Catalan preschool children (and 10 adults) took part in a game in which they had to guess the identity of ten objects which they could touch but not see and then state how certain they were about each guess. An analysis of the children’s reflections on their own uncertainty showed that (a) when preschoolers take a stance, they start to use multimodal uncertainty signals accurately before they are able to self-report their own level of (un)certainty; (b) they use epistemic prosodic and gestural markers before they start using lexical cues; and (c) multimodal cues signaling knowledge state become increasingly more complex over time. These findings suggest that preschoolers’ expression of knowledge states through gestures and prosody reflects an early epistemic sensitivity which develops as they get older.

Introduction
In social encounters it is important for interlocutors to be able to assess and encode the reliability of transmitted information. Efficiently coding and decoding epistemic states is extremely important in a range of social contexts, from everyday interactions like asking for or giving directions to more formalized encounters like business meetings or courtroom statements. As they acquire language, children have to learn how to assess another person’s degree of certainty and how to communicate their own level of certainty. Children’s development of knowledge states has been studied from different research perspectives often without any interaction, in both written and conversational contexts, as well as in behavioral contexts. Research carried out within the field of linguistics and communication has predominantly focused on children’s development of lexical and morphosyntactic markers of uncertainty, with the exception of a couple of studies which have investigated multimodal cues to uncertainty in school children (Swerts & Krahmer, 2005; Visser, Krahmer, & Swerts, 2014). On the other hand, studies in developmental psychology, specifically in the field of metacognition, have been interested in how children develop the ability to monitor their own mental states. While research has found that 2-year-old children are able to signal their ignorance first through behavioral (such as refraining from giving an answer) and gestural cues and shortly after through lexical markers (see Harris, Bartz, & Rowe, 2017 for a review), it is not clear how
preschoolers’ multimodal expression of their own degree of uncertainty develops and in what order the different cues appear. Drawing on both the metacognition and linguistics literature, the present study explores (a) whether preschool children display awareness of their own knowledge state, (b) whether they multimodally signal their epistemic stance by means of facial, gestural, or prosodic markers even before being able to signal it through lexical cues, and (c) how the multimodal expression of epistemic stance changes in the course of development. Below, we first review the literature on the development of knowledge monitoring, then move on to the role that prosody and gesture play in children’s development of epistemic marking, and finally summarize the body of research on how adults and children inform about and signal their epistemic stance.

**Children’s signaling of their knowledge state**

Research on metacognition has addressed the issue of when children become aware of their own level of certainty and are able to report on it (Ghetti, Hembacher, & Coughlin, 2013; Roebers, 2017; Schneider, 2008; Sodian, Thoermer, Kristen, & Perst, 2012; Whitebread et al., 2009 for a review of the different research paradigms). Here we will focus on the paradigm that looks at children’s evaluations of their own knowledge states in the classic total/partial ignorance task (Pratt et al., 1990), in which children are asked whether they know what is inside a box under conditions of total, partial, or no previous knowledge. For example, children are either told or shown what is inside (total knowledge) or they are not told or shown anything (no knowledge, or ignorance condition). Results demonstrate that in the certainty condition, children from age 3 onwards are able to accurately express verbally that they know what is in the box (Pillow, 1989; Pratt & Bryant, 1990; Ruffman & Olson, 1989; Tardif, Wellman, Fung, Liu, & Fang, 2005) and correctly signal their ignorance around the same age (Pratt & Bryant, 1990), or even by age 2 (Rohwer, Kloo, & Perner, 2012). Yet, the picture looks quite different in partial exposure experiments, where children are only exposed to a subset of the items they have to guess. In this case, children struggle until they are about 6 years of age to express their uncertainty verbally. For example, in the study by Rohwer et al. (2012), 4-year-old children were exposed to a set of different objects and then given a closed box containing only one of the objects. The children claimed to know which toy was in the box even though they could not really know for sure which of the several items it was. Only after 5–6 years of age did children correctly deny any knowledge in more than 80% of the trials, and these results remained constant regardless of how many items were included in the set of possibilities. Very young children also seem unable to accurately distinguish between actually knowing and merely guessing. For example, in similar experiments, when asked to indicate which of two boxes contained an object, children could only guess. However, when they guessed correctly, they usually claimed that they knew where the object was, and only when they did not guess correctly did they confess ignorance (Moore, Pure, & Furrow, 1999; Perner & Ruffman, 1995). This tendency to equate guessing correctly with knowing is observable until age 6. This immature understanding of partial knowledge can also be seen when children receive deliberately unclear or ambiguous instructions, such as “Hand me the tall box” when there are two tall boxes, yet judge them to be clear (Beal & Flavell, 1982; Robinson & Robinson, 1982).

Findings in the animal literature (Hampton, 2001; Kornell, Son, & Terrace, 2007; Neldner, Collier-Baker, & Nielsen, 2015; Smith, Shields, & Washburn, 2003), motivated developmental researchers to examine implicit signals of metacognitive abilities displayed in younger children (Balcomb & Gerken, 2008; Call & Carpenter, 2001; Gerken, Balcomb, & Minton, 2011). Studies using tasks that do not require verbal responses have found evidence of sensitivity to their own knowledge state in preschool-aged children and infants (Balcomb & Gerken, 2008; Bernard, Proust, & Clément, 2014; Call & Carpenter, 2001; Gerken et al., 2011; Lyons & Ghetti, 2011; Paulus, Tsalas, Proust, & Sodian, 2014). For example, Call and Carpenter (2001) tested 2- and 3-year-old children using two tubes in one of which food or stickers were hidden. In the total knowledge condition the children saw the food or sticker being hidden, whereas in the partially ignorant condition they did not. When asked whether one or the other tube contained something, it took children much longer to make a decision in the partially ignorant condition. Other studies employing the “opt-out” paradigm (in which child participants are allowed to skip trials...
when they are uncertain about answers, thus avoiding inaccurate responses) have also detected evidence of early metacognitive abilities (Balcomb & Gerken, 2008 among others). Further evidence of children’s early competence has been found in ambiguous reference experiments. In such studies, even children who cannot determine whether an instruction they receive is ambiguous or has an indeterminate referent nevertheless react differently to such instructions, as seen in more frequent eye contact, puzzled expressions, or prolonged reaction times (Patterson, Cosgrove, & O’Brien, 1980; Plummert, 1996; Sekerina, Stromswold, & Hestvik, 2004). Also, it has been shown that 3–5-year-olds signal their knowledge state through behavioral cues, such as selectively skipping trials (Lyons & Ghetti, 2013) or seeking help when reporting they are less certain in a perceptual discrimination task, thus showing the ability to introspect on their confidence and use this introspection to guide their decision (Coughlin, Hembacher, Lyons, & Ghetti, 2015). Side by side with this, researchers have also started to study children’s use of multimodal cues to signal their ignorance. For example, it has been found that when 2-year-old children lack knowledge, they convey ignorance via gestures like shrugs and open-palm hand gestures and only slightly later are they able to verbally acknowledge their unknowingness (see Harris et al., 2017 for an overview). Similarly, Kim, Paulus, Sodian, and Proust (2016) measured how 3- and 4-year-olds monitor their knowledge state through giving verbal judgments and/or by using different gestural cues. In their experiment, they manipulated children’s access to the contents of a box by granting them full access, partial access, or no access to the objects in order to see whether the children would display sensitivity to their own ignorance by refraining to answer when asked to inform another ignorant person. Afterwards, the children were asked to verbally give a judgment on their knowledge state. Importantly, the researchers investigated not only the children’s decisions about informing a third person but also their use of gestural signals of uncertainty (such as tilting their head to one side, shaking their head, shrugging their shoulders, or looking away). The results showed that although children in both age groups were accurately reflecting their knowledge state in the partial and complete ignorance conditions by refraining from informing, they overestimated their knowledge state when asked to verbally report on it. However, interestingly, the 4-year-olds produced significantly more gestural signals of uncertainty than the 3-year-olds, thus displaying early monitoring skills, and gestured more in inverse relation to their degree of certainty.

Early production of gesture and prosody

Most literature from the field of metacognition indicates that children are not able to verbally express their uncertainty (in contrast to certainty and ignorance) until rather late (around 6 years), which raises the question of how children’s lexical, prosodic, and gestural signaling of their (un)certainty develops over time. A number of studies report evidence that gesture comes first in the development of language. Children produce their first deictic gestures between 9 to 12 months, pointing to indicate objects in the environment, for example, well before they start producing their first words (Behne, Liszkowski, Carpenter, & Tomasello, 2012; Camaioni, Perucchini, Bellagamba, & Colonnese, 2004; Liszkowski, Carpenter, Striano, & Tomasello, 2006). Also, once children start to speak, they produce gestures in combination with words (e.g., they point at a cup while saying “mine”) and these gesture-word combinations generally precede the production of two-word combinations (e.g., “my cup”) (Butcher, 2003; Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson & Goldin-Meadow, 2005; Özlüçelik & Goldin-Meadow, 2005). During developmental changes, gesture has also been shown to be an important tool to convey implicit knowledge of a concept that is just emerging. In other words, children who are on the edge of acquiring a new concept use gesture to convey information that clarifies or extends the information expressed in speech, such as when they are learning their first sentences (Butcher & Goldin-Meadow, 2000) or solving difficult cognitive problems (Church & Goldin-Meadow, 1986; Perry, Breckinridge Church, & Goldin-Meadow, 1988).

Yet language development clearly does not end with the acquisition of the first words or the production of sentences. Children learn to manage more complex language skills such as communicating discourse functions and pragmatic messages. While most research has focused on children’s
early production of various types of gestures (whether representational, deictic, or conventional) and their precursor role in lexical acquisition, only a small number of studies have focused on the role of gestures in the acquisition of pragmatic functions. Some of these exceptions have investigated how children learn to express agreement, refusal, and negation (Beauvoir-Hourdel, Morgenstern, & Boutet, 2015; Benazzo & Morgenstern, 2014; Guidetti, 2005). They found that the gestural modality is operational before the verbal modality, with children using conventional gestures such as head shakes and head nods to convey negation and affirmation before they learn to use the corresponding lexical strategies. In relation to knowledge state monitoring and in particular to children’s monitoring of ignorance, as mentioned above, it has been found that 2-year-old children first signal their ignorance through gestures like shrugs and open-palm hand gestures before doing so verbally (see Harris et al., 2017 for an overview). However, little is known about whether children can use different types of gestures and other body cues to signal a partial knowledge state, and whether those cues act as precursors in children’s development of uncertainty signaling skills.

There is increasing evidence that, side by side with gesture, prosody can act as a bootstrapping device in early language development (for an overview, see Esteve-Gibert & Guellaï, 2018; Hübscher & Prieto, 2019). That is, certain types of prosodic features have been shown to guide children’s initial acquisition of word order and syntactic structure (for a conceptualization, see Hirsh-Pasek, Tucker, & Golinkoff, 1996; see also Christophe, Nespor, Guasti, & Van Ooyen, 2003). In addition, there is also evidence that adult use of infant-directed speech, which is characterized by a slower speech rate and exaggerated pitch excursions, helps infants to build the phoneme inventories necessary in order to produce words (Cristia, 2011; Kuhl, Williams, & Meltzoff, 1991; Werker, Dietrich, Kajikawa, Fais, & Amano, 2007). Furthermore, in relation to intonation patterns, it has been found that early intonation patterns appear before the onset of combinatorial speech (Frota, Matos, Cruz, & Vigário, 2016; Prieto, Estrella, Thorson, & Vanrell, 2012). Yet, much less is known about children’s production of complex pragmatic meanings expressed through prosody, and in particular it is not clear how children’s signaling of their uncertain knowledge state through prosody develops.

**Adults’ and children’s production of certainty and uncertainty markers**

The study of epistemic stance, that is, a speaker’s commitment to the truth value of the proposition communicated, has been predominately concerned with the linguistic strategies language users employ to position themselves towards knowledge, with traditional studies mainly focusing on the use of lexical or morphosyntactic features (Conrad & Biber, 2000; Cornillie, 2010; de Haan, 2001; Dendale & Tasmowski, 2001; Heritage, 2012a, 2012b, 2013; Kärkkäinen, 2003; Marin-Arrese, 2011; Zuczkowski, Bongelli, & Riccioni, 2017 among others). Far fewer studies have adopted a multimodal perspective, whereby verbal resources are analyzed alongside voice, facial, and body signals to measure a speaker’s commitment (Borràs-Comes, Roseano, Vanrell, Chen, & Prieto, 2011; Krahmer & Swerts, 2005; Mondada, 2013; Roseano, González, Borràs-Comes, & Prieto, 2016).

When it comes to children, studies on the communication of epistemic stance are even sparser and, again, have mainly looked at the development of lexical/grammatical markers. Previous research has shown that young children start using mental state verbs such as *I know* vs. *I think* to express certainty around age 3 (Diesel & Tomasello, 2001; Shatz, Wellman, & Silber, 1983) and that they start using modal auxiliaries such as *might* to express epistemic modality around 3:6–4 years of age (Papafragou, 1998; Shatz & Wilcox, 1991 for a review). These studies are based on naturalistic data mostly coming from English. Though one study (Lee & Law, 2001) analyzed naturalistic data from Cantonese children, the number of participants was very small (only three children in total). The authors of this study concluded that the direct evidence/uncertainty particle *(lo)* occurred much earlier than the uncertainty/inference particle *(gwaa)*. However, they also noted that while the participants’ mothers used the certainty particle more than 1000 times, they only used the uncertainty particle five times, suggesting a clear influence of exposure. However, findings stemming from bigger and also more controlled data sets are still lacking.
As with the research on adults, the multimodal expression of epistemic stance by children has received little attention, with a few exceptions. Krahmer and Swerts (2005) investigated how Dutch-speaking children (ages 7–8) and adults perceived and produced audiovisual cues to uncertainty. In a first experiment, they applied the Feeling of Knowing paradigm (Hart, 1965) to assess participants' display of their degree of knowledge when answering factual questions. While adults mainly used pauses, fillers (pauses filled with prolonged sounds like *uhh* or *mm*), rising intonation, raised eyebrows, head tilts, and funny faces, children mainly relied on pauses and rising intonation. In a second experiment, adults and children watched the responses that had been recorded in the first experiment and had to judge the speaker's level of uncertainty. In general, the adults' judgments about the recordings were more reliable than the children's. Furthermore, the adults also judged the older children's level of uncertainty more accurately than the younger children's based on their multimodal signals. Also, overall, children in the experiment were better at detecting adults' expressed degree of uncertainty compared to other children's. Later, Visser et al. (2014) investigated the gestural and verbal expressions of uncertainty by children (8–11) using the Feeling of Knowing paradigm as a quiz game in either a collaborative or a competitive setting (two children working as a team vs. two children competing against each other). They found that uncertainty production increased between ages 8 and 11 and that only the older children were affected by the social setting, expressing their confidence more in competitive than collaborative contexts. Finally, Hübscher, Esteve-Gibert, Igualada, and Prieto (2017) tested the ability of 3- to 5-year-old children to detect speaker uncertainty in a comprehension task, which involved the detection of uncertainty in materials that combined from intonational, gestural, and lexical markers. The children, who were divided into two groups by age were asked to indicate by pointing which speaker was uncertain when answering questions. In a between-subjects design, children were either exposed to the lexical condition (where they received lexical and gestural cues to uncertainty) or the intonation condition (where they were exposed to intonational and gestural cues to uncertainty). Within each condition, three different presentation formats were used (audio-only, visual-only and audiovisual) as within-subject variables. The results confirmed that overall, children are able to decode uncertainty states earlier in development when gestural cues are present. Additionally, younger children are better able to infer a speaker’s uncertainty through intonational cues than through lexical marking. Taking this further, the present study aims to test the hypothesis that both prosody and gesture act as forerunners of pragmatic change in children’s signaling of a partial knowledge state. This hypothesis will be tested by looking at children’s production of uncertainty markers and then by performing a cross-sectional analysis.

**The current study**

Based on the findings by Hübscher et al. (2017) and Kim et al. (2016), our hypothesis is that pragmatic gestures and also pragmatic prosody will act as precursors of children’s expression of an uncertain knowledge state. To get a fine-grained picture of how uncertainty signaling develops over the preschool years, we designed an experimental guessing task based on Phan, Meza, Littlewort, Barlett, and Reilly (2010), whereby we manipulated children’s epistemic access to items hidden in a box. Out of 10 objects in total, the children were allowed to see and touch five, while they were not allowed to either see nor touch the other five. The two groups of children (3–4.5 year-olds and 4.5–5 year-olds) were first asked “What is in the box?”. Then, after coming up with an answer, they had to state the degree of certainty they felt about the belief they had communicated. By way of control, a group of adults also carried out the same experimental task.

The aim of the study was threefold. First of all, as a control we tested children’s patterns of knowledge state monitoring by asking them to self-assess their own (un)certainty. Second, the main novel contribution of the study was to assess whether the children could express their uncertain knowledge state through prosodic and gestural means earlier than by means of lexical marking. Third, we sought to determine whether the adults’ marking of uncertainty through multimodal cues would differ from that of the children. In this paper we are interested in children’s signaling of uncertainty and how this is perceived by the interlocutor. Following Morris (2002), we are interested
in all the signals that are being sent out, no matter whether they were intentionally communicated or not, since it is those cues that are being emitted which lead to the overall perception of a speaker’s level of (un)certainty.

We predicted that the children would (1) be more accurate in multimodally signaling their degree of (un)certainty than by self-reporting on it and (2) signal their epistemic stance first and more clearly through gestural and prosodic cues and only later through lexical cues. We also predicted that (3) the adults’ multimodal expression of epistemic stance would be overall more complex than that of the children.

Methodology

Participants

A total of forty children (20 male and 20 female) were recruited at three Catalan public preschools in the metropolitan areas of Barcelona and Girona. Twenty children were 3–4.5 years old ($M = 4.0, SD = 0.47$) and 20 were 4.5–5 ($M = 5.1, SD = 0.53$). Parents were informed about the experiment’s goal and signed a participation consent form prior to their children’s participation in the study. Furthermore, language exposure questionnaires (based on Bosch & Sebastián-Gallés, 2001) were administered to the caregivers in order to ensure that the participating children were predominantly exposed to Catalan (as opposed to Spanish) on a daily basis (mean percentage of overall exposure to Catalan: $M = 88\%, SD = 0.128$). Even though in all public schools in Catalonia the main language of instruction is Catalan, the target schools were chosen based on the high use of Catalan in the school’s catchment area. An additional group of 10 students aged between 19 and 24 ($M = 21.3\%, SD = 1.72$) from the Universitat Pompeu Fabra in Barcelona were recruited to serve as a control group. They were Catalan-dominant, reporting a mean daily usage of Catalan of 89% ($SD = 0.09\%$). These adults signed a written consent prior to their participation in the study and were paid a small amount for participating. This study, including the consent procedure, was approved by the Ethics Board of the Universitat Pompeu Fabra.

Materials

Two sets of objects were used for the guessing game (Figure 1). Set 1 consisted of five common objects (top row in Figure 1: book, spoon, keys, eraser, pen) with which the participants were familiarized before the experiment, and Set 2 consisted of five less common objects (bottom row in Figure 1: tea bag, candle, tape, hazelnut, piece of bark) which the participants did not previously see.

A simple cardboard box measuring 15 cm × 15 cm × 25 cm was covered in wrapping paper for decoration and then two slots were cut at each end, each measuring about 5 cm × 10 cm (see Figure 2). The slot on the opposite side, where the participant would sit, was covered with paper strips, so that the participant could easily put his/her hand inside the box to touch the object, but could not see what was inside.

Experimental procedure

The basic procedure was as follows. An experimenter and participant sat at a small table, facing each other, with the magic box in between. A video camera mounted on a tripod was placed behind the experimenter and set to record the face and upper body of the participant throughout the experiment. On her lap out of sight of the participant, the experimenter held a box containing the ten objects of Set 1 and Set 2. First, the experimenter laid the five objects from Set 1 on the table and told the participant to handle and name each one. That done, the five objects were removed from sight once more. Before starting with the trials, the box was situated on the table and the children were informed that some of the objects that they would be touching inside the magic box would be new whereas others would be the same as they had touched and seen before. The experimenter then took one of the ten objects and placed it in the magic box in such a way that the participant could not see
what she was doing. The experimenter then asked the participant to reach their hand through the strip-concealed slot on their side of the box, feel the object, and say what they thought it was. This procedure was repeated for each of the ten objects, with objects being selected randomly by the experimenter.

To pilot-test the suitability of the target materials, the experiment was carried out first on the ten adult control participants, one at a time, in a quiet room at the Universitat Pompeu Fabra. While these participants were able to guess the objects of Set 1, which they had previously seen, touched, and named, in 100% of the cases, they were able to correctly guess the unfamiliar objects of Set 2 only 50% of the time. Thus while all participants were certain about the Set 1 objects (book, spoon, keys, eraser, and pen), they displayed various degrees of certainty when guessing the other five objects. This pilot study seeming to validate the procedure, the experiment proper involving children was then carried out.

The experiment involving child participants took place in a quiet room at their respective preschools, with each child tested individually. The procedure was identical to that followed in the pilot study. In

Figure 1. Picture of the objects used in the guessing game. Upper row: Set 1. Lower row: Set 2.

Figure 2. Pictures of the curtained slot on the participant's end of the "magic box" (left panel) and the open slot on the opposite end through which the researcher placed objects in the box (right panel).
each trial, immediately upon touching an object the child was asked by the experimenter Què és això? (“What is this?”). Adult participants were asked the same question. However, in the case of the child participants, after supplying an answer to the first question they were additionally asked Com de segur n’estàs? Molt, mig, o poc? (‘How sure are you: very, somewhat, or not very?’). They were told that they could reply to this question either verbally or gesturally by holding out their arm high, at mid-level, or low. The adults were not asked this additional question because it was felt that they signaled their level of certainty reliably enough through their verbal answers, prosody, facial expressions, and gestures.

The experiment lasted around 10 minutes. A total of 100 recorded responses were obtained from the adult control participants (10 adults × 10 items) and 400 recorded responses were obtained from the child participants (40 children × 10 items).

**Data coding**

The resulting total of 500 responses were first given a binary score according to whether the participant had correctly guessed the object at hand (1 = correct, 0 = not correct). Next, in the case of the 400 responses by children, the child’s self-reported degree of certainty (or epistemic stance) was given a binary score, with ‘very certain’ (molt segur) assigned a value of 1, while the two degrees of uncertainty (“somewhat certain” and “not very certain”) were subsumed under the same cover category of “uncertainty” and assigned a value of 0. The two initial levels of uncertainty (“somewhat certain” and “not very certain”) were condensed into one level since we were interested in a range of expression of uncertainty, rather than investigating different degrees of uncertainty separately. Furthermore, this binary separation allows for logistic regression analyses. Note that when the children signaled complete ignorance of the object (i.e., expressed unknowingness), nonverbally and/or verbally, we did not ask them to report on their degree of certainty since they had not come up with a hypothesis. There were 33 such instances (16.5%) for the group of younger children but only two (1%) for the group of older children. These cases were excluded from further analysis since our study focuses on partial knowledge rather than ignorance.

Finally, the contents of the 500 audiovisual recordings were orthographically transcribed and labeled for their lexical, prosodic, and gestural information by means of ELAN (Lausberg & Sloetjes, 2009) by the first author for the children, and by the second author for the adults. The goal here was to have an external assessment of the participants’ epistemic stance independent of their own self-reported input. The two coders met on several occasions during the initial stages of the coding process in order to look at examples, clarify doubts, and refine the coding system. In the final stages the third author was brought into the discussions as well.

Since one of the three goals of this study was to assess the contribution of lexical marking to epistemic stance relative to prosodic or gestural marking, the data were labeled lexically, prosodically, and gesturally, as follows:

**Lexical coding**

Here, the category “lexical” comprises all lexical and morphosyntactic elements, as well as instances of “thinking aloud”, where participants appeared to be verbalizing their thoughts as they touched the object in the box. Thus, in the orthographic transcriptions of the recordings, note was made of all occurrences of epistemic adverbs (e.g., potser “maybe”), epistemic verbs (e.g., crec que “I think”), morphosyntactic cues like the conditional mood (com una xinxeta podríem dir “like a drawing pin, we could say”), references to the thinking process (ai no, està dur “oh no, it’s hard”), and vague language (una cosa així “a thing like”). Vague language has been defined as language which is fuzzy, general, and imprecise, has a low semantic content, and is heavily dependent on shared contextual knowledge for its meaning (Channell, 1994). People are vague either because they lack precise knowledge or because they lack the goal of being precise; in our experiment, the motivation was clearly the former. Such lexical markers were scored as either present (= 1) or absent (= 0).
Prosodic coding

The oral data were prosodically labeled following the Cat_ToBI system (Prieto, 2014). Only nuclear configurations were labeled because it is this part of the contour that typically conveys the pragmatic meaning of the utterance (see Ladd, 1996, among others). The rising and rising-falling nuclear pitch contours (L + H* L%, L* H%, L + H* !H%) were labeled as uncertainty pitch contours, while the falling pitch contours (L* L%, H + L* L%) were labeled as certainty pitch contours. Two other prosodic features were also labeled as uncertainty markers: fillers like *uhh or *mm and final vowel elongations such as in un boliii (“a pe-e-e-en”). The presence of one or more markers was scored as present (= 1) or absent (= 0).

Gesture coding

The coding scheme for face and body signals was agreed among the authors after a series of exploratory analyses of the two datasets (children and adults) (see Figure 3). It is based on Allwood’s (2007) MUMIN multimodal coding scheme, with some modifications, and on Ekman, Friesen, & Hager, (2002) Facial Action Coding System (FACS). To decide which of the array of facial and body cues that participants produced during the guessing task were specifically aimed at signaling uncertainty, we conducted a literature review. Note that when deciding which signals to regard as signals of uncertainty we did not restrict ourselves to solely intentional signals but also included signals that were most likely completely unconscious but which nevertheless offered information about the person’s cognitive state. This is consistent with Morris (2002) view that:

... what matters with gesturing is not what signals we think we are sending out, but what signals are being received. The observers of our acts will make no distinction between our intentional and our unintentional, incidental gestures. In some way, our incidental gestures are the more illuminating of the two, if only for the very fact that we do not think of them as gestures, and therefore we do not censor and manipulate them so strictly. (p. 21)

It is well known that the upper part of the face, namely eyebrows and eyelids, play a significant part in conveying various cognitive processes such as attention, reflection, concentration, or mnemonic effort (Ekman, 1979). Besides these meanings, the eyebrows can also communicate uncertainty and doubt. For instance, according to Eibl-Eibesfeldt (1974) raised eyebrows can convey meanings associated with doubt, questioning, and emphasis during conversation. Frowning is another eyebrow movement associated with uncertainty. According to Maatman, Gratch, and Marsella (2005), frowning and averted gaze are often linked to a speaker’s communication of uncertainty, while Givens (2001) reports a variety of cues associated with uncertainty or doubt, including facial expressions (eyebrow frowns, eye movements, lip-pouting, lip-pursing), head movements (head-shakes, head tilts), and gestures like palm-up open hand gestures, shoulder shrugs, and adaptors. These last, also called self-manipulators, are largely unconscious hand movements typically involving scratching, touching, or covering a part of the face or body which are performed without communicative intention and denote psychological discomfort and anxiety (Givens, 2001).

With regard to the eyelids, De Sanctis (1902) and later on Bitti, Bonfiglioli, Melani, Caterina, and Garotti (2014) observed a marked reduction of the eye aperture and a tightening of the eyelids during mnemonic effort. The study conducted by Bitti et al. (2014) offered further proof that eyebrow raising and squinting (or tightened) eyelids are reliable behavioral cues signaling speakers’ uncertainty. In a question-answer task, the authors differentiate between answers where speakers communicate their lack of knowledge; answers where speakers communicate their uncertainty; and answers where speakers, although uncertain, try to retrieve the information requested by the speaker. In the first case, there is no eyebrow-raising, in the second case the verbal answer is accompanied by eyebrow-raising, while in the third case together with the eyebrow-raising there also occurs a squinting of the eyelids and gaze aversion.

Although not an uncertainty marker per se, a co-speech signal that can sometimes contribute to conveying uncertainty is the nose wrinkle. Usually known in the literature as a cue of disgust.
**Figure 3.** Categories of face and body signals annotated in the data set.

<table>
<thead>
<tr>
<th>Eyebrows</th>
<th>Eyelids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raised</td>
<td>Furrowed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lips</th>
<th>Nose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressed/stretching lips</td>
<td>Corners down</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Head</th>
<th>Shoulders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tilt/Cant</td>
<td>Shake</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Manual gestures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iconic</td>
</tr>
</tbody>
</table>
in situations related to food and bad odors in general (Chapman, Kim, Susskind, & Anderson, 2009; Ekman & Friesen, 2003; Rozin, Haidt, & McCauley, 1999), this gesture typically implies a distancing from the disgusting situation or object. If we take certainty and knowledge as the ideals we all tend to in communication, wrinkling one’s nose while distancing oneself from an unpleasant, uncertain situation can be interpreted as an uncertainty cue.

In general, the lower part of the face has received less attention in the literature focusing on the multimodal communication of uncertainty. Among the few authors who investigated mouth configurations during uncertain answers, Krahmer and Swerts (2005) mention that both adults and children, when uncertain or unknowing, produced “funny faces”, that is, a kind of “marked facial expression”. According to Krahmer and Swerts (2005), a face is perceived as funny if in its composition there is a “funny” configuration of the lips, such as lip corner depression (AU 15), lip stretching (AU 20) or lip pressing (AU 24), manifested in combination with eyes opened wide (AU 5) and possibly some brow movement as well. In our data set, when uncertain, both children and adults pressed and/or stretched their lips, and when they had no idea at all (i.e., were “unknowing”), they lowered the angles of their lips.

As far as head movements are concerned, head tilts have often been found to correlate with expressions of uncertainty and lexical repairs (Heylen, 2005; Lee & Marsella, 2006; Marsi & Van Rooden, 2007), indicating insecurity, shyness, or lack of knowledge (Allwood, 1998). Headshakes have been typically correlated with negation (Ekman & Friesen, 1969; Kendon, 2004; Robinson & Heritage, 2016), intensification, and inclusivity (Goodwin, 1980; McClave, 2001), and more recently with high commitment to beliefs (Vincze & Poggi, 2017). Here we noted that negation headshakes co-occurred with the formulation of a hypothesis and therefore interpreted such headshakes as indicating a high degree of uncertainty in the speaker.

In our corpus we also noticed a large number of shoulder shrugs performed by both children and adults. This signal – shoulders first raised and then going back down to their initial position – is a polysemous item which may assume, depending on the context, quite diverse meanings, such as obviousness, lack of knowledge, or non-commitment (because of either carelessness or powerlessness) (Debras, 2015; Debras & Cienki, 2012; Jokinen & Allwood, 2010). Sometimes, participants (both children and adults) shrugged their shoulders in the initial phase of the guessing, and then came up with a hypothesis in the final phase. Participants hence moved from an initial unknowing epistemic stance to an uncertain one.

In browsing our adult corpus, we noticed that, along with facial and head cues, adults also performed two types of hand gestures: metaphoric and iconic. The Palm Up Open Hand (PUOH) gesture (Cienki & Müller, 2008; Kendon, 2004; Müller, 2004) is a metaphoric gesture performed with palm open and turned upwards, and with fingers extended more or less loosely. The PUOH is a polysemous gesture as well which may take on different meanings derived from two basic domains of action: (1) giving, showing, or offering an object by presenting it on the open hand, or (2) displaying an empty hand to indicate the fact of not having something (Müller, 2004). As pointed out by Müller (2004), actions performed with the open hand are not limited to representing the manipulation of concrete objects but may also illustrate the handling of abstract concepts. Specifically, PUOH may represent the idea that an abstract concept is either visible (i.e., patently true) or missing (i.e., one is showing one’s empty hands to the interlocutor). In other words, depending on the context, this gesture can signal either obviousness, or lack of knowledge. In the context of our experiment, participants’ use of PUOH while guessing objects from the unseen category was clearly intended to communicate lack of knowledge.

A second type of hand gesture found in our adult data was iconic gestures. Participants used these gestures to represent the characteristics of the objects they were trying to guess. In the illustration of an iconic gesture in Figure 3 below, a participant mimics touching the pointed tip of a candle. As widely noted in the literature, iconic gestures seem to play a functional role in lexical access or word retrieval (Beattie & Coughlan, 1999; Krauss, 1998), so the participant may be performing the gesture here to help recall the word he is seeking.
As in the case of prosodic and lexical coding, the presence of one or more markers was scored as present (= 1) or absent (= 0).

**Overall epistemic stance score**

The epistemic stance of the participant was assessed for each trial (10 trials per participant) and assigned one of three values: “certain”, “uncertain”, or “unknowing”. This coder assessment was based on all the information that had been conveyed multimodally; that is, it took into account the all lexical, prosodic, and gestural markers of uncertainty captured by the labeling processes described above. If the child or adult participant signaled one or more of the cues mentioned below, the trial was coded as uncertain, if none of these cues was present, then the trial was coded as certain.

The “unknowing” label was applied when a child participant failed to self-report a degree of certainty about his/her guess. As noted, these responses were excluded from further analysis.

Thus, for the 466 fully coded video-recorded responses that remained, a total of 14 tiers were annotated: (1) the *orthographic tier* containing the transcription of the verbal message and annotation of the lexical/morphosyntactic markers; (2) the self-assessment tier (in the case of children only) containing the child’s own reflection on his/her degree of certainty (high, mid-level, low); (3) the guessed correctly tier (yes or no); (4) the “ToBI” tier containing the intonational transcription; (5) the other prosodic markers tier, where filled pauses and final vowel elongations were noted; (6) the “eyebrow tier” (raised, furrowed); (7) the eyelid tier (squinted/tightened); (8) the lips tier (Pressed/stretched lips, lower lip forward, lip corners down); (9) the nose tier (wrinkled); (10) the head tier (tilted, shaking); (11) the shoulder tier (shrug); (12) the manual gesture tier (iconic, metaphoric); (13) adaptors (self vs. other); and (14) the *coder assessment tier* containing the coders’ interpretation of the speaker’s degree of certainty (certain, uncertain, unknowing).

In order to illustrate children’s multimodal signalling of uncertainty, we will describe the lexical, gestural, and prosodic coding in more detail through a detailed analysis of the guessing response by a 5-year-old child (see Figure 4).
While looking at the experimenter, the child comes up with a guess *una plastillina* (“a play dough”). Lexically there is no uncertainty marker present, however, prosodically the child signals uncertainty by employing a rising intonation. Also, at the end of the guess he marks uncertainty through tightened eyelids as if trying to zoom in with his gaze in the intent of seeing more clearly. The eyelid signal is followed by a protrusion of the lower lip, a facial expression typically occurring in situations of regret (in this case regret of not being able to come up with a certain answer).

**Reliability test**

An inter-rater reliability test was carried out to make sure that the two coders were consistent in terms of their epistemic coding of data as “certain” or “uncertain”. Twenty percent of the database (i.e., 80 items guessed by the children and 20 items guessed by the adults) was randomly selected, with care taken to ensure that objects from Sets 1 and 2 were uniformly represented across speakers. Three coders (all members of the Prosodic Studies Group at the Universitat Pompeu Fabra) were asked to independently annotate this subset of the audiovisual recordings. The sections of video they were exposed to were edited so that they saw the children after they had come up with a hypothesis but before they self-evaluated the certainty of their guess, the idea being that the child’s reckoning of their own certainty might bias the rater’s judgment. The Kappa statistic for rater annotations was obtained. Since four raters in total were involved (for each data set there was one official coder + three external raters), the Fleiss fixed marginal statistical measure was used. Fleiss equally arbitrary guidelines characterize kappas over 0.75 as excellent, 0.40–0.75 as fair to good, and below 0.40 as poor. The fixed marginal kappa statistic obtained for epistemic assessment was 0.79 for the coding of data from the children and 0.86 for the coding of data from the adults. These scores thus reveal high agreement among internal and external raters for epistemic assessment.

**Results**

First, we analyzed the percentage of objects that were guessed correctly across the two conditions (objects from Set 1 vs. objects from Set 2), for all three age groups (children aged 3–4.5, children aged 4.5–5, adults). Results showed that for Set 1 the 3-year-olds guessed the objects correctly in 96% of cases, the older children in 94% of cases, and the adults in 100% of cases. For Set 2, on the other hand, both the younger and the older children only correctly guessed the objects in 22% of cases, whereas the adults guessed correctly in 44% of cases. The results clearly show that while both groups of children and the adults guessed the Set 1 objects in most cases, they both performed much more poorly with Set 2, even though, unsurprisingly, the adults performed better in this latter task.

**Self and coder assessment of children’s knowledge state**

Next, we analyzed the children’s self-reported epistemic stance in relation to their correct and incorrect guesses of the target object, as well as the coders’ assessments in relation to the children’s self-assessments. The goal here was to establish whether the children’s being sure about their guesses correlated with guessing correctly. We also measured their correct guesses against the coders’ rating of their being certain or not. The data set from the children was also broken down by age group (younger or older) to see the effect of age on their performance.

As noted above, for both coders’ and self-reported epistemic assessments certainty was scored as 1 and uncertainty was scored as 0. All the cases where the children did not come up with a hypothesis (when they were completely ignorant) were excluded from the data set. Figure 5 shows the mean proportion of certainty scores as rated by children (self-assessment) or coders (external assessment) across guessing success (0 guessed incorrectly; 1 guessed correctly), broken down by age group. The graphs show that while preschool children were highly accurate in self-assessing their own state (see the correctly guessed items), they are not accurate in self-assessing their own uncertain epistemic
state (see the incorrectly guessed items), and that they tend to overstate their own knowledge in this case. Interestingly, in the case of older children’s incorrectly guessed items there is a contrast between the less accurate self-assessment and the more accurate coder assessment.

In a further step, a Generalized Linear Mixed Model (GLMM) with Mean Rate (Certainty rate) as a dependent variable (Binomial distribution, Logit link) was carried out using IBM SPSS Statistics v24 software. A random intercept was set for Participant. The following variables were set as fixed factors: Age Group (younger children, older children), Guessed Correctly (0, 1), and Epistemic Assessment (self, external). The results revealed a significant main effect of Guessed Correctly ($F(1, 714) = 98,253, p < .001$). Both self-reported and coder ratings indicated significantly more certainty when the children guessed the object correctly than when they did not guess it correctly. Furthermore, there was a significant interaction between Age Group and Epistemic Assessment ($F(1, 714) = 3,905, p = .049$). Interestingly, in incorrectly guessed items, older children self-reported a higher degree of certainty than the external rater perceived ($p < .022$). This result indicates that, at this age, children are not able to properly reflect and verbally report on their uncertainty state: they verbally state that they are certain of their answer, but at the same time their face and prosody convey uncertainty.

**Children’s use of uncertainty lexicon, gesture, and prosody**

We then examined whether children’s and adults’ marking of uncertainty varied in relation to age and cue typology (lexical, prosodic, and body and facial cues). To do so we included all the responses which were assessed as uncertain by both children and raters coders, or by either one or the other. Figure 6 presents the mean proportion of lexical, prosodic, and gestural cues of uncertainty present in the data. Furthermore, Figure 7 presents the mean occurrence of uncertainty cues (prosodic, gestural and lexical) depending on whether the object was guessed correctly or not (i.e. whether it was known or unknown). As it can be seen, children’s signaling of uncertain cues mainly occurred in trials in which they guessed the object incorrectly. The fact that we also find some cases in which

![Figure 5. Mean certainty scores as rated by children (self-assessment) or by coders (external assessment), across younger and older children groups, and across guessing success.](image-url)
uncertainty was signaled even though the item was guessed correctly, is not very surprising, as sometimes one is not so sure but can still come up with the right guess.

A GLMM with Mean Presence (or Proportion rate) as a dependent variable (Binomial distribution, Logit link) was carried out. The following effects were set as fixed factors: Age Group (younger children, older children, adults) and Cue (lexical, prosodic, gestural). A significant main effect of Age

![Figure 6. Mean presence of lexical, prosodic, and gestural cues of uncertainty across the three age groups.](image)

![Figure 7. Mean presence of lexical, prosodic, and gestural cues of uncertainty across the three age groups when guessed incorrectly vs. guessed correctly.](image)
Group was found \((F (1, 612) = 30,884, p < .001)\). Adults produced significantly more cues than the younger and older children, and the older children produced significantly more cues than the younger children. Furthermore, there was a main effect of Cue \((F (1, 612) = 24,302, p < .001)\). There were significantly more prosodic cues than lexical cues \((p < .001)\), significantly more prosodic than gestural cues \((p < .001)\), and significantly more gestural than lexical cues \((p < .001)\). No significant interactions were found between Age Group and Cue.

A close inspection of the results shows, not surprisingly, that adults’ multimodal communication of uncertainty – both lexical, prosodic and gestural – is richer than children’s. Adults not only employ more epistemic markers, whether lexical, prosodic, or gestural, but they also use a greater variety of cues. First, adults employ an array of lexical and morphosyntactic markers such as epistemic verbs or “thinking aloud” talk which do not yet occur in the children’s repertoire. For example, while younger children are not yet producing any lexical markers of uncertainty, in the older children’s group only one occurrence of sembla “it seems” (1% of the uncertain trials) and two occurrences of crec que “I think” (2%) were found. By contrast, in the adult data 16 occurrences of lexical markers (44.44%) were found. In other words, adults employed a total of 16 lexical cues: besides the two mentioned above, they employed other epistemic verbs such as he pensat “I thought” and m’imagino “I imagine” and many cases of no ho sé “I don’t know”. Adults used the thinking aloud strategy in eight cases (22%) and the morphosyntactic strategy in 14 guesses (39%). For their part, children in the older age group employed the thinking-aloud strategy in five cases (6%), and one child in this group even employed the morphosyntactic strategy podria ser una goma “it might be an eraser” (1%).

With regard to prosody, children in the younger age group extensively used the rising and rising-falling uncertainty pitch contours (22 occurrences, 67% of the trials) and the older age group used them even more (65 occurrences, 73% of the trials). Additionally, the younger children started to use fillers (6 occurrences, 19%) and vowel lengthening (3 occurrences, 9%), strategies increasingly put to use over the preschool years given that the older children used fillers in 29 trials (33%) and vowel lengthening in nine trials (10%). In contrast, the adults used uncertainty pitch contours 27 times (75%), fillers 17 times (47%), and vowel lengthening 22 times (61%). However, it is important to highlight that the younger group of children (3–4.5 years old) successfully used target-like prosodic expressions of uncertainty such as rising and rising-falling pitch contours.

Regarding body and facial cues, children in the younger age group typically used their head and eyebrows to signal uncertainty. While younger children produced raised eyebrows in ten guesses (31%), older children produced them in 21 guesses (24%) and additionally produced furrowed eyebrows in two cases (2%). Also, both younger and older children produced head tilts (3 (9%) and 4 (5%) times respectively). The adults’ use of eyebrows and head was more pronounced than the children’s. The adults produced raised eyebrows in 16 guesses (44%) and furrowed eyebrows in 16 guesses (44%), while they tilted and shook their head in eight (22%) and seven guesses (19%) respectively. Furthermore, younger children produced shoulder shrugs in five guesses (16%), older children produced them in one guess (1%), while adults shrugged their shoulders in eight guesses (22%).

There are other facial signals such as nose wrinkles, squinted eyelids, and various lip movements (pressed/stretched lips and corners down) which were only encountered in the older children’s group (but to a very limited extent: three occurrences of nose wrinkles, seven occurrences of squinted eyelids, seven occurrences of stretched lips, and six occurrences of lips corners down. Adults, on the other hand, wrinkled their nose in 11 guesses (31%), squinted their eyelids in 6 (17%), pressed/stretched their lips in 9 cases (26%) and pulled the lip corners down in 7 cases (19%).

Furthermore, adults produced a variety of hand gestures (four iconic gestures (11%) and three metaphoric gestures (8%)), which were not encountered in the child data. It is worth mentioning that the younger children did successfully use some facial/gestural markers of knowledge state, such as raised and furrowed eyebrows, head tilts, and shoulder shrugs, but only the older children are able to employ all the facial and body signals that are in the adult data, like wrinkled nose, squinted eyelids and various lip configurations (pressed/stretched lips and corners down).
To sum up, while children start to signal their uncertainty at age 3 through prosodic and gestural cues before using lexical signals, over the preschool years their signaling of epistemic stance through multimodal signals clearly increases, however it is far from being fully developed at age 5.

Discussion and conclusion

To our knowledge, this is the first study to explore experimentally whether and how preschool children aged 3–5, as compared to adults, express uncertainty multimodally. The study presents a fine-grained analysis of speakers’ (both children’s and adults’) signaling of epistemic stance by means of gestural, prosodic, and lexical patterns while controlling for interlocutors’ (i.e. coders’) awareness and interpretation of the exhibited stance. The three questions that the present study aimed to address were whether preschool children are able to multimodally signal their epistemic stance before they are able to reflect and report on the strength of their beliefs; whether children exploit prosody and gesture to signal their knowledge state before they use lexical items; and finally, how children’s multimodal marking of uncertainty develops over the preschool years and how it differs from adult patterns.

First of all, the results of the children’s self-reporting seen here confirm that young children tend to overestimate their own knowledge state, which is in line with previous research on children’s difficulties in expressing their uncertainty in words (Kim et al., 2016; Rohwer et al., 2012). In cases where children assessed themselves as certain and guessed the object correctly, this corresponded with coder assessments by adult researchers. In the cases where children did not guess the object correctly, both younger and older children reported themselves as being significantly less certain. However, whereas the older children were perceived as significantly less certain by the coder assessments as compared to their own self-assessment, this was not the case for the younger children. This might be an indication that younger children’s uncertainty signaling is only starting to develop, as adult raters were not able to perceive any uncertainty markers, whether prosodic, gestural or lexical. The finding that older children self-reported a higher degree of certainty than was perceived by the raters is similar to that reported by Kim et al. (2016). While both studies reveal that 4–5-year-old children signal their uncertainty through gestural cues, our study shows that, alongside gestural cues, children also make use of specific prosodic cues before they are capable of verbally reporting on their uncertainty. This combination of gestural and prosodic cues allowed raters in our study to perceive the children as uncertain even when the children assessed themselves as certain.

This leads to the second research question, namely the relative amount of lexical, prosodic, and gestural marking in the expression of the uncertain stance across age groups. Our analysis of the utterances produced by 3- and 5-year-old children showed that 3-year-old children used prosodic and gestural cues to express uncertainty, but not yet lexical markers such as epistemic uncertainty adverbs. Comparing these results with previous studies on children’s acquisition of uncertainty markers, the current study shows similar results to those in Swerts and Krahmer’s (2005) study with 7- and 8-year-old children, but at a much earlier age than previously reported. Similar to previous studies, gesture also comes first in children’s signaling of uncertainty, and the present results provide further evidence that gesture is also exploited early by children to mark different pragmatic strategies, as has been found for agreement, refusal, negation, and total ignorance (Beaupoil-Hourdel et al., 2015; Benazzo & Morgenstern, 2014; Guidetti, 2005; Morgenstern, 2014) and also in children’s narrative development (Demir, Levine, & Goldin-Meadow, 2015). Our study confirms and extends this general finding, revealing that multimodal signals of partial knowledge state precede the development of the ability to verbally express this state. Crucially, gesture in our study clearly goes hand in hand with prosody, which is used as an even stronger marker of uncertainty. Prosodic cues of uncertainty such as rising intonation, final vowel elongation, and filled pauses were all found to be widely used in children’s guesses. Interestingly, based on a qualitative analysis of our data, we found that children seem to mark their uncertainty by manipulating their speech phonetically. In a future study it would be interesting to also take into account phonetic factors in children’s marking of uncertainty, such as syllable duration and
intensity, and also measures of voice quality. Doing so might lead to more robust results with regard to how children mark their epistemic stance through prosodic cues.

Comparing our results to previous studies on children’s lexical cues to uncertainty, the current study shows a later appearance of such cues than pointed out by other previous studies (see Papafragou, 1998; Shatz & Wilcox, 1991 for modal auxiliaries); and (O’Neill & Atance, 2000 for modal adverbs). There are several possible explanations for this result. First of all, most previous studies have investigated children’s lexical signaling of uncertainty by using naturalistic observations and often base their conclusions on a very small sample of children. As pointed out by previous studies (see e.g. Lee & Law, 2001), the input provided by caregiver speech might play a big role in children’s production of uncertainty cues, something which has not been investigated in naturalistic investigations which are uncontrolled in nature and where children are in interaction with other adults. In our experiment, this factor was controlled for because the children were not exposed to adults’ use of markers of uncertainty whatsoever. Second, since the current study tested children in one particular experimental setting, it cannot be ruled out that children might produce lexical cues to uncertainty earlier in other settings. On the other hand, the present findings suggest possible directions for developmental and more applied research, because they show that prosodic, facial and body cues seem to be used as early markers of knowledge state, appearing before children are able to use lexical cues such as adverbs and modal auxiliaries.

Comparing children’s versus adults’ use of gesture during speech, it would seem that gestural cues steadily develop over the preschool years and thereafter into adulthood. So while gesture together with prosody comes first in children’s signaling of their uncertainty, not only does the use of gestural and body cues seem to increase significantly over the preschool years, but so does the variety of articulators used. There seems to be a trend in 3-year-old children to mainly use their head, shoulders, and eyebrows to mark their uncertainty, while other articulators, such as eyelids, lips, nose, and hands seem to be used only later on starting at 5 years of age and even more so in adulthood. In general, these findings are in line with the two previous studies on children’s multimodal signaling of uncertainty (Krahmer & Swerts, 2005; Visser et al., 2014). Those studies showed that children’s expressiveness continues to develop between 7 and 8 years and between 8 and 11 and is characterized by the following verbal or visual features: filled pauses (“hmm” or “err”) and pauses in speech, rising intonation (typical of questioning), eyebrow movements, smiles, and marked facial expressions. Thus it does not seem surprising that children’s multimodal expression of uncertainty increases significantly over the preschool years. Nevertheless, our results are novel in the sense that they take a detailed look at the individual gestural articulators to assess children’s early multimodal expression of uncertainty. We have found that besides the eyebrows and an overall marked facial expression, the head too is employed early by children as an articulator to express uncertainty (by tilting it to the side). Also compared to Krahmer and Swerts (2005) study, which analyzed “funny” faces created by a combination of facial movements, in the present study we analyzed facial cues separately, showing that while certain articulators such as the eyebrows were employed earlier by children, uncertainty facial marking involving the nose, lips, and eyelids increased in frequency over the preschool years. Finally, regarding prosodic cues, while Krahmer and Swerts (2005) found that fillers only play a marginal role in children’s signaling of uncertainty, our data show that children employ them already very early and in 30 percent of the uncertain guesses in the older age group.

To sum up, the ability to signal uncertainty to an interlocutor is crucial in communication. Speakers signal their belief state by using multimodal strategies, and interlocutors need to take all of them into account when inferring a speaker’s level of certainty. The gaining of such abilities is therefore an important step in children’s pragmatic development. The current study is the first to investigate the multimodal development of epistemic marking over the preschool years, and it has revealed that both prosodic and gestural patterns are actively exploited in the first stages of the acquisition of the communication of epistemic stance. Children make use of prosody and gesture before they employ lexical cues to signal their stance, and also before they can accurately report on their feeling of knowing.

The results of the present study give rise to further interesting questions. While in the present study we were able to demonstrate that the ability to signal uncertainty through prosody and gesture appears earlier in children, it is not yet clear whether there exists any causal relation between these earlier appearing cues
and the later appearing lexical cues. In a further study it would thus be of interest to test the facilitating role that prosody and gesture may play in children’s uncertainty signaling. Also, an important aspect which should be studied in the future is whether and how certain signals (individual prosodic, gestural and lexical) are correlated and whether there are patterns of co-occurrence of co-localization that can be described (for example, if a child exhibits a nose wrinkle, does the child also exhibit an eyebrow raise on that trial?). Similarly, the correlations between the accuracy of self-reports and the presence of individual markers of uncertainty could be investigated, in order to tap further into the issue of whether and how the accuracy of children’s self-reports correlates with the uncertainty perceived externally by the raters. Furthermore, widening the analysis of uncertainty behavior to more natural and spontaneous contexts taking into account other social behaviors such as requesting for help and asking questions could offer a more comprehensive view of how children acquire the ability to signal uncertainty. Finally, it would also be of interest to study the social aspect of (un)certainty signaling through observing, for example, when children start to use uncertainty or hedging devices as a face-saving strategy in particular social contexts. One method to explore this might be to expose children to different interlocutors and see whether their multimodal signaling of uncertainty changes depending on the person they are talking to.

To summarize, our results suggest that multimodal competence plays an integral part in children’s early development of epistemic stance marking, and might signal important upcoming changes in children’s emerging pragmatic abilities.

Acknowledgments

Preliminary versions of this work were presented at the workshop on Audiovisual Speech Processing and Language Learning in Barcelona (2016), the iGesto conference in Porto (2017), and the Language as a Form of Action conference in Rome (2017). Our special thanks go to Sotaro Kita and Isabella Poggi, who gave us valuable comments and feedback. Many thanks also to the anonymous reviewers whose comments helped to improve the manuscript. Also, the first author would like to express her gratitude to the members of her PhD thesis committee, Judith Holler, Aliyah Morgenstern, and Marc Swerts for their valuable comments on the manuscript. Furthermore, we would like to express our gratitude to the children, parents, and teachers at the Escola Sant Martí, Escola La Farigola del Clot, Escola Pública Dr. Estalella Graells preschools and the undergraduate students who cooperated so generously. Many thanks also go to Judith Llanes and Olga Kushch for helping us with the data collection, to the interreliability raters Evangelia Kiagia, Irene Cañada, and Cristina Sánchez-Conde, and finally to Joan Borràs-Comes for his invaluable help with the statistical analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by the Spanish Ministry of Science and Innovation (grant FFI2015-66533-P “Intonational and gestural meaning in language”), and a grant awarded by the Generalitat de Catalunya (2017 SGR _ 971) to the Prosodic Studies Group. The first author also acknowledges an FPI grant from the Spanish Ministry of Science and Innovation (BES-2013-065019), while the second author acknowledges the Italian National Project PRIN 2012C8BJ3X, “Certainty and uncertainty in biomedical scientific communication”, which partly supported the research.

References


