Acquiring verbal reference:
the interplay of cognitive, linguistic, and general learning capacities

Abstract

Verbal reference is the ability to use language to communicate about objects, events, or ideas, even if they are not witnessed directly, such as past events or faraway places. It rests on a three-way link between words, their referents, and mental representations of those referents. A foundational human capacity, verbal reference extends the communicative power of language beyond the here-and-now, enabling access to language-mediated learning and thus fueling cognitive development. Here, we consider how and when this capacity develops. The existing literature suggests that a nascent form of verbal reference emerges around infants’ first birthdays and becomes increasingly robust by their second. We propose that acquiring verbal reference requires an interplay among cognitive, linguistic, and general learning capacities. We discuss the powerful developmental advantages of verbal reference and suggest new research directions to advance our understanding of how it emerges.
Acquiring verbal reference:
the interplay of cognitive, linguistic, and general learning capacities


Human language permits us to call to mind objects, events, and ideas that we cannot witness directly. We can share information that does not, or cannot, manifest perceptually at the time of learning (Deacon, 1997; Miller, 1990), including historical facts, hypothetical scenarios, and scientific constructs. Intuitive and effortless for most adults, such communication presents a challenge for infants. Consider, for example, everyday utterances like “Where did I leave my keys?” or “Mommy is in the garage, fixing the sprinkler”. In these conversations, and countless others, the referent of the word (keys, sprinkler) is invisible when the word is mentioned. As a result, infants cannot identify the referent of the novel word by mapping a visible object to its name. Instead, if they are to succeed, infants must go beyond the here-and-now: they must somehow link the novel word to a mental representation of the absent object.

This capacity, known as ‘verbal reference’, rests upon a three-way link between words, their referents, and mental representations of those referents (Bloom, 1993; Clark & Bangerter, 2004; Clark & Wilkes-Gibbs, 1986; Waxman & Gelman, 2009). Verbal reference, thus, requires three key constituent components: (1) an ability to establish and retrieve mental representations, (2) knowing that speech communicates about absent entities, and (3) recognizing the referential specificity of words, which implies precise links between words and various aspects or states of...
the represented phenomena (e.g., both “it’s broken” and “it’s a car” can index the same object but one word indicates its state, while the other indicates its kind). We propose that it is the interplay among these constituent components, supported by general learning capacities, such as statistical and cross-situational learning, memory, and causal inferences, that enables the development of verbal reference.

The goal of this paper is to bring together different pieces of evidence on the development of verbal reference and its constituent components to gain a fuller understanding of this fundamental human capacity. First, we review evidence that the constituent components of verbal reference are in place by 12 months, with the first evidence of verbal reference emerging shortly after – at the beginning of infants’ second year. We then describe the developmental advantages of attaining verbal reference, focusing specifically on infants’ lexical acquisition and access to language-mediated learning. Finally, we propose new research directions to advance our understanding of the developmental origins and consequences of verbal reference.

2. The development of verbal reference in infancy.

In reviewing the development of verbal reference, we devote particular attention to infants’ ability to identify word referents that were absent at the time of mention (referred to as ‘absent reference’, ‘fragmented reference’, or ‘displaced reference’ in various publications; we will use the term ‘absent reference’). We focus on absent reference because we are interested in knowing whether infants link words to mental representations, rather than establishing mere associations between words and co-present referents. We therefore propose that the strongest evidence for infants’ comprehension of verbal reference will come from experiments in which a
word and its intended referent are uncoupled in time and space (e.g., “Mommy is at the store”, “Let’s find your ball”).

The first signs of absent reference comprehension emerge around infants’ first birthday (e.g., Gallerani et al., 2009; Osina et al., 2013, 2014). At this developmental moment, infants begin to understand reference to recently hidden objects, but their success is dependent on “visual anchors”. That is, they appear to require visual reminders of currently absent entities to retrieve them from memory (Ganea, 2005; Ganea & Saylor, 2013b). For example, twelve-month-olds can remember the properties of a mentioned hidden object, such as its color, by looking to an anchor of a matching property, such as an occluder of the same color (Saylor, 2004). They also produce communicative behaviors, such as looking or pointing to the anchors. For example, in Gallerani et al. (2009), infants tended to look and gesture towards a door behind which a person was hiding, when their mothers’ mentioned that person (see also Saylor & Baldwin, 2004). Twelve-month-olds’ success with ‘anchored’ absent reference suggests that they have begun to link words and mental representations of absent objects. This success is likely the foundation for infants’ increasing sophistication in communicating about perceptually unavailable entities without needing anchors.

Indeed, by about 14-16 months, infants retrieve mental representations of absent objects or events, even in the absence of anchors (Hendrickson & Sundara, 2017; Luchkina et al., 2020) and successfully update their representations on the basis of verbal descriptions alone. Ganea et al. (2016), for example, showed 16-month-olds two animals near a table and told them subsequently (without showing the scene again) that one of the animals was now on the table. Infants looked significantly longer to the revealed scene if it mismatched the verbal description, suggesting that they had built expectations about the scene based on the language input. Infants
at this age also begin to incorporate their social knowledge to interpret absent reference. For example, they take into account individuals’ prior interactions with objects to interpret their referentially ambiguous requests for hidden objects (Saylor & Ganea, 2007).

As infants approach their second birthday, their comprehension of absent reference becomes sufficiently robust to enable them to establish entirely new representations based on language input (Baldwin, 1993; Ferguson et al., 2014). For example, in Ferguson et al. 19-month-olds identified a referent of a novel word (“dax”), making an inference about its referent’s animacy based on verbal information alone (“the dax is crying”). Infants at this age also update their beliefs about an object’s location or its state (e.g., wet, broken) on the basis of language alone (Ganea et al., 2007; Ganea & Harris, 2010).

In sum, by 24 months, infants successfully form, retrieve, and modify representations of objects or events based on verbal information alone and successfully identify real-world counterparts of those representations when they become perceptually available (see Ganea & Saylor, 2013, for a fuller review of absent reference in infancy). This suggests that infants have formed a stable three-way link between words, their referents, and mental representations of those referents and rely on language to learn new information about perceptually unavailable things. How do infants achieve this? In the following sections, we consider three constituent components of verbal reference (Fig. 1) and their role in the development of the three-way link.

3. **Constituent components of verbal reference.**

3.1. **Establishing and retrieving representations of absent objects or events.**

The three-way link enabling verbal reference requires that infants establish representations of objects or events and retrieve those representations later. The foundation for
moving beyond the here-and-now, this ability emerges in the first months and increases rapidly over the first year of life (e.g., Ganea & Saylor, 2013; Munakata, 2001).

Central to the ability to establish and retrieve representations of absent objects is infants’ memory capacity. The first evidence of infants’ forming a representation of a hidden object have been recorded in 3- to 4-month-olds, who were habituated to a rotating screen that could flip 180°. Later, a box was placed behind the screen, such that the screen would first obscure it, and then be stopped by it, if rotated. Infants dishabituated to the screen rotating 180° but not 120° (consistent with the presence of the box), suggesting that they formed a representation of the box and made inferences about physical impossibility of a 180°-rotation even when they could no longer see the box (Baillargeon, 1987; Baillargeon & DeVos, 1991). At 5 months, infants remember not only the existence of hidden objects, but their properties, such as size and location (e.g., Baillargeon et al., 1989; Baillargeon & Graber, 1987; see Spelke et al., 1992, for a review). Five-month-olds also search for an object in its’ former location after a few seconds of delay from object presentation (e.g., Clifton et al., 1991; Hood & Willatts, 1986; see Baillargeon, 1993, for a review). By 6 months infants can retain a memory of an event for at least 24 hours and by 9-10 months they remember sequences of events over the span of months (Bauer, 2002; Bauer et al., 2000; Carver et al., 2000; Mandler & McDonough, 1995; see Bauer et al., 2010, for a review).

At 12 months, infants do not only have stable representations of objects and events, but can retrieve such representations using conventional symbols, such as pointing gestures. For example, 12-month-olds are surprised if the revealed location of an object hidden behind one of two occluders contradicts the location indicated by experimenter’s points (Gliga & Csibra, 2009). Moreover, 12-month-old infants access their representations of absent objects by looking
at their former locations (‘anchors’) and point to those locations to request the objects (Behne et al., 2012; Bohn et al., 2018; Gräfenhain et al., 2009; Rüther & Liszkowski, 2020; Tomasello et al., 2007; see Liszkowski et al., 2007, for a review). Thus, by 12 months, infants have sufficient memory capacity to maintain representations of absent objects and use gestures, like pointing, to retrieve those representations.

3.2. **Speech communicates about present and absent things.**

Although infants appear to communicate about absent entities with pointing, it is not clear whether they extend the same communicative power to words. Do they also link words to mental representations of absent objects or events?

There is evidence that by 6 months, infants expect that language (but not other human-produced sounds, such as coughing) involves a transfer of information about unobservable mental states (Cheung et al., 2012; Vouloumanos et al., 2014). In Vouloumanos et al. (2014), for example, infants first observed an actor (the ‘communicator’) try and fail to reach for one of two objects. She then turned to another actor (the ‘recipient’), who could reach both objects but did not know which object the recipient wanted, and produced an utterance. Six-month-olds looked significantly longer when the recipient handed the speaker the object she did not want, thus indicating that they expected the communicator’s utterance to transmit her preference to the recipient.

Such an expectation requires that infants’ make causal inferences and realize that others’ mental states govern their actions. Upon observing the communicator try but fail to complete an action, infants must infer the her intention. When the communicator directs her speech to the recipient, who completes the intended action, infants must infer that the communicator’s speech and the recipient’s action are causally connected. Indeed, infants begin to understand that human
actions, such as pushing, cause inanimate objects to move by at least 4 months (Leslie, 1982). Soon after, by 5 months, infants realize that their own communicative actions elicit social responses (e.g., Goldstein et al., 2009). By 6 months, infants appear to attribute goals to others’ actions (e.g., Woodward, 1998) and take into account others’ access to information to interpret their actions (e.g., Luo & Johnson, 2009; Southgate & Vernetti, 2014). This evidence suggests that infants in Vouloumanos et al.’s studies indeed understood that the two actors communicated about intentions and preferences via speech. Combined with the ability to establish mental representations and retrieve them via words, this understanding enables infants to make an inferential leap and realize that speech is a conduit for accessing mental representations and communicating about them.

In sum, by 6 months, infants recognize that speech communicates not just about present but also about absent or unobservable entities. This recognition becomes even more robust by twelve months, when infants realize that speech is a better way to learn about mental states than even direct observations of someone’s actions (Vouloumanos et al., 2012). Moreover, 12-month-olds recognize that words are conventional forms that are shared among speakers of the same language (Buresh & Woodward, 2007; Colomer & Sebastian-Galles, 2020; Henderson & Woodward, 2012) and extend the same communicative properties to foreign speech (Vouloumanos, 2018).

What remains unclear is how and when infants realize that words not only refer to objects and events (or their mental representations), but offer even greater precision to refer to objects as individuals, as members of kinds, or their distinct properties (color, parts) and states (wet, dirty).

3.3. Referential precision of words.
Referential precision, a fundamental feature of language, is not available in gestural indices, such as pointing. For example, pointing can direct attention (there is a running man), but only words provide precision about what can communicate which aspect specific aspects and properties of that scene (the running man is injured, the running man is my friend). When do infants’ links to words become more precise than their links to pointing and other indexical gestures (see Novack & Waxman, 2020, for a review on infants’ gestural communication)?

This precision is evident by 12 months, when infants are sensitive not only to whether objects are named, but how they are named (Waxman & Braun, 2005). For example, applying the same word consistently to a set of distinct individual objects (e.g., images of different fish) highlights commonalities among them and promotes object categorization (Althaus & Plunkett, 2016; Ferguson & Waxman, 2017; LaTourrette & Waxman, 2020; Waxman & Markow, 1995). Applying a different word to each object highlights distinctions among them and promotes object individuation (Ferguson & Waxman, 2017; LaTourrette & Waxman, 2020; Waxman & Braun, 2005). By 14 months, infants distinguish novel words presented as count nouns (e.g., “this one is a blick”) from those presented as adjectives (e.g., “this is a blick one”), and map novel count nouns, but not adjectives, specifically to object kinds (e.g., Booth & Waxman, 2009; see Waxman & Gelman, 2010, for a review).

This distinction requires that infants have mechanisms allowing them to discover different kinds of words (e.g., nouns, verbs, adjectives) and map them to different kinds of meaning, even if they reference the same object. Such mechanisms are likely provided by general learning capacities, such as statistical learning, that emerge in early infancy and support infants’ discovery of structure in the language input (see Smith et al., 2014, for a review). For example, by 8 months, infants successfully use statistical regularities to identify words in a
continuous stream of speech (Aslin et al., 1998; Jusczyk & Hohne, 1997; Saffran et al., 1996; see Saffran, 2020, for a review). At this age infants also recruit statistical learning to distill abstract rules, such as sound patterns, from auditory input (Ferguson & Lew-Williams, 2016). By at least 12 months, infants use statistical learning to learn grammatical rules (e.g., Gomez & Gerken, 1999). Infants also identify the correct referents of novel words by cross-referencing word-referent co-occurrences over different situations (known as cross-situational comparison or cross-situational learning; e.g., Ichinco et al., 2009; Smith & Yu, 2008; Trueswell et al., 2013).

In sum, by the beginning of the second year infants recognize that unlike pointing, words are a precise referential instrument that can communicate about specific parts or aspects of present and absent objects or events.

4. **Bringing it all together: what does the infant gain with verbal reference?**

Thus far, our discussion has centered around the constituent components and general learning capacities that enable the development of verbal reference. In this section we focus on downstream advantages that attaining verbal reference contributes to cognitive and linguistic development, including advances in lexical acquisition and access to language-mediated learning.

First, the emergence of verbal reference streamlines infants’ ability to leverage relevant cues to word meaning. To establish precise word meanings, infants often need to integrate multiple cues. For example, when there are multiple candidate referents of a novel word, infants must rely on extralinguistic referential cues, such as the speaker’s eye gaze (e.g., Pruden et al., 2006) to infer the correct meaning of the word. This challenge is further complicated by limited informativeness of each individual cue to meaning. For example, corpus analyses of video-
recordings of parents interacting with their 6-to-18-month-old infants revealed that no single social signal (eye gaze, pointing, hand position) is sufficient to resolve this problem of reference, especially for younger infants (Frank et al., 2013). Infants begin to integrate multiple cues to meaning at around 12-14 months (e.g., Beier & Carey, 2014; Pruden et al., 2006) – the same age when the first evidence of absent reference comprehension is observed. Thus, verbal reference may beget infants’ ability to incorporate multiple sources of relevant information to infer word meanings.

Evidence from a computational approach is also consistent with the possibility that verbal reference promotes cue integration. Frank et al., (2009), for example, developed a word-learning model that incorporates referential intention. The words speakers utter do not just depend on the set of plausible referents, but are a function of what the speakers intend to say and how those intentions can be translated into the language they are speaking. Tested on the CHILDES corpus (MacWhinney, 2000), this model outperformed comparison models that did not include referential intention and accounted for rapid word learning phenomena in infants.

Finally, the argument that verbal reference promotes cue integration is also supported by evidence from adult word-learning studies. Wang and Mintz (2018) demonstrated that word-referent mappings are better retained by adults when they are told that words co-occurring with images refer to those images (as opposed to simply marking them as a class of “word” objects). These findings demonstrate that even experienced language users’ ability to take advantage of the available cues to meaning is affected by their beliefs about the referential nature of the utterances they hear. For beginner language learners verbal reference becomes the fulcrum for leveraging social, contextual, and distributional cues.
Second, attaining verbal reference enables language-mediated learning. Recall that by the end of the second year, infants successfully learn novel facts about familiar objects from language. For example, Ganea and colleagues told 22-month-olds that “Lucy (a familiar toy animal, who was hidden from view during this conversation) got wet”; later, when children were asked to find Lucy, they chose a wet toy, rather than the dry one that they had previously seen (Ganea et al., 2007). By 30 months, children can create detailed novel representations of entirely unfamiliar objects (e.g., a novel creatures’ body shape, color, and habitat) and identify those objects when they become perceptually available (Saylor et al., 2016). This capacity for language-mediated learning is critical to young children’s further cognitive development, as it allows them to gain access to the knowledge about things they do not experience directly.

Notably, in many cases, infants’ major source of language input is overheard rather than child-directed speech (e.g., Sperry et al., 2019). Realizing the referential nature of words may be the mechanism enabling infants to recognize the informational value of non-directed speech. For example, Floor and Akhtar (2006) showed that 18-month-olds successfully learn novel words both when addressed directly and when overhearing a conversation between two adults. This capacity further extends young children’s learning opportunities and allows them to acquire vocabulary and fact knowledge even in environments that do not provide ample child-directed speech (see also Schieffelin, 1985; Silva et al., 2010).

5. Conclusion, future directions.

In this review, our goal was (1) to better understand the processes underlying the development of verbal reference in infancy by tracing the development of its constituent components and general learning capacities, and (2) to illustrate the relation between verbal
reference, infants’ success integrating multiple cues to meaning, and the development of language-mediated learning. We have shown that verbal reference is sufficiently robust to support language-mediated learning by the 24-30 months of age. Building on this foundation, we propose several new directions of research that will help address outstanding questions about the development of verbal reference.

First, does verbal reference require more than the development of its constituent components? Our literature review shows that most of these components are in place by 9 months. Why then does absent reference, a hallmark of verbal reference, emerge later, by 12-14 months and remain fragile until the end of the second year? Furthermore, even though each constituent component has been observed in other species (e.g., Bohn et al., 2015; Flemming et al., 2013; Kaminski et al., 2004; Lyn & Savage-Rumbaugh, 2000; Maginnity & Grace, 2014; Wobber et al., 2014), verbal reference appears to be unique to humans. Even great apes trained in sign language do not respond to messages about the states or locations of absent objects (Savage-Rumbaugh et al., 1993). What enables human infants to bring the constituent components together and develop verbal reference?

The answer may lie in infants’ experience with social interactions, especially those involving language. Infants appear to infuse many signals, including non-linguistic ones, with referential status when they are presented in reciprocal social (also referred to as ‘socially contingent’; see Csibra, 2010) interactions. For example, Ferguson and Waxman, (2016) show that sine-wave tone sequences, presented in conjunction with objects, fail to support categorization in infants. However, when these sequences are first demonstrated embedded in a reciprocal social interaction, they become infused with language properties – enough to permit sine-wave tone sequences to support object categorization (as words do; see also Taurzin &
Gergely, 2018; Waxman & Markow, 1995). While infants show this remarkable success in a lab setting, it is likely that this effect is short-lived. To establish a robust understanding that a given signal is communicative and referentially specific, infants likely need a large amount of frequently recurring observations. It is also possible that infants need to pay particular attention to reciprocal social interactions. Current evidence suggests that such interactions become privileged by infants’ attention by the end of the first year (see Thiele et al., 2021). Thus, it is possible that by 12-14 months, but not earlier, infants have accumulated sufficient experience with reciprocal social interactions to infer that language is a referentially precise communicative system. Alternatively, it is possible that verbal reference emerges earlier. Some evidence hints at this possibility. An electrophysiological investigation found that even 9-month-olds detect semantic mismatches if they hear a word in the absence of an object (e.g., ‘Here comes the duck!’) and then see an object with a different name (e.g., a cookie; Parise & Csibra, 2012). At 9 months infants also build expectations about the number of objects of different kinds based on the number of distinct labels they hear before the objects are revealed (Dewar & Xu, 2007; Xu et al., 2005). These findings may reflect infants’ referential understanding of words. Future research designed to test these hypotheses will promote our understanding of the developmental origins of verbal reference.

Second, what enables us to represent abstract concepts and communicate about them? It is possible that absent reference comprehension scaffolds this capacity by enabling us to establish representations of never-seen objects or events from language input alone. Recall that in Ferguson et al., (2014) 19-month-olds were able to identify a referent of a novel word they learned in the absence of any objects based on the animacy of a verb argument. By 24 months, children impose constraints on the meaning of entirely novel words from their linguistic
properties, such as verb transitivity (Arunachalam et al., 2013; Ferguson et al., 2018; Yuan & Fisher, 2009). It is possible that with more experience, children eventually move beyond concrete nouns and verbs and can use linguistic context to form representations for entities that have no stable perceptual form. Indeed, Della Rosa et al., (2010) showed that unlike concrete concepts, abstract concepts tend to be acquired from linguistic contexts: the more abstract the concept, the less likely perceptual experience is to shape it (see also Borghi et al., 2019). Future investigations designed to link infants’ command of absent reference and their subsequent ability to learn abstract concepts will shed light on the role of verbal reference in the development of abstract knowledge.

Finally, a productive avenue of future research would be to test our hypotheses about the role of verbal reference in cue integration and language mediated learning. Is there a causal link to verbal reference? Current evidence is based on correlational evidence and requires empirical testing.

To conclude, the development of verbal reference, which requires an elaborate interplay between cognitive, linguistic, and general learning capacities, is a paramount milestone in language acquisition achieved in the first 2 years of life. Verbal reference enables us to share information across time and space, and to extend our representational capacity beyond anything we have witnessed directly. With verbal reference, we can communicate about abstract concepts like justice and learn about subatomic particles like quark or molecular structures like the DNA’s double helix (aka ‘hard words’; see Gleitman et al., 2005). Evidence concerning the development of verbal reference, however, remains sparse. We look forward to future avenues of research designed to trace how it emerges and unfolds in first years of life.
References

https://doi.org/10.1111/desc.12358

https://doi.org/10.1080/01690965.2011.641744

https://doi.org/10.1111/1467-9280.00063


https://doi.org/10.1016/S0885-2014(89)90040-3

https://doi.org/10.1016/S0885-2014(87)80014-X


