Available online at www.sciencedirect.com

ScienceDirect

Lingua xxx (2018) xxx–xxx

Lingua

www.elsevier.com/locate/lingua

Intentional inference during infants' observational word learning

Xiaobei Zheng [a], Xiangzhi Meng [b,*], Yinglin Ji [a,*]

[a] Research Centre for Language and Cognition, Shenzhen University, Shenzhen, China
[b] School of Psychological and Cognitive Sciences, and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

Received 4 August 2017; received in revised form 18 February 2018; accepted 18 February 2018

Abstract

Word learning theories propose a dynamic development mechanism in which infants recruit statistical and perceptual cues at an earlier stage and social cues at a later stage, but it still remains unclear what promotes transition between the two stages. Recent studies have shown that the ostensive cue promotes infants' social learning, de-emphasizing, and even denying, the contribution of their development of teleological cognition. The present study was operationalized to explore the influence of the bare teleological inference by teaching infants words in a monadic observational context where they were the only participants, and where no ostensive cues were provided by other agents. We utilized Woodward's (1998) experimental scheme and presented 12- and 17-month-old infants with a hand or a rod repeatedly approaching and grasping or touching a pair of objects, paired with two novel word labels respectively. In the test phase, where the position of objects was changed, two word–object associations were either switched or not. No detection of the changes was found for the non-intentional rod in both age groups. In contrast, with the intentional hand agent, the older group showed detection of the referential relation between word labels and objects, but the younger group simply showed sensitivity to the position. These findings provide an important insight into issues such as how infants used their teleological inference in word learning, and how multiple cues are integrated during the transition of infant word learning from statistic and perceptual learning to social and interactive learning.

Keywords: Infants; Word learning; Intentional inference; Observational learning; Habituation

1. Introduction

Infants acquire their first words at the end of their first year of life. The speed is slow at the beginning of their word learning, but dramatically increases at around 18 months of age (Bloom, 1973; Dromi, 1987; McCarthy, 1954; Nelson, 1973). With the acceleration of their learning processes, the dominant way that infants acquire a word changes from statistical learning to social learning (Hollich et al., 2000; Golinkoff and Hirsh-Pasek, 2006). The social–pragmatic view of word learning suggests that joint attention is a core part of the latter mode of learning. Infants map a word label to an entity to which they and adults jointly attended (Baldwin, 1993; Bruner, 1983; Tomasello, 1992). Infants who are more exposed to word labels in episodes of joint attention have a larger size of productive vocabulary (Tomasello and Farrar, 1986). They

* Corresponding authors.
E-mail addresses: mengxzh@pku.edu.cn (X. Meng), yinglin.ji@szu.edu.cn (Y. Ji).

https://doi.org/10.1016/j.lingua.2018.02.007
0024-3841/

also recognize the target of both nonverbal deictic referential gestures (Liebal et al., 2009, 2010) and verbal labels (Akhtar et al., 1996; Tomasello and Akhtar, 1995) on the basis of joint attention.

The episode of joint attention is considered to be grounded in referential intention, which plays an important role in word learning (Sabbagh and Baldwin, 2005). Theoretically, linguistic communication is considered as a cooperative activity between the speaker and the listener (Clark, 1996). In cooperative activities, such as working together to bake a cake, there is an accepted purpose or goal towards which participants jointly work (Grice, 1989). So, cooperative participants should have some ability to understand the goals, intentions, and perceptions of their partner. Empirically, evidence has also shown that while adults attempt to provide appropriate words for entities within the attentional focus of infants, the infants themselves also coordinate their attention to the intended referent guided by the adult's focus of attention (Baldwin, 1991; Baldwin et al., 1996).

In light of the above referential intention account, infants should be able to map word labels to the intended entities, even though the learning occurs in an observational context. In addition, field studies find that infants learn words through overhearing in a polyadic situation where two additional interlocutors join in a conversation and each express their communicative intent to the other. In some cultures, observational learning is dominant in infant learning (de Leon, 1998; Pye, 1986; Rogoff, 2003). Some empirical evidence also demonstrates that infants do acquire a word label via observation, and their learning performance is not measurably better in face-to-face interaction than in observation (Akhtar, 2005; Akhtar et al., 2001; Floor and Akhtar, 2006; Gampe et al., 2012; see Shneidman and Woodward, 2016 for a review). For example, in Gampe et al.'s (2012) study, the researchers addressed either infants or a second experimenter and introduced a novel object with a word label. The 18-month-old infants acquired word labels equally well in both interaction and observation conditions.

However, one problem of the aforementioned studies is that infants were all involved in an ostensive context where rich cues, such as eye contact and contingent responses, were provided. The ostensive context is believed to be a prerequisite for the cultural relevance of the speaker's communicative intent. Specifically, in a cooperative communication activity, participants have to express these purposes or goals demonstratively (Sperber and Wilson, 1986). Seen in this light, the essence of the social learning may be the ''overtness'' or ''openness'' specified by the ostensive cues. The cognitive mechanism that infants use in their learning may be their inherent sensitivity to the ''openness'' of conversation. For example, Moll et al. (2011) found that 2-year-old children could adjust the referential domains of the speaker's utterance, depending on whether their interlocutors are co-present. This suggests that the ''openness'', which is normally signalled via eye contact in the situation of co-presence, influences children's cognitive representation of the speaker's intended meaning.

In the past decade, empirical evidence shows that intentional inference does not contribute much to infants' inductive learning and generalization. Indeed, infants understand that the speaker usually gazes at the entity about which he or she is talking even before they truly understand the intention of the speaker (Perner and Ruffman, 2005). Evidence shows that with only teleological inference, infants can learn inductively and will simply interpret particular entities individually. Only when infants are presented with ostensive cues will they generalize their interpretation of the agent's goal and intention to the congeneric population (Egyed et al., 2013; Gergely et al., 2007) and entities (Butler et al., 2015; Futó et al., 2010; Topál et al., 2008). For example, in Egyed et al.'s (2013) study, after watching an adult consistently demonstrating two opposite dispositional attitudes respectively on two different objects, infants learned the objects' qualitative property in a generalized way based on object rather than individual when the adult additionally provided eye-contact and greeting in the ostensive context.

For the same reason, without ostensive cues, infants may not acquire a word, given that a word label is a generalized concept acquired during infants' inductive learning and generalization (Gelman and Markman, 1986). The context where the speaker has a teleological intention that is not ''open'' may be not intuitively felt to be communicatively purposeful.

It has not been reported whether infants learn words with a teleological intention, because the teleological context has not been distinguished from the ostensive context in empirical research until the above generalization studies were reported recently. In previous investigations of observational word learning, both teleological and ostensive elements were co-present in the episodes, even though those ostensive cues were not directed to the infants. The ostensive cues expressed to another recipient may contribute to infants' understanding of different perspectives of individuals and generalization of such understanding during their learning.

In this context, the present study was particularly interested in whether infants could learn words in monadic observation in which they were the only participants with communicative competence and no ostensive cues were offered by other entities. Specifically, we aimed to test whether infants could form a reliable referential interpretation, rather than a fragile perceptual link, with the bare teleological information. If they did so, it can be implied that the word learning with social cognition is triggered by the development of intentional inference and teleological cognition. Otherwise, as suggested by the aforementioned studies, infants' innate sensitivity to ostensive context is what urges them to transform their dominant mode of word acquisition from statistical to social learning.

A considerable number of teleological inference studies have been conducted in such monadic contexts and have found that infants demonstrate their ability to infer another individual's motivations and to ascribe goals to the agent's

Please cite this article in press as: Zheng, X., et al., Intentional inference during infants' observational word learning. Lingua (2018), https://doi.org/10.1016/j.lingua.2018.02.007