Neural Networks 106 (2018) 194–204

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Social babbling: The emergence of symbolic gestures and words

Laura Cohen *, Aude Billard

Learning Algorithms and Systems Laboratory, School of Engineering, EPFL, Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 2 October 2017
Received in revised form 20 June 2018
Accepted 27 June 2018
Available online 18 July 2018

Keywords:
Language development
Symbol grounding
Communication
Social cognition
Reinforcement learning
Parental responsiveness

ABSTRACT

Language acquisition theories classically distinguish passive language understanding from active language production. However, recent findings show that brain areas such as Broca's region are shared in language understanding and production. Furthermore, these areas are also implicated in understanding and producing goal-oriented actions. These observations question the passive view of language development. In this work, we propose a cognitive developmental model of symbol acquisition, coherent with an active view of language learning. For that purpose, we introduce the concept of *social babbling*. In this view, symbols are learned in the same way as goal-oriented actions in the context of specific caregiver–infant interactions. We show that this model allows a virtual agent to learn both symbolic words and gestures to refer to objects while interacting with a caregiver. We validate our model by reproducing results from studies on the influence of parental responsiveness on infants language acquisition.

1. Introduction

In this last 15 years, an important amount of energy has been spent to adopt an embodied cognition perspective of language acquisition. We will try to demonstrate here that those approaches focus mostly on a passive role of the learning agent during symbol learning. This classical view posits that infants learn words by listening to a knowledgeable caregiver, while observing and/or manipulating objects. However, infants are able to express meanings well before they produce their first words. For example, babbling infants use vocalizations to request objects and actions from their caregiver. During this phase, their verbal utterances are not yet imitations of adults' words (see Halliday, 2006, in particular chapter one on infancy and protolanguage). This suggests an active and creative role of the learner in the acquisition of language. In this view, the learner actively tries to achieve specific goals by using language as a tool. This *functionalism* of language is defined by Bates et al. as the idea that the form of natural language is created, acquired and used in the service of functions (Bates, Thal, & MacWhinney, 1991). In this work, we propose an alternative solution to state of the art approaches of symbolic word acquisition, based on a functionalist view of language. This approach is inspired by findings on shared brain areas between goal-directed behavior and language (Fazio, Cantagallo, Craighero, DAusilio, Roy, Pozzo, et al., 2009; Nishitani, Schürmann, Amunts, & Hari, 2005), as well as developmental observations (Halliday, 2006).

Language development is typically considered to begin between 6 and 10 months. At this early age, Halliday refers to the child's verbal utterances as protolinguistic (*Phase 1*) (Halliday, 2006). He describes some distinctive properties of this child's first language. Each utterance consists of only one element, that does not correspond yet to words, as utterances do not share phonological properties with the English lexicon. They are not imitations of adults words, but are rather spontaneously created by the child. Other studies report that the phonological properties of this first stage speech are dominated by biological factors, such as the nature of the vocal tract (Vihman, Ferguson, & Elbert, 1986). This protolinguistic system is initially only personal, as the utterances are individual-specific. It is later replaced by words of the child's first language. However, Halliday considers these utterances as already part of a language, as they are composed of constant content-expression pairs: a specific vocal posture (the expression) is uniquely associated with a specific meaning (the content). This requirement is referred to as the *systematicity* of language. Other studies report that protolinguistic infants in their first year already possess a sound-meaning system. For example, D'Odorico and Franco (1991) shows that different infants produce selectively specific vocalization types depending on the communication context. Furthermore, babbling infants are able to express distinct communicative functions before their first words (Karousou & López-Ornat, 2013). In particular, they use vocalizations to regulate the action of their caregiver toward concrete goals. Halliday describes this first function of language as imperative: infants start using language for the caregiver to retrieve distant objects for them (Halliday, 2006). The informative function of language appears only much later, around 22 months. Halliday defines *functionality* as the second requirement of a language, that is the fact that each meaning is derivable from a function. The development of

* Corresponding author.
E-mail address: laura.cohen.94@gmail.com (L. Cohen).

https://doi.org/10.1016/j.neunet.2018.06.016
0893-6080/

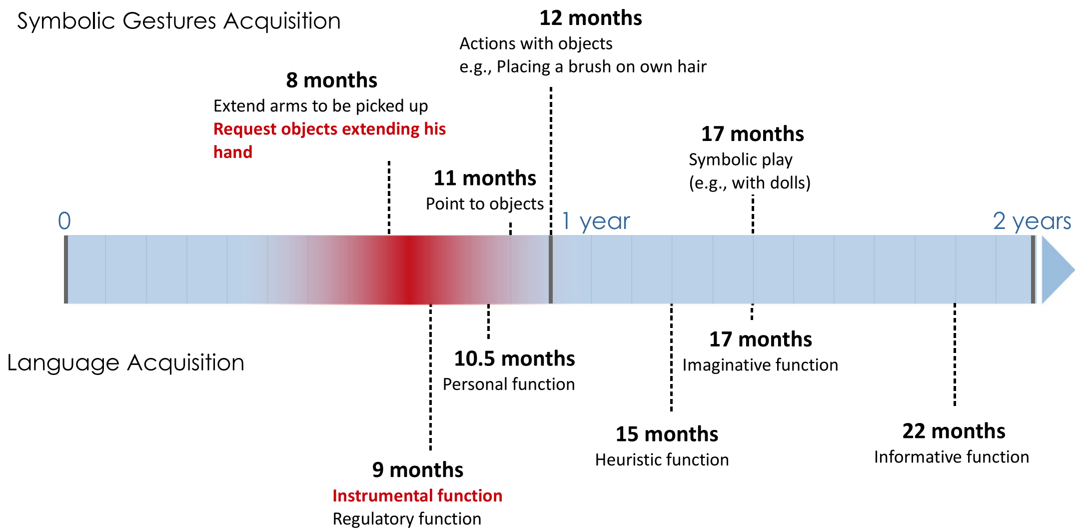

Fig. 1. Developmental stages of symbolic gestures (top) and language (bottom) from birth to 2 years of life, according to Halliday (2006) for language and Caselli et al. (2012) for gestures. The period of interest of this work is displayed in red. The functions of these early stages of language are the following: to satisfy a desire (instrumental), to make requests (regulatory), to express thoughts and feelings (personal), to explore the environment (heuristic), to play "Let's pretend" (imaginative) and to communicate information (informative).

symbolic gestures in early childhood follows a similar pattern, often shortly preceding its verbal equivalent (Bates & Dick, 2002; Caselli, Rinaldi, Stefanini, & Volterra, 2012). For example, a child starts requesting objects by extending the hand shortly before producing verbal utterances of the same kind; see Fig. 1 for details of symbolic gesture and language developmental milestones. This suggests that language builds upon general cognitive abilities such as gesture production and understanding, that would rely on common neural structures (Bates & Dick, 2002). In a functionalist view of language, a strong parallel emerges between language and goal-directed actions: by means of a speech act (Austin, 1975), the infant tries to reach a concrete rewarding goal. Following the same line of thought, language has been related to a tool by numerous authors (Borghi, Scorolli, Caligiore, Baldassarre, & Tummolini, 2013; Mirolli & Parisi, 2011; Nazzi & Gopnik, 2001; Tylén, Weed, Wallentin, Roepstorff, & Frith, 2010). However, this stance is taken metaphorically, as language is mainly thought as a mental tool that enhances our cognitive abilities by means of providing representational capacities. We argue here that in the course of development, language is initially a concrete tool, i.e., with direct consequences on the state of the world. In this sense, it follows the same principles as the development of goal-directed actions. Strong evidence to support this claim lies in the similar neural pathways between language and goal-directed action processes.

1.1. Language and the brain

Two principal areas in the brain are classically associated with language, namely Broca's area and Wernicke's area. Broca's region is located in the left ventro lateral prefrontal cortex (PFC), and includes Brodman's areas (BA) 44 and BA 45, extending anteriorly to BA 47 (see Fig. 2.A). Wernicke's region is located in the left superior temporal cortex (BA 22 and 42). Broca's area was initially associated with language production (i.e., phonology), while Wernicke's area was considered the center of language understanding. More recently, however, various studies have demonstrated a wider involvement of Broca's area in numerous cognitive and perceptual tasks. A functional specialization of Broca's sub-region is proposed, with phonologic (BA 44), syntactic (BA 45/BA 44), and semantic functions (BA 47/BA 45) (Goucha & Friederici, 2015; Hagoort, 2005). In this view, Broca's area mediates a cascade of activation from the most ventral area (BA 47) to premotor (PM) and motor cortices, directly adjacent to BA 44. Adjacent to BA 47, the orbitofrontal part of the PFC (OFC) is directly connected to the limbic system, and thus to reward centers in the brain. For example, in Gilbert and Fiez (2004), authors observed an effect of reward (monetary incentive) on activation of BA 47 during a word memorization task. Other findings converge in attributing an important role of BA 47 in reward-based decision-making cognition (Elliott, Friston, & Dolan, 2000; Ernst, Nelson, McClure, Monk, Munson, Eshel, et al., 2004; Koch, Schachtzabel, Wagner, Reichenbach, Sauer, & Schlösser, 2008).

Interestingly, similar regions are involved in goal-directed behavior, as the frontal lobe exerts control over behavior in a pathway that begins in the OFC and from there projects to PFC, PM and finally motor cortex (Kandel, Schwartz, Jessell, Siegelbaum, & Hudspeth, 2000; Ridderinkhof, Van Den Wildenberg, Segalowitz, & Carter, 2004). In the same line, neuroanatomical investigations indicate a functional link between language and action areas (Fazio et al., 2009; Nishitani et al., 2005; Pulvermüller, 2005; Willems & Hagoort, 2007). For example, Broca's area is also active during the execution of hand and arm goal-directed movements (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Furthermore, the mirror neuron system in humans appears to include Broca's area, which was interpreted by Rizzolati et al. as a common mechanism for language understanding and action recognition that relies on mirror neurons (Rizzolatti & Arbib, 1998). Although these observations are arguments in favor of the general concept of common mechanisms for language and action, the mirror neuron approach focuses on the natural selection aspect of language, and does not explain how language is learned during the infant's development. In this work, we focus on the developmental aspect of language learning, as opposed to the preformist theory which states that language does not necessitate learning but rather relies on innate structures such as the Universal Grammar proposed by Chomsky (2006).

For this purpose, we propose the following functionalist cognitive model of language development:

- **Broca's region** is a strong associative region, with bidirectional inputs from a) reward/goal areas and b) premotor and motor cortex. In the first stages of language development, we propose that this region is key to the infant learning to associate symbolic actions (imperative speech or gestures)