Adv Dev Educ Psychol, 2024, 5(1): 175-184
DOI: 10.25082/ADEP.2024.01.001

REVIEW

Empirical research in early infancy language acquisition: A nonsystematic review of literature

Jing An

The College of Sinology and Chinese Studies, Beijing Language and Culture University, Beijing 100083, China

Correspondence to: Jing An, The College of Sinology and Chinese Studies, Beijing Language and Culture University, Beijing 100083, China;
E-mail: jingan19930804@gmail.com

Received: January 8, 2024;
Accepted: March 8, 2024;
Published: March 13, 2024.

Citation: An, J. (2024). Empirical research in early infancy language acquisition: A nonsystematic review of literature. *Advances in Developmental and Educational Psychology*, **5**(1), 175-184.
https://doi.org/10.25082/ADEP.2024.01.001

Abstract: This paper unsystematically reviewed the journal publications written in English in the period from 2010 to 2020 on the topic of early infancy language acquisition. The review was through two aspects: language comprehension, and language production. Additionally, this paper also reviewed several frequently used and new research methods and tools. For early infancy language comprehension, empirical studies found evidences on infants' comprehension of words' meaning through sounds, especially the comprehension of nouns, the latter was proven to have relevance with nouns' familiarity and cross-word relations in the nouns. Besides, empirical studies found evidences on infants' distinguishment between concrete words and abstract concepts. An important factor to influence infancy language comprehension is the social environment that the baby was exposed to, mainly the mother's speech and her daily life routine (for example, her work). For early infancy language production, empirical studies found evidences on infants' ability to associate similar sounds with different objects, and measurement for infants' language outputs were brought forward in another empirical study. Empirical studies also found infants' language outputs match to words and objects from the environment. Lastly, this paper reviewed the most frequently used technical methods: fMRI and fNIRS technology for investigating neural mechanisms of infancy language processing. There are other new research methods, include large-sample database analysis and quantitative modeling, corpus analysis, language inputs sampling, language model for infancy language acquisition.

Keywords: infancy, language acquisition, language model

1 Introduction

Language comprehension and language production consisted the holistic process of language development. The mechanism of infancy language development, or, infancy language acquisition, helps scientists to understand the unsolved questions of language. Therefore, this paper will review the early infancy language acquisition through these two aspects, in a nonsystematic approach. Besides, there are some developing researching tools and new methods in the field of infancy language development, therefore related literature are also included. In the field of infants' early language acquisition, scholars have conducted many empirical researches. Some of these researches are from the perspective of early language comprehension, and could be analyzed through investigation on language inputs (Bergelson & Swingley, 2012; Bergelson & Swingley, 2015; Bergelson & Aslin, 2017a; Garrison et al., 2020; Bergelson & Aslin, 2017b; Bergelson & Swingley, 2013; Laing & Bergelson, 2019; Bergelson & Swingley, 2018), some researches are from the perspective of early language production, and could be explored through infancy language outputs, in other words, babies' babbling (Laing & Bergelson, 2020; Swingley & Humphrey, 2018; Yoshida et al., 2009). There are also some studies conducted from the perspective of developing new researching tools and methods in relation to early infancy language acquisition and development (Bergelson et al., 2018; Adriaans & Swingley, 2017; Buerkin-Pontrelli et al., 2018; Aslin et al., 2015).

2 Methods

A comprehensive literature search was conducted, using Google Scholar, which was limited to journal publications written in English between 2010 and 2020.

The following search terms were used: "infancy, infant"; "language acquisition, infancy language acquisition"; "empirical research"; "infancy language development"; "infancy language comprehension, early language comprehension"; "infancy language production, early language production"; and "early language acquisition researching method". These search strategies yielded more than 100 hits. An initial screening process resulted in the exclusion of duplicate articles, conference abstracts, and expert opinions. A final selection of 16 articles was made, including empirical research journal articles, database and corpus-based studies, language model-based study and systematic review.

3 Early infancy language comprehension

A coin has two sides, one side of early infancy language acquisition is infancy language comprehension. Researches on this aspect were conducted through language inputs of the infants. Scholars have conducted in-depth research on infants' word understanding—-mainly nouns understanding in relation to words' familiarity and cross-words relation; how they connect word with sounds; how they distinguish between concrete words and abstract concepts; how the social environment especially mother's speech and her daily life routine influence the infants' language acquisition and development.

3.1 Comprehension of words' meaning through sounds

Bergelson and Swingley (2012) demonstrated that instead of between 9 and 15 months of age, in fact, infants already know the meanings of several common words from the age of 6 months onward. It is previously widely accepted that infants begin learning their native language not by learning words, but by discovering features of the speech signal: consonants, vowels, and combinations of these sounds. Learning to understand words, as opposed to just perceiving their sounds, was said to come later, between 9 and 15 months of age, when infants develop a capacity for interpreting others' goals and intentions. Bergelson and Swingley tried to demonstrate that this consensus about the developmental sequence of human language learning is flawed: they presented 6- to 9-month-old infants with sets of pictures to view while their parent named a picture in each set. Over this entire age range, infants directed their gaze to the named pictures, indicating their understanding of spoken words. Because the words were not trained in the laboratory, the results show that even young infants learn ordinary words through daily experience with language. This surprising accomplishment indicates that, contrary to prevailing beliefs, either infants can already grasp the referential intentions of adults at 6 months or infants can learn words before this ability emerges. The precocious discovery of word meanings suggests a perspective in which learning vocabulary and learning the sound structure of spoken language go hand in hand as language acquisition begins. Though evidential, this conclusion draw from the experiments was not rigorous enough. Considering that the babies had already had 6-9 months of life experience with their parents (or caregivers), a possible explanation could also be that the babies were receiving a certain influence from the behavior of their parents or caregivers while showing instinctive sensitivity in order to survive, therefore, infants focused their gaze on the picture named by the parent does not necessarily indicate their understanding of spoken language or that they have learned the word. In addition, children do not develop theory of mind until they are about 4 years old, therefore, in terms of grasping the referential intentions of adults, another possible explanation is that babies instinctively sensitively respond to the caregiver's behavior, but they cannot be said to be able to grasp the caregiver's intentions.

3.2 Comprehension distinguishing concrete words and abstract concepts

Bergelson and Swingley (2013) continued to develop their research on this direction, they came up with that young infants' learning of words for abstract concepts like 'all gone' and 'eat' in contrast to their learning of more concrete words like 'apple' and 'shoe' may follow a relatively protracted developmental course. They conducted a new experiment to examine this hypothesis. Parents named one of two events shown in side-by-side videos while their 6-16-month-old infants (n = 98) watched. On average, infants successfully looked at the named video by 10 months, but not earlier, and infants' looking at the named referent increased robustly at around 14 months. They thought that 6-month-olds already understand concrete words in this task (Bergelson & Swingley, 2012). A video-corpus analysis of unscripted mother-infant interaction showed that mothers used the tested abstract words less often in the presence of their referent events than they used concrete words in the presence of their referent objects. They suggest that referential uncertainty in abstract words' teaching conditions may explain the later acquisition of abstract than concrete words, and they thought the changes in social-cognitive abilities plays a role over the 6-14 months period. They discussed that one possibility was that the developmental difference concerns the requirements for learning more abstract words. This hypothesis had two versions: One was that the same learning machinery is at work in learning concrete and abstract words, but the statistics of abstract words are more complex and therefore demand more data to resolve, which is manifested there in the later age at which evidence of learning is found; The other was that learning abstract words demands skills that do not begin to emerge until around 10 months, such as the capacity for reading others' intentions. This ability could prove more useful for understanding abstract words than concrete words, because abstract words may have fewer correlated perceptual features. That is, shape, size, color, movement, and texture range less freely for things called 'juice' than for situations called 'all gone!' or 'uh-oh.' While our task does not itself require gaze-following, abstract words are, to a greater degree, expressions of the parent's perspective, and as such their learning might depend more on skills of intention-reading. At 2015 they replicated and extended their findings (Bergelson & Swingley, 2015) from experimental reports which have shown that infants of 6-9 months know the meanings of some common words. With a new set of items, they showed that when young infants (age 6-16 months, n = 49) are presented with side-by-side video clips depicting various common early words, and one clip is named in a sentence, they look at the named video at above-chance rates. They reported anew that infants understand common words by 6-9 months and that performance increases substantially around 14 months. The results implied that 6- to 9-month-olds' failure to understand words not referring to objects (verbs, adjectives, performatives) in a similar prior study was not attributable to the use of dynamic video depictions. Thus, 6- to 9-month-olds' experience of spoken language includes some understanding of common words for concrete objects, but relatively impoverished comprehension of other words.

3.3 Comprehension of nouns

3.3.1 Familiarity of nouns

Bergelson and Aslin investigated one step further at 2017 (Bergelson & Aslin, 2017a). Their study investigated infants' knowledge about familiar nouns. Infants (n = 46, 12-20-month-olds) saw two-image displays of familiar objects, or one familiar and one novel object. Infants heard either a matching word (*e.g.* "foot" when seeing foot and juice), a related word (*e.g.* "sock" when seeing foot and juice) or a nonce word (*e.g.* "fep" when seeing a novel object and dog). Across the whole sample, infants reliably fixated the referent on matching and nonce trials. On the critical related trials they found increasingly less looking to the incorrect (but related) image with age. These results suggested that one-year-olds look at familiar objects both when they hear them labeled and when they hear related labels, to similar degrees, but over the second year increasingly rely on semantic fit. The researchers suggested that infants' initial semantic representations are imprecise, and continue to sharpen over the second postnatal year.

Garrison et al. (2020) also studied nouns comprehension in infants, they found that familiarity plays a small role in nouns comprehension in infants aged 12-18 months. They pointed out infants amass thousands of hours of experience with particular items, each of which is representative of a broader category that often shares perceptual features. Robust word comprehension requires generalizing known labels to new category members. While young infants have been found to look at common nouns when they are named aloud, the role of item familiarity had not been well-examined. This study compared 12-18-month-olds' word comprehension in the context of pairs of their own items (*e.g.* photos of their own shoe and ball) versus new tokens from the same category (*e.g.* a new shoe and ball). Their results replicated previous work showing that noun comprehension improves rapidly over the second year, while also suggested that item familiarity appears to play a far smaller role in comprehension in this age-range. This in turn suggested that even before age two, ready generalization beyond particular experiences is an intrinsic component of lexical development.

3.3.2 Cross-word relations

In 2017, Bergelson and Aslin (2017b) also came up with the question that although the surprising finding that even 6-month-olds understand common nouns (Bergelson & Swingley, 2012), however, is their early lexicon structured and acquired like older learners? They tested 6-month-olds for a hallmark of the mature lexicon: cross-word relations. They also examined whether properties of the home environment that have been linked with lexical knowledge in older children were detectable in the initial stage of comprehension. They used a new dataset, which includes in-lab comprehension and home measures from the same infants. They found evidence for cross-word structure: on seeing two images of common nouns, infants looked significantly more at named target images when the competitor images were semantically unrelated (*e.g.*, milk and foot) than when they were related (*e.g.*, milk and juice), just as older learners did. They further found initial evidence for home-lab links: common noun "copresence" (i.e., whether words' referents were present and attended to in home recordings) correlated with in-lab comprehension. These findings suggested that, even in neophyte word learners, cross-word relations are formed early and the home learning environment measurably helps shape the lexicon from the outset.

3.4 Social environment of comprehension

In 2018, Bergelson and Swingley (2018) had studied the impact of social environment on infants' comprehension. They studied from a phonetic perspective on infants' understanding of words when the speaker was unfamiliar and when the pronunciation varied. They pointed out that to understand spoken words, listeners must appropriately interpret co-occurring talker characteristics and speech-sound content. This ability was tested in 6-14-months-olds by measuring their looking to named food and body-part images. In the new talker condition (n = 90), pictures were named by an unfamiliar voice; in the mispronunciation condition (n = 98), infants' mothers "mispronounced" the words (*e.g.*, nazz for nose). 6-7-month-olds fixated target images above chance across conditions, understanding novel talkers, and mothers' phonologically deviant speech, equally. 11-14-months-olds also understood new talkers, but performed poorly with mispronounced speech, indicating sensitivity to phonological deviation. Between these ages performance was mixed. These findings highlighted the changing roles of acoustic and phonetic variability in early word comprehension, as infants learn which variations alter meaning. Taken together, their results suggested that speech-sound reorganization occurs in tandem with lexical growth over the first postnatal year. Although infants might reveal a performance cost with more thorough individual testing, here infants were unbeholden to highly familiar tokens during early word comprehension. Infants' lexical representations allow them to connect novel word tokens to the conceptual categories they denote, at the same (modest) levels demonstrated for correctly-pronounced maternal speech. Between 6 and 14 months, they reorganize their speech-sound interpretation to continue accepting words spoken with new voices but to exclude realizations that deviate from the conventional form in their phonology. This research may have a considerable connection with the rapid and effective acquiring of the oral second language of infants, because if infants are not restricted by familiar symbols in the early word comprehension process - while adults are just the opposite, who are rather constrained by familiar symbols - then infants can easily associate new word symbols with the concepts these symbols refer to as correctly as understanding their mother's correctly pronounced speech, then it would not be so surprising that infants can quickly learn a new language that their mother does not use, especially in listening and speaking.

Laing and Bergelson (2019) examined the literature on the effects of mothers' work status on infant language development and it was mixed, with little focus on varying work schedules and early vocabulary. They used naturalistic data to analyze the productive vocabulary of 44 17-month-olds in relation to mothers' work status (full time, part time, stay at home) at 6 and 18 months. Infants who experienced a combination of care from mothers and other caretakers had larger productive vocabularies than infants in solely full-time maternal or solely other-caretaker care. Their results draw from naturalistic data to suggest that this care combination may be particularly beneficial for early lexical development.

4 Early infancy language production

Another side of early infancy language acquisition is infancy language production, or, language outputs, which means babies' babbling. Many in-depth empirical researches were conducted in this aspect.

4.1 Ability to associate sounds with different objects

Yoshida et al. (2009) came up with a question that, can infants, in the very first stages of word learning, use their perceptual sensitivity to the phonetics of speech while learning words? Before them, a common idea was infants of 14 months cannot learn two similar sounding words unless there is substantial contextual support. Their experiment advanced the understanding of this failure by testing whether the source of infants' difficulty lies in the learning or testing phase. Infants were taught to associate two similar sounding words with two different objects, and tested using a visual choice method rather than the standard Switch task. The results revealed that 14-month-olds are capable of learning and mapping two similar sounding labels; they can apply phonetic detail in new words. Yoshida et al. discussed the relation between infants' concurrent failure and the developmental transition to success in the Switch task, and suggested that infants' learning abilities were previously obscured by the demands of the testing phase in the experimental task.

4.2 Production measurement

Swingley and Humphrey (2018) evaluated which features of spoken language aid infant word learning, used a corpus of infant directed speech (Brent & Siskind, 2001), which was characterized on several linguistic dimensions and statistically related to the infants' vocabulary outcomes word-by-word. Comprehension (at 12, 15 months) and production (15 months) were predicted by frequency, frequency of occurrence in one-word utterances, concreteness, utterance length, and typical duration. These features had been proposed to influence learning before, but in this research their relative contributions were measured. Mothers' data predicted learning in their own children better than in other children; thus, vocabulary was measurably aligned within families. These analyses provided a quantitative basis for claims concerning the relevance of several properties of maternal English speech in facilitating early word learning.

4.3 Production matches to objects from the environment

Laing and Bergelson (2020) found early infant production of babbling is matched to words and objects from the infant's environment. Infants' early babbling allows them to engage in proto-conversations with caretakers, well before clearly articulated, meaningful words are part of their productive lexicon. Moreover, the well-rehearsed sounds from babble serve as a perceptual 'filter', drawing infants' attention towards words that match the sounds they can reliably produce. Using naturalistic home recordings of 44 10-11-month-olds (an age with high variability in early speech sound production), this study tested whether infants' early consonant productions match words and objects in their environment. They found that infants' babble matches the consonants produced in their caregivers' speech. Infants with a well-established consonant repertoire also match their babble to objects in their environment. Their findings showed that infants' early consonant productions are shaped by their input: by 10 months, the sounds of babble match what infants see and hear. Their results provided first-step evidence for how the transition to word production might take place, as infants bridge the gap between babble and words through the articulatory matching of input and output. That is, Vocal Motor Schema (VMS) may draw infants' attention to words they could produce in the input, making them more memorable and therefore good candidates for the early lexicon. Their results suggested that Vocal Motor Schema (VMS) consonants may prompt articulatory 'matching' between what an infant hears in the input and their own vocalizations, independent of any lexical knowledge of the word in question. They proposed that this may support more speech-like consonant production, as infants move from the canonical syllables of babble to the establishment of an inventory of speech sounds. The causal relationship here was questionable to some extent: did the baby babble first and then pay attention to the matching word, or did the baby see the thing and hear its word (thus understanding the simple correspondence between sound and thing) and then try to babble? In other words, did the Vocal Motor Schema driven the baby to babble and learn or it was because of something else? Is the reason of early lexicon vocal? Or it was through the combination of visual, vocal and something else? How do deaf infants develop early lexicon? These questions remain worth studying. But it is evidential that vocal experience plays an important role in early lexicon acquisition.

5 Methods and new tools in the field

Good research practice needs developed methods and up-to-date tools. Researches on developing new researching methods and tools in relation to the field of early infancy language acquisition were also reviewed, in an unsystematic approach.

5.1 The most frequently used technical methods

Over the past two decades, the field of cognitive neuroscience has relied heavily on hemodynamic measures of blood oxygenation in local regions of the brain to make inferences about underlying cognitive processes. These same functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS) techniques have been adapted for use with human infants.

Aslin, Shukla and Emberson (2015) reviewed the advantages and disadvantages of these two neuroimaging methods for studies of infant cognition, with a particular emphasis on their technical limitations and the linking hypotheses that are used to draw conclusions from correlational data. fMRI provides a rich picture of the spatial distribution of estimated neural activity, but it has its disadvantages: First, MRI recordings require rigid head stabilization; currently available algorithms can tolerate head movements of only a few millimeters over a 5-10 min scan, resulting in elimination of scans that exceed this threshold. Moreover, the head coil and associated mirror for viewing visual stimuli constrain the angular size and proximity of visual displays, making it difficult to reliably present visual stimuli to infants. Second, the MRI scanner produces substantial acoustic noise from the gradients that, even with adequate hearing protection, can be distracting or induce startle responses that limit the ability to hear subtle acoustic differences, or attend to nonauditory stimuli. Third, safety issues inherent to any MR scan (*e.g.*, heating of body tissues, especially if they come in contact with any ferrous metal, and damage to the auditory system from inadequate hearing protection), while posing limited dangers, are more difficult to mitigate in a nonverbal participant such as an infant. Despite taking special precautions to deal with these safety concerns, it has proven to be very difficult to eliminate movement artifacts in infants and very young children. Thus, the vast majority of fMRI studies with infants are conducted while they are sleeping. This enables the successful collection of structural MRI and fMRI recordings in this population without movement artifacts, but it severely limits the kinds of stimuli and behavioral paradigms that can be used with infants to study the neural correlates of cognitive development.

The disadvantages of fMRI are counterpointing to the major advantages of fNIRS: First, the fNIRS cap that holds the fiber-optic bundles to deliver near-infrared light to the emitters and gather near-infrared light from the detectors can be lightweight and comfortable to wear. Second, because this cap attaches the emitters and detectors to the scalp, the fNIRS recordings do not require rigid head stabilization and therefore enable the infant to be positioned to maximize attention to experimental stimuli or to engage in motor responses that would not be feasible in the MRI scanner environment. Third, there is no acoustic noise from the fNIRS machine (in contrast to noisy MRI gradients), thereby reducing distractions and maximizing the fidelity of auditory stimuli. Fourth, there are no safety concerns because the intensity of the IR-light sources is well below the FDA guidelines for damage to neural tissue. Thus, fNIRS studies can readily be conducted with developmental and special populations who would not tolerate the constraints of the fMRI environment.

The disadvantages of fNIRS had been overcame at a certain level comparing to the past, but some disadvantages remain: First, the spatial resolution of fNIRS is currently 2-3 cm, and as noted above, the signal recorded from a given channel samples a 3D volume that is not precisely known and projects that volume onto a midchannel location on the surface of the scalp. Second, fNIRS does not provide coregistration with anatomical images of the participant's brain. Rather, either a separate structural MRI must be obtained and the location of the emitters and detectors aligned with external skull landmarks, or average structural MRI templates for age- or headsize-appropriate infants can be used to provide approximate coregistration using average skull landmarks. Third, fNIRS can only record from the surface of the cortex (although sulci are clearly within the range of accessible depths). Thus, brain regions more than 2 cm from the surface of the scalp are outside the range of fNIRS recordings, including many cortical areas in the ventral temporal cortex (*e.g.*, the fusiform gyrus) and all medial-brain structures (*e.g.*, hippocampus). Fourth, fNIRS recordings obtained from participants across different ages must consider changes in skull thickness and signal degradation introduced by the presence of hair (particularly if the hair is dense and highly pigmented). Finally, because fNIRS signals sample photons that passed twice through the scalp, skull, and surface vasculature as they travel from emitter to detector, changes in these signals are contaminated by these non-neural factors. Thus, without taking precautions to reduce the effects of systemic vascular responses (*e.g.*, due to changes in blood pressure mediated by arousal), measurements of changes in cortical oxygenation may be masked by these much larger non-neural, supracortical signals.

Studies used both fMRI and fNIRS to examine infants' neural activity patterns in response to speech stimuli. Aslin, Shukla and Emberson (2015) highlighted the prospects of improving the quality of fNIRS data from infants to address in a more sophisticated way how cognitive development is mediated by changes in underlying neural mechanisms. They also summarized key findings in the late 1990s to the early 2010s in several domains of infant cognition including language processing and learning: Early studies examining the pattern of neural activity in infants to speech stimuli using both fNIRS and fMRI were directed to the question of cortical selectivity (i.e., lateralization) and to functional specialization [i.e., speech versus nonspeech (Dehaene-Lambertz et al. 2002; Peña et al. 2003)]. Subsequent studies have addressed three questions: (a) What are the acoustic properties of speech that engender language-like processing (Peña et al., 2003; Dehaene-Lambertz et al., 2002; Poeppel, 2003; Redcay et al., 2008; Bortfeld et al., 2009; Telkemeyer et al., 2009; Telkemeyer et al., 2011; Arimitsu et al., 2011; Mahmoudzadeh et al., 2013), (b) How does speech processing interact with other cognitive capacities such as memory and attention, which involves a form of automatic (implicit) learning (Ross-Sheehy et al., 2004; Nakano et al., 2008; Nakano et al., 2009; Blaser & Kaldy, 2010; Benavides-Varela et al., 2011; Benavides-Varela et al., 2012), and (c) What are the early learning mechanisms that might underlie the extraction of linguistic structure present in the speech input? For language acquisition is not just about rote memorization of speech input, but crucially depends on the ability to understand and to generate novel utterances. Studies with infants have demonstrated that humans are capable of extracting relevant language like structures from brief exposure to speech (Marcus et al., 1999; Gómez, 2002; Gervain et al., 2008; Wagner et al., 2011; Gervain et al., 2012; Marchetto & Bonatti, 2013; Gerken et al., 2014).

5.2 Large-sample database analysis and quantitative modeling method

Adriaans and Swingley (2017) used large-sample database analysis and quantitative modeling methods to study the impact of prosodic exaggeration in infant speech on the learnability of vowels. Acoustic analyses of natural infant-directed speech have suggested that phonetic categories are not presented to learners as separable clusters of sounds in acoustic space. Adriaans & Swingley's study was a step toward explaining how infants begin to solve this problem, they proposed that the exaggerated prosody characteristic of infant-directed speech may highlight for infants certain speech-sound tokens that collectively form more readily identifiable categories. A database was presented, containing vowel measurements in a large sample of natural American English infant-directed speech. Analyses of the vowel space showed that prosodic exaggeration in infant-directed speech has the potential to support distributional vowel learning by providing the learner with a subset of "high-quality" tokens that infants might attend to preferentially. Categorization models trained on prosodically exaggerated tokens outperformed models that were trained on tokens that were not exaggerated. Though focused on more prominent, exaggerated tokens does not provide a solution to the categorization problem, it would make it easier to solve. They stated that the large experimental literature on infants' early language trajectories was not matched by a comprehensive description of children's language environment, so it was difficult to quantitatively track children's typical developmental paths, and quantitative modeling of early speech and vocabulary development remains to be studied. Additionally, as of the time their study was published, no research had addressed on the topic of the speaker's underlying motivations or intentions.

5.3 Corpus analysis method

In addition, scholars in the field of early language acquisition also use corpus analysis methods to study infants' vocabulary and phonetic learning. These corpus analyses can be cross-language. Swingley and Alarcon (2018) analyzed a corpus of native Spanish speakers and concluded that vocabulary learning may contribute to infants' phonological learning. The advantage of using a Spanish corpus is that in Spanish samples, phonetic context does not obscure a given vowel's occurrence in different words. In their first year, infants begin to learn the speech sounds of their language. This process is typically modeled as an unsupervised clustering problem in which phonetically similar speech sound tokens are grouped into phonetic categories by infants using their domain-general inference abilities. Swingley and Alarcon (2018) argued that maternal speech is too phonetically variable for this account to be plausible, and they provided phonetic evidence from Spanish showing that infant-directed Spanish vowels are more readily clustered over word types than over vowel tokens. The results suggested that infants' early adaptation to native-language phonetics depends on their word-form lexicon, implicating a much wider range of potential sources of influence on infants' developmental trajectories in language learning. However, this study also had its shortcomings: The pronunciation characteristic of Spanish is that other phonemes play a role in the presentation of vowels, which is not common in most languages; Moreover, the corpus consist of approximately 2,000 vowels used in this study is too small comparing to the life experience of a 6- to 8-month-old infant.

5.4 New language inputs sampling method

Bergelson et al. (2018) provided information on sampling methods for natural language input from infants. They pointed out, measurements of infants' quotidian experiences provide critical information about early development. However, the role of sampling methods in providing these measurements was rarely examined. They directly compared language input from hour-long video-recordings and daylong audio-recordings within the same group of 44 infants at 6 and 7 months. They compared 12 measures of language quantity and lexical diversity, talker variability, utterance-type, and object presence, finding moderate correlations across recording-types. However, video-recordings generally featured far denser noun input across these measures compared to the daylong audio-recordings, more akin to 'peak' audio hours (though not as high in talkers and word types). Although audio-recordings captured about 10 times more awake-time than videos, the noun input in them was only 2-4 times greater. Notably, whether they compared videos to daylong audio-recordings or peak audio times, videos featured relatively fewer declaratives and more questions; furthermore, the most common video recorded nouns were less consistent across families than the top audio-recording nouns were. Thus, hour-long videos and daylong audio-recordings revealed fairly divergent pictures of the language infants hear and learn from in their daily lives. They suggested that short video-recordings provide a dense and somewhat different sample of infants' language experiences, rather than a typical one, and should be used cautiously for extrapolation about common words, talkers, utterance-types, and contexts at larger timescales. If theories of language development are to be held accountable to 'facts on the ground' from observational data, greater care is needed to unpack the ramifications of sampling methods of early language input.

5.5 Language model in relation to infancy language acquisition

From the perspective of new tools, Buerkin-Pontrelli, Coffey and Swingley (2018) also explored the question of whether current statistics-based concepts of language learning can be extended to infants' natural language experience. They implemented a statistical-clustering word-segmentation model and sent its outputs to an implementation of a "frame" based form class tagger and, separately, to a simple word-order heuristic parser. They tested this pipeline model on various input types, ranging from quite idealized (orthographic words) to more naturalistic resyllabified corpora. They asked how these modeled capacities work together when they receive the noisy outputs of upstream word finding processes as input, which more closely resembles the scenario infants face in language acquisition. Their assessment was that the present analyses, on balance, probably underestimate the difficulties infants face, because true phonetic variability is more severe than acknowledged by our input corpora. Even so, the results suggested that while some word-finding is feasible statistically, successful form-class identification cannot be accomplished in this pipeline using frequent frames, excepting a gross and still error-ridden division into unlabeled clusters of nouns, verbs, and miscellany. Frequency imbalances and their ordering might suggest aspects of word order to infants, but the imbalances were slight in their resyllabified data, so that one might question whether they were salient enough to drive infant intuitions. In their view, it is likely that the statistical outputs they modeled were somewhat informative, but clearly insufficient. The solution, they suggested, was probably not to reduce estimates of what infants know, but rather to find ways to incorporate knowledge of word meaning into the relevant computations. Although this study wasn't perfect, the idea of extending the current statistical-based language learning concept to human natural language experience is worthy of reference. In the field of artificial intelligence, especially natural language processing, this way of thinking may be of great use in the future. For human individuals who are unable to express natural speech autonomously, such as human infants or persons with aphasia, communication could be simplified by the usage of language models and/or brain-computer interfaces technology.

6 Conclusion

This paper unsystematically reviewed the journal publications written in English in the period from 2010 to 2020 on the topic of early infancy language acquisition. The review was through two aspects: language comprehension, and language production. Additionally, this paper also reviewed several frequently used and new research methods and tools.

For early infancy language comprehension, empirical studies found evidences on infants' comprehension of words' meaning through sounds, especially the comprehension of nouns, the latter was proven to have relevance with nouns' familiarity and cross-word relations in the nouns. Besides, empirical studies found evidences on infants' distinguishment between concrete words and abstract concepts. An important factor to influence infancy language comprehension is the social environment that the baby was exposed to, mainly the mother's speech and her daily life routine (for example, her work). For early infancy language production, empirical studies found evidences on infants' ability to associate similar sounds with different objects, and measurement for infants' language outputs were brought forward in another empirical study. Empirical studies also found infants' language outputs match to words and objects from the environment. Lastly, this paper reviewed the most frequently used technical methods: fMRI and fNIRS technology for investigating neural mechanisms of infancy language processing. There are other new research methods, include large-sample database analysis and quantitative modeling, corpus analysis, language inputs sampling, language model for infancy language acquisition.

Conflict of interest

The author declares no conflicts of interest in this work.

References

Adriaans, F., & Swingley, D. (2017). Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. The Journal of the Acoustical Society of America, 141(5), 3070–3078. https://doi.org/10.1121/1.4982246

Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., Ikeda, K., Takahashi, T., & Minagawa-Kawai, Y. (2011). Functional Hemispheric Specialization in Processing Phonemic and Prosodic Auditory Changes in Neonates. Frontiers in Psychology, 2. https://doi.org/10.3389/fpsyg.2011.00202

Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic Correlates of Cognition in Human Infants. Annual Review of Psychology, 66(1), 349–379. https://doi.org/10.1146/annurev-psych-010213-115108

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2018). Day by day, hour by hour: Naturalistic language input to infants. Developmental Science, 22(1). Portico. https://doi.org/10.1111/desc.12715

Bergelson, E., & Aslin, R. (2017). Semantic Specificity in One-Year-Olds' Word Comprehension. Language Learning and Development, 13(4), 481–501. https://doi.org/10.1080/15475441.2017.1324308

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. Proceedings of the National Academy of Sciences, 114(49), 12916–12921. https://doi.org/10.1073/pnas.1712966114

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. Proceedings of the National Academy of Sciences, 109(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. Cognition, 127(3), 391–397. https://doi.org/10.1016/j.cognition.2013.02.011

Bergelson, E., & Swingley, D. (2014). Early Word Comprehension in Infants: Replication and Extension. Language Learning and Development, 11(4), 369–380. https://doi.org/10.1080/15475441.2014.979387

Bergelson, E., & Swingley, D. (2017). Young Infants' Word Comprehension Given An Unfamiliar Talker or Altered Pronunciations. Child Development, 89(5), 1567–1576. Portico. https://doi.org/10.1111/cdev.12888

Benavides-Varela, S., Gómez, D. M., & Mehler, J. (2011). Studying Neonates' Language and Memory Capacities with Functional Near-Infrared Spectroscopy. Frontiers in Psychology, 2. https://doi.org/10.3389/fpsyg.2011.00064

Benavides-Varela, S., Hochmann, J.-R., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. Proceedings of the National Academy of Sciences, 109(44), 17908–17913. https://doi.org/10.1073/pnas.1205413109

Bortfeld, H., Fava, E., & Boas, D. A. (2009). Identifying Cortical Lateralization of Speech Processing in Infants Using Near-Infrared Spectroscopy. Developmental Neuropsychology, 34(1), 52–65. https://doi.org/10.1080/87565640802564481

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. Cognition, 81(2), B33–B44. https://doi.org/10.1016/s0010-0277(01)00122-6

Blaser, E., & Kaldy, Z. (2010). Infants Get Five Stars on Iconic Memory Tests. Psychological Science, 21(11), 1643–1645. https://doi.org/10.1177/0956797610385358

Buerkin-Pontrelli, A., Coffey, J., & Swingley, D. (2018). Outputs as inputs: Sequential Models of the Products of Infant 'Statistical Learning' of Language. Paper presented at the CogSci.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional Neuroimaging of Speech Perception in Infants. Science, 298(5600), 2013–2015. https://doi.org/10.1126/science.1077066

Garrison, H., Baudet, G., Breitfeld, E., Aberman, A., & Bergelson, E. (2020). Familiarity plays a small role in noun comprehension at 12–18 months. Infancy, 25(4), 458–477. Portico. https://doi.org/10.1111/infa.12333

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. Proceedings of the National Academy of Sciences, 105(37), 14222–14227. https://doi.org/10.1073/pnas.0806530105

Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at Birth: The Newborn Brain Detects Identity Relations and Sequential Position in Speech. Journal of Cognitive Neuroscience, 24(3), 564–574. https://doi.org/10.1162/jocn_a_00157

Gerken, L. A., Dawson, C., Chatila, R., & Tenenbaum, J. (2015). Surprise! infants consider possible bases of generalization for a single input example. Developmental Science, 18(1), 80. https://doi.org/10.1111/desc.12183

Gómez, R. L. (2002). Variability and Detection of Invariant Structure. Psychological Science, 13(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Laing, C., & Bergelson, E. (2020). From babble to words: Infants' early productions match words and objects in their environment. Cognitive Psychology, 122, 101308. https://doi.org/10.1016/j.cogpsych.2020.101308

Laing, C., & Bergelson, E. (2018). Mothers' Work Status and 17-Month-Olds' Productive Vocabulary. Infancy, 24(1), 101–109. Portico. https://doi.org/10.1111/infa.12265

Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., Grebe, R., & Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. Proceedings of the National Academy of Sciences, 110(12), 4846–4851. https://doi.org/10.1073/pnas.1212220110

Marchetto, E., & Bonatti, L. L. (2013). Words and possible words in early language acquisition. Cognitive Psychology, 67(3), 130–150. https://doi.org/10.1016/j.cogpsych.2013.08.001

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. Science, 283(5398), 77–80. https://doi.org/10.1126/science.283.5398.77

Nakano, T., Homae, F., Watanabe, H., & Taga, G. (2008). Anticipatory Cortical Activation Precedes Auditory Events in Sleeping Infants. PLoS ONE, 3(12), e3912. https://doi.org/10.1371/journal.pone.0003912

Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2008). Prefrontal Cortical Involvement in Young Infants' Analysis of Novelty. Cerebral Cortex, 19(2), 455–463. https://doi.org/10.1093/cercor/bhn096

Peña, M., Maki, A., Kovacić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. Proceedings of the National Academy of Sciences, 100(20), 11702–11705. https://doi.org/10.1073/pnas.1934290100

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time.' Speech Communication, 41(1), 245–255. https://doi.org/10.1016/s0167-6393(02)00107-3

Redcay, E., Haist, F., & Courchesne, E. (2008). Functional neuroimaging of speech perception during a pivotal period in language acquisition. Developmental Science, 11(2), 237–252. Portico. https://doi.org/10.1111/j.1467-7687.2008.00674.x

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The Development of Visual Short-Term Memory Capacity in Infants. Child Development, 74(6), 1807–1822. Portico. https://doi.org/10.1046/j.1467-8624.2003.00639.x

Swingley, D., & Alarcon, C. (2018). Lexical Learning May Contribute to Phonetic Learning in Infants: A Corpus Analysis of Maternal Spanish. Cognitive Science, 42(5), 1618–1641. Portico. https://doi.org/10.1111/cogs.12620

Swingley, D., & Humphrey, C. (2017). Quantitative Linguistic Predictors of Infants' Learning of Specific English Words. Child Development, 89(4), 1247–1267. Portico. https://doi.org/10.1111/cdev.12731

Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., & Wartenburger, I. (2009). Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds. The Journal of Neuroscience, 29(47), 14726–14733. https://doi.org/10.1523/jneurosci.1246-09.2009

Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., & Wartenburger, I. (2011). Acoustic Processing of Temporally Modulated Sounds in Infants: Evidence from a Combined Near-Infrared Spectroscopy and EEG Study. Frontiers in Psychology, 1. https://doi.org/10.3389/fpsyg.2011.00062

Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. Developmental Science, 12(3), 412–418. Portico. https://doi.org/10.1111/j.1467-7687.2008.00789.x