Infant Protophones: Prediction of Language Development and Mechanisms of Action

Authors: Liu Qiaoyun, Liu Min, Siqi Chen, Zhijia Xu, Liu Qiaoyun

Date: 2023-02-01T00:00:00+00:00

Abstract

Prelinguistic vocalizations are speech-like vocalizations that constitute the foundation for subsequent language development. Existing research has demonstrated that the quantity or frequency of prelinguistic vocalizations in infants and toddlers, the proportion or frequency of canonical syllables, consonant diversity, and communicative vocalizations predict expressive language; however, findings regarding the prediction of receptive language have been inconsistent, and results concerning the predictive value of canonical babbling onset age for expressive word onset age and expressive vocabulary size have also been inconsistent. The predictive relationship between prelinguistic vocalizations and language development can be partially explained by three mechanisms: prelinguistic vocalizations provide the foundational basis for language production, create learning opportunities for infants, and elicit social responsive behaviors. Future research may consider investigating the causal relationship between prelinguistic vocalizations and language development, the moderating effects of relevant factors on the predictive power of prelinguistic vocalizations for language development, the influence of language assessment methods and prelinguistic vocalization metrics themselves, the most valuable prelinguistic vocalizations for language intervention in children with language disorders, and the dynamic interplay between prelinguistic vocalizations and social responses.

Full Text

Preamble

Speech-like vocalizations, also known as protophones, refer to the various types of sounds that infants produce before they primarily use spoken language for communication. These sounds develop along a continuum and gradually become more speech-like (Oller, 2000; Schoen Simmons, 2021). In typically developing (TD) children, speech-like vocalizations progress through several stages: quasi-vowels (0–2 months), goos (1–4 months), full vowels (or fully resonant nuclei) and

marginal babbling (3–8 months), and finally canonical babbling (5–10 months) (Morgan & Wren, 2018; Oller, 2000). Quasi-vowels are produced when the vocal tract is in a relatively natural state and has not yet formed the vocal tract shape for vowels, resulting in lighter and shorter sounds. Full vowels are produced when the vocal tract shape for vowels has formed, resulting in louder and longer sounds. Goos are vocalizations produced when the tongue is positioned near the back of the oral cavity (Buder et al., 2013). Both canonical and marginal babbling consist of consonant-vowel (CV) syllable combinations. The differences between them are: first, the transition between consonant and vowel in canonical babbling is rapid (typically less than 120ms, though some studies consider it less than 250ms), while the transition time in marginal babbling is typically greater than 120ms (or greater than 250ms according to some studies). Second, marginal babbling may lack a full vowel as the syllable nucleus and may not have a substantial consonant-vowel transition, whereas canonical babbling must have a full vowel as the syllable nucleus and must have a substantial consonant-vowel transition (Buder et al., 2013; Goldstein & Schwade, 2008).

As a precursor to speech development, speech-like vocalizations constitute the majority of infant vocalizations (Oller et al., 2021), and their predictive value for language abilities has been extensively studied in both TD children and children with language disorders. However, these studies are primarily correlational, and causal relationship studies are still lacking, with inconsistent findings. For example, some studies have shown that the onset age of speech-like vocalizations predicts the onset age of expressive words (McGillion et al., 2017), while other studies have found no such predictive relationship (Lang et al., 2021). Additionally, some studies have found that the quantity of speech-like vocalizations predicts receptive language in infants and toddlers (Weismer et al., 2010), whereas others have not found this relationship (Werwach et al., 2021). Therefore, which speech-like vocalizations are correlated with language development in infants and toddlers, and what are the mechanisms underlying these correlations? This article first reviews empirical evidence for correlations between infant speech-like vocalizations and language development, then summarizes the underlying mechanisms that can explain these correlations, and finally proposes future research directions. By organizing and summarizing existing literature, this article aims to further understand how infant speech-like vocalizations help them transition from the prelinguistic period to the linguistic period, thereby providing a scientific basis for how to better promote infant language development.

2. Measurement of Speech-Like Vocalizations and Language Abilities

2.1 Data Collection

Current research primarily employs three methods to collect infant vocalization data: first, the Language ENvironment Analysis (LENA) system (e.g., Brook-

man et al., 2020); second, play interaction sample analysis (e.g., Lyakso et al., 2014); and third, parent report (Werwach et al., 2021). Regarding meta-analyses of full-day language samples using LENA, which comprehensively analyzed the relationship between speech-like vocalization quantity and expressive or receptive language in preterm infants (born one month early) to TD toddlers (up to 46 months), children with hearing loss, and children with Autism Spectrum Disorder (ASD), results indicated that speech-like vocalization quantity correlates with language abilities (including both receptive and expressive) (Wang et al., 2020). Two subsequent studies using LENA to measure vocalization quantity in TD infants supported these meta-analysis findings, showing that vocalization quantity at 12–14 months predicted expressive vocabulary at 18–20 months (Brookman et al., 2020; Ha et al., 2022). Play interaction sample analysis studies have shown that speech-like vocalization quantity in 9-month-old TD infants correlates with expressive vocabulary at 12 and 24 months (Lyakso et al., 2014), and that vocalization frequency in ASD toddlers (average age 23 and 33 months) predicted expressive language one year later (McDaniel et al., 2020a; Nevill et al., 2019) but did not predict receptive language (Nevill et al., 2019). Another study of 30-month-old children with ASD and developmental delay found that vocalization frequency in ASD toddlers correlated with concurrent receptive and expressive language, whereas children with developmental delay did not show the same pattern (Weismer et al., 2010). Parent report analysis results showed that speech-like vocalization quantity in 6-month-old TD infants significantly predicted expressive vocabulary at 12 months but did not significantly predict receptive vocabulary (Werwach et al., 2021).

Overall, results predicting expressive language from speech-like vocalization quantity are relatively consistent, whereas results predicting receptive language show discrepancies. These inconsistencies may be due to different methods of language ability testing and different characteristics of receptive and expressive language. For example, when summarizing previous differences in language measurement results, Nevill et al. (2019) found that direct testing of children yielded higher expressive than receptive language scores, whereas indirect testing (such as parent report) showed the opposite pattern of higher receptive than expressive language. Their study also found similar patterns among different directly administered language scales. Receptive language abilities are more implicit and difficult to fully capture through testing, whereas expressive language is more explicit and easier to reflect through testing. Future research could investigate the influence of language testing methods.

2.3 Phonetic Quality of Vocalizations

2.3.1 Canonical Syllables Canonical syllable frequency or ratio shows good predictive value for expressive language in infants and toddlers, whereas non-canonical syllable frequency does not predict expressive or receptive language. Studies of TD children have shown that canonical syllable ratio or true canonical syllable ratio in 9- and 12-month-old infants predicts expressive vocabulary

at 21 and 24 months (Chapman et al., 2003; Yankowitz et al., 2022). However, predictions of expressive vocabulary from canonical syllable ratio in 6-8-month-old infants are inconsistent. Yankowitz et al. (2022) found that canonical syllable ratio in 6-month-old infants did not predict expressive vocabulary at 24 months, whereas another study found that canonical syllable ratio in 6-8-month-old infants predicted expressive vocabulary at 18–20 months (Kim & Ha, 2022). The reason may be age differences among infants. Yankowitz et al.' s study included only 6-month-old infants, whose canonical syllable ratio was too low (around 2% on average). Kim and Ha' s study included 8-month-old infants, whose canonical syllable ratio was slightly higher (around 15% on average).

Studies on the relationship between non-canonical syllable quantity and language abilities have shown that non-canonical syllables in 13-month-old infants are not correlated with expressive or receptive vocabulary (Lopez et al., 2020). This study also found that canonical syllable quantity was not correlated with receptive vocabulary. Studies of children with language disorders have found similar results. For example, the number of canonical syllables produced per minute by 9-month-old infants with ASD correlated with expressive language at 12 months, whereas non-canonical syllable quantity did not correlate with expressive language (Talbott, 2014). Infants with fragile X syndrome who used canonical syllables at 9 months had better expressive and receptive language at 24 months than those who did not use canonical syllables, and higher early canonical syllable ratios predicted better later expressive language (Hamrick et al., 2019). Lack of canonical syllables in early infancy may indicate later language delay (Lohmander et al., 2017; Rubin, 2021). Overall, canonical syllable frequency or ratio predicts expressive language, with only a few studies examining the prediction of receptive language by canonical syllable frequency and showing inconsistent results. Additionally, age may be a factor influencing the predictive relationship between canonical syllables and expressive language.

2.3.2 Babbling Onset The predictive value of canonical babbling onset age for expressive word onset age and expressive vocabulary shows inconsistent results. First, some studies have used the stable production of at least one or two supraglottal consonants as the criterion for canonical babbling onset (CBO) and examined whether CBO predicts the onset age of first meaningful expressive words or receptive and expressive vocabulary (e.g., Majorano et al., 2014; McGillion et al., 2017). CBO correlates with meaningful expressive word onset age, such that smaller CBO predicts earlier meaningful expressive word onset (Keren-Portnoy et al., 2009; McGillion et al., 2017). However, CBO shows inconsistent predictions for receptive and expressive vocabulary. Majorano et al. (2014) found that CBO predicted expressive vocabulary at 12 months (i.e., smaller CBO predicted larger vocabulary at 12 months) but did not predict receptive vocabulary at 12 months or expressive vocabulary at 18 months. McGillion et al. (2017) found that CBO did not predict receptive vocabulary at 18 months but did predict expressive vocabulary.

Second, whether the production includes two reduplicated babbles or canonical babbling ratio or multisyllabicity (MULTI) are other commonly used indicators for determining CBO. Studies using these indicators have shown that CBO is not correlated with meaningful expressive word onset age (Fagan, 2009; Lang et al., 2021), and there is large variation in the age difference between CBO and meaningful expressive word onset, with time intervals ranging from 2 to 11 months (Lang et al., 2021). However, Jung and Houston (2020) examined the relationship between CBO (using canonical babbling ratio as the indicator) and language abilities in cochlear-implanted children with an average activation age of 21 months and found that CBO predicted expressive vocabulary 24 months after cochlear implant activation.

Thus, existing research does not support the conclusion that earlier babbling leads to better expressive language. One reason may be that different indicators lead to substantial variation in CBO across studies. For example, the median CBO age in McGillion et al.' s (2017) study, which used stable production of two consecutive supraglottal consonants as the criterion, was 10 months, whereas the maximum average CBO age in Lang et al.' s (2021) study, which used six indicators including canonical babbling ratio and MULTI, was 8 months. Different CBOs may represent different underlying abilities in infants. Although CBO captures the emergence of oral-motor readiness for word production (Oller, 2000), for many infants, time points at or before 9 months are too early and insufficient to indicate higher levels of oral-motor ability (Lang et al., 2021). However, infants who meet the criterion of stable production of two different consonants may have better oral-motor abilities, enabling them to focus attention and remember word forms and meanings (McCune & Vihman, 2001), which partially explains why studies using different consonant production as the CBO criterion have more positive results. Additionally, prerequisites for meaningful word production and development may extend beyond phonetic abilities alone. Infants' cognitive and communicative skills may facilitate word learning, and asynchronous development of these skills may affect meaningful expressive word onset age (Lang et al., 2021). Based on these results and analyses, investigations of canonical babbling onset and its relationship with later expressive vocabulary need to comprehensively consider factors such as communication, cognition, and different CBO indicators.

2.3.3 Consonant Diversity Consonant diversity refers to consonant inventory and diversity of key consonants used in communication acts. Studies on consonant inventory as a predictor of language abilities have found that consonant inventory in TD infants at 11 months predicts mean length of utterance in morphemes (MLU) and vocabulary diversity in spontaneous language samples at 24 months but does not predict expressive vocabulary on the MCDI. Consonant inventory at 18 months predicts MLU, vocabulary diversity, and expressive vocabulary at 24 months, with consonant inventory at 18 months accounting for more variance than at 11 months. However, consonant inventory at 7 months does not predict MLU, vocabulary diversity, or expressive

vocabulary at 24 months (Gerhold et al., 2020). Researchers suggest this may be because 7-month-old infants are too young, and their consonant inventory is not yet sufficient to predict subsequent expressive language. A subsequent study of TD infants and infants with mild-to-moderate hearing loss who wore hearing aids before 6 months showed that consonant inventory at 18 months correlated with expressive vocabulary at 24 months (Persson et al., 2022). In children with cleft palate, consonant inventory at 9 months before surgical repair negatively correlated with the number of different words children could use at 39 months, whereas after surgery at 21 months, the correlation was positive. Researchers suggest that preoperative results may not represent true outcomes. Additionally, consonant inventory at 21 months negatively correlated with MLU at 39 months (Beckett, 2017). From these opposing results regarding consonant inventory' s prediction of MLU, this indicator may not be a good predictor of MLU.

More studies have begun to use diversity of key consonants used in communication acts as an indicator to explore its relationship with language, with relevant research primarily focusing on children with ASD and TD children. Studies of TD toddlers have found that diversity of key consonants at 20 months predicts expressive language at 33 months but does not predict receptive language abilities. Diversity of key consonants at 14 months does not predict later receptive or expressive language (Watt et al., 2006). Combined with the above results on consonant inventory, the predictive effect of consonant diversity on expressive language becomes more apparent around 18 months. Studies of children with ASD have shown that diversity of key consonants in ASD children with an average age of 21 months predicts expressive language at age 3 (Wetherby et al., 2007). Subsequent studies of ASD toddlers and toddlers with fragile X syndrome have yielded the same results (Fielding-Gebhardt & Warren, 2019; McDaniel et al., 2020a; Saul & Norbury, 2020; Woynaroski et al., 2017; Yoder et al., 2015). Thus, numerous studies indicate that consonant diversity predicts expressive language. These results partially support the notion that consonant production may share motor capabilities with speech and may indicate that children are attempting to produce words (Woynaroski et al., 2016).

Comprehensive review of existing research on the phonetic quality of infant speech-like vocalizations reveals that canonical syllable frequency or ratio and consonant diversity show relatively consistent results in predicting language abilities, particularly expressive language, though age may be a moderating factor. In contrast, canonical babbling onset shows more variable results in predicting language abilities, likely influenced by other factors.

2.4 Communicative Vocalizations

The social pragmatic theory of word learning posits that intentional prelinguistic communication is a prerequisite for intersubjective symbolic use (Tomasello, 2008). Communicative vocalizations refer to vocalizations combined with eye gaze or gestures that are clearly directed toward a communication partner (Mc-

Daniel et al., 2020a, 2020b). Research on TD children and children with language development disorders indicates that intentional communicative vocalizations effectively predict vocabulary development. Donnellan et al. (2020) assessed the relationship between vocalization frequency in mother-infant interactions in natural contexts in 11–12-month-old infants and expressive vocabulary at 15–24 months (average 19 months). Results showed that when not considering whether vocalizations were coordinated with eye gaze, more canonical syllable vocalizations predicted more expressive vocabulary, whereas more non-canonical syllable vocalizations predicted fewer expressive words. When vocalizations were coordinated with eye gaze, more total vocalizations (including both canonical and non-canonical) predicted more expressive vocabulary, indicating that intentional communicative vocalization behaviors predicted expressive vocabulary. Therefore, researchers suggest that the most valuable predictor of later language in infants is the frequency of early coordinated gaze communicative vocalizations.

The predictive value of communicative vocalizations for expressive language has also been verified in children with ASD. McDaniel et al. (2019) studied children with ASD with an average age of 35 months who were nonverbal or had fewer than 20 expressive words and found that the number of communicative vocalizations in ASD toddlers effectively predicted expressive language at 4, 8, and 12 months later. Subsequent research also supported this result, showing that the number of vocal communicative acts and the proportion of communicative vocalizations in semi-structured play interactions in ASD children with an average age of 23 months effectively predicted expressive language 12 months later (McDaniel et al., 2020a). However, these studies did not specifically analyze different types of vocal communicative acts (i.e., vocalizations coordinated with eye gaze, vocalizations coordinated with gestures, vocalizations coordinated with both eye gaze and gestures). Therefore, it cannot be determined which specific type of vocal communicative act predicts expressive language in ASD children. Children with ASD have impairments in eye contact (Bradshaw et al., 2021), leading to the reasonable inference that vocalizations coordinated with eye gaze may be fewer in ASD children. Recent research has confirmed that children with ASD with vocabulary below 20 words have fewer vocalizations coordinated with eye gaze and vocalizations coordinated with both eye gaze and gestures than TD children with comparable vocabulary, but vocalizations coordinated with gestures alone show no significant difference (Murillo et al., 2021). Thus, specific vocal communicative acts predicting expressive language in ASD children may differ from those in TD children, necessitating further research.

In summary, communicative vocalizations are effective predictors of later expressive language. Canonical syllable communicative vocalizations may be stronger predictors of expressive language than non-canonical syllable communicative vocalizations. Additionally, communicative vocalizations predict expressive language in both TD children and children with language development disorders. However, which vocal communicative acts predict expressive language development in specific types of language disorder requires further investigation.

3. Mechanisms of Speech-Like Vocalizations Facilitating Language Development

Early research demonstrated that a prerequisite for language production is the development of specific motor skills, such as the ability to produce syllables (Vihman et al., 1985). Recent research has revealed that interactive patterns between infant speech-like vocalizations and social responses also provide possible sources for language development. The following sections explore three possible mechanisms by which speech-like vocalizations facilitate language development: vocalizations provide the basis for language production, vocalizations create an effective learning state, and vocalizations elicit social responsive behaviors.

3.1 Vocalizations Provide the Basis for Language Production

The first mechanism emphasizes that the phonetic and functional features of speech-like vocalizations form the foundation for speech production. There is a natural logic that early features in development form the basis for later features, and individuals tend to maintain consistency between the emergence order and developmental order of behavioral capacities (Newman, 2016). From this logic, protophones are generated by strong endogenous motivation shortly after birth (Long et al., 2022), and their emerging speech features can provide a foundation for later speech development.

3.1.1 Phonetic Continuity First, the earliest speech-like vocalizations (such as quasi-vowels) are mostly produced during infants' solitary exploratory play, are highly endogenous, and result from deep natural selection (Long et al., 2020). Early vocalizations form the basis for subsequent vocalizations; without this foundation, subsequent canonical syllables, communicative vocalizations, and vocal imitation cannot emerge (Oller et al., 2019). Oller et al. (2019) found that both preterm infants (born two months early) and full-term infants produced a minimum of 1.4 protophones per minute, greatly exceeding another common early vocalization—crying. Without the earliest protophones, the basis for adults to engage in vocal interaction with infants would be greatly reduced, as adults are more likely to respond to infants' speech-like vocalizations than to non-speech-like vocalizations (Warlaumont et al., 2014).

Second, there is strong phonetic continuity between canonical syllables or babbling and early words, as the vast majority of words in natural language consist of canonical syllables (Oller et al., 2019). Babbling syllables have similar consonant-vowel combinations to early word syllables, such as coronal consonants combined with front vowels (Lahrouchi & Kern, 2018). Approximately 50%–80% of consonants in the first 10 words are the same as those in speech-like vocalizations (Keren-Portnoy et al., 2009). The distribution of labial, coronal, and dorsal consonants does not differ significantly between babbling and the first 100 words (van der Feest et al., 2020). Therefore, during the initial stage of word learning, when expressive vocabulary is less than 100 words, infant word acquisition better fits the phonological dominance hypothesis, which posits that

word learning depends more on the articulatory system and speech production abilities.

3.1.2 Functional Flexibility Functional flexibility refers to the ability to use any word or sentence to express different emotions in different contexts, which is essential for language communication and a fundamental property of language (Jhang & Oller, 2017). Infants'speech-like vocalizations already exhibit functional flexibility, meaning that infants produce vocalizations with multiple emotional meanings in different contexts (Oller et al., 2013). Oller et al. (2013) found that 3–4-month-old infants' speech-like vocalizations have stronger functional flexibility than non-speech-like vocalizations (such as crying and laughter). Infants' speech-like vocalizations can express positive, neutral, and negative emotional functions, whereas crying or laughter functions are more rigid and can only express negative or positive emotional functions. Further research has found that functional flexibility in speech-like vocalizations can appear in infants as young as 1 month (Jhang & Oller, 2017). Therefore, the functional flexibility exhibited in infant speech-like vocalizations may be a crucial step in language development.

3.2 Vocalizations Create an Effective Learning State

The second mechanism emphasizes the impact of speech-like vocalizations on infants' own language learning state, suggesting that infants' speech-like vocalizations may help them create an effective state conducive to language learning and signal this to others. This effective state involves focused attention, effective speech perception, and high-motivation social engagement.

3.2.1 Attentional State Object-directed vocalizations (ODV) refer to sounds infants produce when looking at nearby or held objects. ODV indicates that infants are in a state of focused attention, which helps them learn object features and the connection between objects and corresponding symbols (Goldstein et al., 2010). Goldstein et al. (2010) used experimental research to verify this mechanism. The experiment included an object exploration test and a preferential looking task. First, the object exploration test identified high-vocalization objects and low-vocalization objects for 12-month-old infants. Then, in the preferential looking task, original versions and shape-distorted versions of high- or low-vocalization objects were presented in pairs while observing infants' looking time. The logic of this manipulation was that infants who had learned more visual features of high-vocalization objects would show greater preference for the shape-distorted versions when presented simultaneously with the original versions, looking at them longer. Results showed that infants looked longer at the novel versions of high-vocalization objects and learned more about their features. This supports the view that ODV marks a state of focused attention and that infants may be particularly sensitive to perceptual information following vocalizations. The second experiment further linked the attentional state created by ODV to language learning, finding that

infants were more likely to associate linguistic symbols with objects following ODV than without vocalization.

3.2.2 Perceptual State Speech-like vocalizations facilitate infants' perception of phonetic categories in speech streams. Skilled babbling can serve as a perceptual "filter," drawing infants' attention to words that match sounds they can reliably produce (Laing & Bergelson, 2020). Vilain et al. (2019) used an intersensory matching procedure to explore the impact of 6- and 9-month-old infants' babbling production abilities on consonant perception. The study found that compared to infants without babbling production abilities, infants who could stably produce babbling and repeatedly produce consonants /b/ and /d/ could match speech containing these consonants with visual information. When the auditory consonant materials were changed to /v/ and /z/, which infants in both age groups had not produced, neither group could complete the intersensory matching task.

Further research indicates that as age increases, infants with stronger speech-like vocalization abilities are more likely to perceive consonants that are difficult to articulate. Lorenzini and Nazzi (2020) used a head-turn preference procedure to explore 11- and 14-month-old infants' perception of familiar words containing consonants with different articulation difficulties. They divided infants in both age groups into high consonant ability and low consonant ability groups based on their existing consonant abilities. The study found that 11-month-old high consonant ability infants showed longer perception times for words, but no significant effect of consonant articulation difficulty was found. However, 14-month-old high consonant ability infants showed longer perception times for words containing consonants with higher articulation difficulty. In summary, infants' speech-like vocalization abilities facilitate their speech perception and thereby promote language learning.

3.2.3 Motivational State It is crucial for language acquisition that infants discover they can influence others' behavior rather than passively participating. Infants learn the social functions of speech-like vocalizations between 2-5 months. In the still-face task, 5-month-old infants increase speech-like vocalizations to expect adults with still faces to respond to their vocalizations, indicating that infants at this age can consciously use speech-like vocalizations to influence adults (Bigelow et al., 2018; Elmlinger et al., 2022). Recognizing the effectiveness of speech-like vocalizations in eliciting responses—that is, infants' ability to actively guide interactions and cause caregivers to respond—is an important step toward language communication (Wu & Gros-Louis, 2017). Wu and Gros-Louis (2017) tested 10-month-old infants' vocalizations under three conditions (mother responds to infant communicative acts and shares attention and interest in toys with infant, mother only responds to infant communicative acts without attending to toys, mother ignores infant communicative acts) and their relationship with language abilities at 15 months. Results showed that infants who could produce more vocalizations to influence maternal behavior

when mothers ignored their vocal communicative acts had better language abilities at 15 months, indicating that infants who are aware of the social impact of vocalizations have an advantage in language development.

3.3 Vocalizations Elicit Social Responsive Behaviors

The third mechanism emphasizes the social function of speech-like vocalizations, whereby infants intentionally or unintentionally elicit responsive behaviors from social partners and create social interactions, thereby facilitating their language development (Albert et al., 2018; Cohen & Billard, 2018; Donnellan et al., 2020). During free play, social partners respond to approximately 40–70% of infants' speech-like vocalizations (e.g., Athari et al., 2021; Gros-Louis et al., 2006; Lieberman et al., 2019), with verbal responses significantly exceeding non-verbal responses (Gros-Louis & Miller, 2018). The contingent, directive, and scaffolding features of social responsive behaviors increase infants' chances of successfully learning language (Tamis-LeMonda et al., 2014).

3.3.1 Contingent Responsiveness Social partners (primarily caregivers) provide contingent responses that are dependent on infants' prior speech-like vocalizations and temporally continuous with them (e.g., Fagan & Doveikis, 2017; Gros-Louis & Miller, 2018). First, social partners' contingent responses depend on infants' prior speech-like vocalizations. For example, when an infant vocalizes "ba" while playing with a ball, the mother responds with "ball." This dependency may increase the likelihood that infants hear words corresponding to the most salient objects and events, thereby promoting statistical learning (Tamis-LeMonda et al., 2014). Second, most maternal responses to infant vocalizations occur within 2 seconds, with almost all responses occurring within 3 seconds (Pretzer et al., 2019; Van Egeren et al., 2001; Yoo et al., 2018). This close temporal connection is crucial for infants' language development because the likelihood of two events being linked depends on their co-occurrence within a short time window (Rovee-Collier, 1995). Contingent responses strengthen infants' attention, leading to strong preferences for objects that caregivers manipulate, enabling infants to better learn from infant-directed speech (Masek et al., 2021; Mason et al., 2019). Event-related potential studies have preliminarily shown that compared to non-contingent verbal responses, infants show attention to and possibly engage in lexical-semantic processing of caregivers' contingent verbal responses (Lam-Cassettari et al., 2021).

The benefits of social partners' contingent responses for infants' phonetic and lexical development have been confirmed. Experimental studies have shown that compared to non-contingent verbal responses, when mothers provide contingent responses to 9.5-month-old infants' speech-like vocalizations, infants' vocalizations incorporate mothers' phonetic patterns (Goldstein & Schwade, 2008). Observational studies of 12-month-old infants have shown that compared to no contingent response, when caregivers contingently respond to infants' CV vocalizations, infants produce more CV syllables, and when they contingently respond

to infants'V vocalizations, infants'V syllable production decreases (Gros-Louis & Miller, 2018). This indicates that the feedback loop between infant vocalizations and adult contingent responses can help infants produce more advanced sounds and subsequently more complex phonetic patterns in first words. Research on children with ASD also shows that parents' contingent verbal responses predict expressive vocabulary in children with ASD (McDaniel et al., 2017).

3.3.2 Directive Information When caregivers respond to infants' speech-like vocalizations, directive information is generally reflected in their language through object naming, description, and questioning about objects or events (Tamis-LeMonda et al., 2014). Observational and experimental studies of 8-14-month-old infants have shown that compared to less informative verbal responses (such as statements unrelated to the infant' s environment, affirming or prohibiting infant behavior), mothers provide more verbal responses containing directive information (such as naming objects infants attend to, describing infants' states or objects they attend to, asking questions about objects infants attend to) in response to infants' speech-like vocalizations (Albert et al., 2018; Fagan & Doveikis, 2019; Hong, 2017). Compared to less informative responses, caregivers' verbal responses containing directive information correlate with infants' concurrent receptive and expressive vocabulary (Lopez et al., 2020) or predict infants' future expressive vocabulary (Goldstein & Schwade, 2010).

Further analysis has found that among caregivers' directive information responses, responses to infants' ODV speech-like vocalizations are significantly more frequent than responses to non-ODV speech-like vocalizations (Albert, 2021; Albert et al., 2018), and infants produce significantly more ODV vocalizations than vocalizations directed toward adults or without specific direction (Hong, 2017). A study on the predictive value of mothers' verbal responses to 9-month-old infants' ODV speech-like vocalizations for vocabulary at 15 months showed that verbal responses containing object names that infants attended to significantly predicted later expressive vocabulary (Goldstein & Schwade, 2010). That is, mothers who respond to infants' ODV speech-like vocalizations by providing object names may help infants identify the connection between vocalizations and objects in the environment, thereby promoting word learning.

3.3.3 Scaffolding Support Social partners' responsive behaviors are coordinated with infants' developmental levels and provide scaffolding for infant language development. Infants' speech-like vocalizations help caregivers produce simpler, more learnable language (Elmlinger et al., 2019, August). Caregivers' language in contingent responses to infants' speech-like vocalizations has fewer word types than infant-directed but non-contingent language (Elmlinger et al., 2021), specifically manifested as caregivers using fewer different words, fewer words per utterance, a higher proportion of single-word utterances (Elmlinger et al., 2022, July; Elmlinger et al., 2019), and more repeated words (Elmlinger et al., 2021). The repetitive feature of maternal language input at infant age 7 months predicts expressive vocabulary at 24 months (Newman et al., 2016).

Compared to having different target words in successive sentences, 2-year-old toddlers more easily acquire target words when they appear repeatedly in successive sentences (Schwab & Lew-Williams, 2016). Maternal repetitive utterances at toddler age 21 months predict vocabulary at 30 months (Casla et al., 2022). Thus, infant speech-like vocalizations elicit simplified adult language that promotes language learning in infants and toddlers.

Existing research on mechanisms by which speech-like vocalizations may facilitate language development has deepened understanding of why speech-like vocalizations can promote language development. However, most studies have explored either infant speech-like vocalizations or social responses separately. Research involving dynamic interactions between the two has also focused on how the phonetics of infant speech-like vocalizations become more mature with social responses. Future research could attempt to explore deeper dynamic interactions.

4. Future Directions

4.1 Causal Relationship Between Speech-Like Vocalizations and Language Development

The above review shows that existing research has found correlations between early speech-like vocalizations and subsequent language development in infants and toddlers; however, experimental studies verifying the causal relationship are lacking. First, future research could use experimental methods, focusing on infants in the transition period from speech-like vocalizations to first words, to manipulate social contingent response methods and explore the relationship between infant speech-like vocalizations and word learning. Recent research has begun to explore the impact of manipulating parental contingent responses on infant vocalizations and expressive vocabulary. An experimental group where parents produced more infant-directed speech and had more interactive engagement with infants showed more babbling and expressive vocabulary than a control group (Ferjan Ramírez et al., 2019). However, this study did not manipulate parents to respond only to infants' speech-like vocalizations, nor did it distinguish different types of speech-like vocalizations. As previously discussed, different speech-like vocalizations have different effects on language, making it necessary to control for speech-like vocalization types in future research. Second, the causal relationship between speech-like vocalizations and word learning could be determined by intervening on speech-like vocalizations in prelinguistic children with language disorders.

4.2 Moderating Effects of Related Factors

The above results on different speech-like vocalizations predicting language show that some indicators have inconsistent or even highly variable results in predicting language development. The reason may be the influence of other related factors. For example, the severity of social communication disorders affects fu-

ture language abilities in children with ASD to some extent (Thurm et al., 2015). Does social communication disorder moderate the relationship between speech-like vocalizations and language abilities? Cognition affects language development and may be associated with language abilities from as early as 3 months (Perszyk & Waxman, 2018), but whether cognition moderates the relationship between speech-like vocalizations and language development remains unclear. The above review found that age of TD infants may affect the relationship between speech-like vocalizations and language development. Approximately 60% of children with ASD are nonverbal or communicate with only minimal words before age 5 (Maltman et al., 2021), and approximately 30% of children with ASD remain extremely limited in speech at age 5 and beyond (Tager-Flusberg & Kasari, 2013). Is age a moderating factor in the prediction of language abilities by speech-like vocalizations in children with ASD? Future research could explore the moderating effects of factors such as disorder severity, cognition, and age on the relationship between speech-like vocalizations and language development.

4.3 Influence of Language Ability Testing Methods

First, the inconsistent prediction of expressive and receptive language by speech-like vocalizations may be due to language ability testing methods. Future research could control for speech-like vocalization indicator types, child age, and disorder type, using different language testing methods to explore whether speech-like vocalizations consistently predict receptive and expressive language. Second, speech-like vocalization indicators themselves may affect their relationship with language abilities. For example, research on canonical babbling ratio criteria for predicting future language difficulties shows that criteria of 0.14 and 0.15 have similar predictive sensitivity but differ substantially in specificity. This suggests that the criterion for canonical babbling ratio may be a factor affecting the prediction of language by speech-like vocalizations (Nyman et al., 2021). Future research could also explore the influence of speech-like vocalization indicators themselves.

4.4 Most Valuable Speech-Like Vocalizations for Language Disorder Interventions

A long-term goal for children with language disorders is to improve their language abilities, and intervening on the speech-like vocalizations that most affect their language abilities is an important way to achieve this goal. As discussed above, different speech-like vocalizations have different predictive validity for language development. Therefore, exploring which speech-like vocalizations have the most predictive value for different types of disorders is crucial. Although some studies have explored whether speech-like vocalizations in children with ASD predict expressive language (Blume et al., 2021; McDaniel et al., 2018), they have not analyzed which specific speech-like vocalizations more effectively predict expressive language. As previously mentioned, children with ASD have impairments in eye contact, so communicative acts combining eye

gaze with vocalizations may also be correspondingly reduced. Do communicative vocalizations coordinated with eye gaze and those coordinated with gestures predict language equally? What is the most predictive vocal communicative act for children with ASD? Future research could further explore these questions.

4.5 Dynamic Interaction Between Speech-Like Vocalizations and Social Responses

First, explore the dynamic interaction between speech-like vocalizations and social responses in TD children. Some research has examined how the interaction between infant speech-like vocalizations and social responses affects infant phonetics, showing that caregivers are more likely to respond to infants' speech-like vocalizations, and infants' vocalizations become more mature after caregiver responses (Warlaumont et al., 2014). However, deeper dynamic interactions remain unclear, such as changes in infants' attention after social contingent responses, changes in object processing, changes in matching adult responses with referents, and how caregivers adjust their behaviors in response to changes in infants' speech-like vocalizations and related behaviors. Future research could explore these aspects, especially longitudinally investigating the development of infant vocalizations and changes in social response interactions.

Second, explore the dynamic interaction between speech-like vocalization disorders and social responses. Numerous studies support that children with ASD, speech sound disorders, Rett syndrome, and other conditions show delayed or atypical development of speech-like vocalizations (Bartl-Pokorny et al., 2022; Garrido et al., 2017; Overby et al., 2020; Roche et al., 2018; Yankowitz et al., 2022). However, the dynamic interaction between these children' s speech-like vocalizations and social responses is unclear. A recent study on social contingency in infants with delayed babbling development and TD infants found that parents did not differ significantly in the number of responses to speech-like vocalizations, but parents' responses to infants with delayed babbling development were more often acknowledgments (such as "yes"), whereas responses to TD infants were more often imitations or expansions of their vocalizations (Lieberman et al., 2019). Compared to maternal regulatory language (such as "look here," "put it down"), TD infants vocalized more in response to maternal referential language (such as "that' s a puppy") (Kuchirko et al., 2018). Additionally, there is a significant correlation between the time interval of caregivers' contingent responses to TD children and TD children' s responses to caregivers' contingent responses, whereas this relationship does not exist between children with language delay and their caregivers. The average interval of coordinated responses between 9-month-old infants and caregivers is 2 seconds, and each unit increase in interval reduces language abilities at ages 2–3 by 0.53 units (Northrup & Iverson, 2015). Future research could further explore how speech-like vocalizations in children with language disorders affect contingent responses and how contingent responses in turn affect their speech-like vocalizations.

5. Conclusion

The quantity or frequency of speech-like vocalizations, canonical syllable ratio or frequency, consonant diversity, and communicative vocalizations in infants and toddlers predict language development. The following perspectives explain how speech-like vocalizations facilitate language development: speech-like vocalizations provide a phonetic basis for early words and prepare for functional flexibility in language; speech-like vocalizations mark infants' attentional state, speech perception state, and motivational state for social interaction that are conducive to learning; and speech-like vocalizations elicit social responses that provide contingent support, directive information, and scaffolding for language development. Future research should explore the causal relationship between speech-like vocalizations and language development, the most valuable speech-like vocalizations for language intervention in children with language disorders, and the mechanisms underlying dynamic interactions between TD children and children with language disorders and social responses, to provide deeper explanations of how speech-like vocalizations develop into language. Research findings will help promote language development in TD children and inform interventions for children with language disorders.

References

Albert, R. R. (2021). Teacher talk: Infant vocal cues affect non-lead infant teachers' responding. *Early Childhood Research Quarterly*, 55, 326–335.

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2018). The social functions of babbling: Acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science*, 21(5), Article e12641. https://doi.org/10.1111/desc.12641

Athari, P., Dey, R., & Rvachew, S. (2021). Vocal imitation between mothers and infants. *Infant Behavior and Development*, 63, Article 101531. https://doi.org/10.1016/j.infbeh.2021.101531

Bartl-Pokorny, K. D., Pokorny, F. B., Garrido, D., Schuller, B. W., Zhang, D., & Marschik, P. B. (2022). Vocalisation repertoire at the end of the first year of life: An exploratory comparison of Rett syndrome and typical development. *Journal of Developmental and Physical Disabilities*, 34(6), 1053–1069. https://doi.org/10.1007/s10882-022-09837-w

Beckett, L. N. (2017). Profiling speech and language outcomes of children with cleft palate at 39 months of age: Examining predictors and identifying speech and language characteristics (Unpublished master' s thesis). The University of Utah, Salt Lake City.

Bigelow, A. E., Power, M., Bulmer, M., & Gerrior, K. (2018). The effect of maternal mirroring behavior on infants' early social bidding during the still-face task. *Infancy*, 23(3), 367–385.

Blume, J., Wittke, K., Naigles, L., & Mastergeorge, A. M. (2021). Language

growth in young children with autism: Interactions between language production and social communication. *Journal of Autism and Developmental Disorders*, 51(2), 644–665.

Bradshaw, J., McCracken, C., Pileggi, M., Brane, N., Delehanty, A., Day, T., ··· Wetherby, A. (2021). Early social communication development in infants with autism spectrum disorder. *Child Development*, 92(6), 2224–

Brookman, R., Kalashnikova, M., Conti, J., Xu Rattanasone, N., Grant, K. A., Demuth, K., & Burnham, D. (2020). Depression and anxiety in the postnatal period: An examination of infants' home language environment, vocalizations, and expressive language abilities. *Child Development*, 91(6), e1211–e1230.

Buder, E. H., Warlaumont, A. S., Oller, D. K., Peter, B., & MacLeod, A. (2013). An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. In B. Peter, & A. A. N. MacLeod (Ed.), *Comprehensive perspectives on child speech development and disorders: Pathways from linguistic theory to clinical practice* (pp. 103–134). Hauppauge, NY: Nova Science Publisher.

Casla, M., Méndez-Cabezas, C., Montero, I., Murillo, E., Nieva, S., & Rodríguez, J. (2022). Spontaneous verbal repetition in toddler-adult conversations: A longitudinal study with Spanish-speaking two-year-olds. *Journal of Child Language*, 49(2), 266–301.

Chapman, K. L., Hardin-Jones, M., & Halter, K. A. (2003). The relationship between early speech and later speech and language performance for children with cleft lip and palate. *Clinical Linguistics & Phonetics*, 17(3),

Cohen, L., & Billard, A. (2018). Social babbling: The emergence of symbolic gestures and words. *Neural Networks*, 106, 194–204.

Donnellan, E., Bannard, C., Mcgillion, M. L., Slocombe, K. E., & Matthews, D. E. (2020). Infants' intentionally communicative vocalizations elicit responses from caregivers and are the best predictors of the transition to language: A longitudinal investigation of infants'vocalisations, gestures, and word production. *Developmental Science*, 23(1), Article e12843. https://doi.org/10.1111/desc.12843

Elmlinger, S., Goldstein, M., & Casillas, M. (2022, July). Immature vocalizations simplify the speech of Tseltal Mayan and US caregivers. Paper presented at the proceedings of the Annual Meeting of the Cognitive Science Society, Toronto, Canada.

Elmlinger, S. L., Park, D., Schwade, J. A., & Goldstein, M. H. (2021). Comparing word diversity versus amount of speech in parents' responses to infants' prelinguistic vocalizations. *IEEE Transactions on Cognitive and Developmental Systems*, 14(3), 1036–1045.

Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019). The ecology of prelinguistic vocal learning: Parents simplify the structure of their speech in response to babbling. *Journal of Child Language*, 46(5), 998–1011.

Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019, August). Babbling elicits simplified caregiver speech: Findings from natural interaction and simulation. Paper presented at the 2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), Oslo, Norway.

Elmlinger, S. L., Schwade, J. A., Vollmer, L., & Goldstein, M. H. (2022). Learning how to learn from social feedback: origins early vocal development. *Developmental Science*, Article e13296. https://doi.org/10.1111/desc.13296

Fagan, M. K. (2009). Mean length of utterance before words and grammar: Longitudinal trends and developmental implications of infant vocalizations. *Journal of Child Language*, 36(3), 495–527.

Fagan, M. K., & Doveikis, K. N. (2017). Ordinary interactions challenge proposals that maternal verbal responses shape infant vocal development. *Journal of Speech Language and Hearing Research*, 60(10), 2819–2827.

Fagan, M. K., & Doveikis, K. N. (2019). What mothers do after infants vocalize: Implications for vocal development or word learning? *Journal of Speech Language and Hearing Research*, 62(8), 2680–2690.

Ferjan Ramírez, N., Lytle, S. R., Fish, M., & Kuhl, P. K. (2019). Parent coaching at 6 and 10 months improves language outcomes at 14 months: A randomized controlled trial. *Developmental Science*, 22(3), Article e12762. https://doi.org/10.1111/desc.12762

Fielding-Gebhardt, H., & Warren, S. F. (2019). Early predictors of later expressive language in boys with fragile X syndrome. *American Journal on Intellectual and Developmental Disabilities*, 124(1), 11–24.

Garrido, D., Watson, L. R., Carballo, G., Garcia-Retamero, R., & Crais, E. R. (2017). Infants at-risk for autism spectrum disorder: Patterns of vocalizations at 14 months. *Autism Research*, 10(8), 1372–1383.

Gerhold, K., Torrington Eaton, C., Newman, R. S., & Bernstein Ratner, N. (2020). Early phonological predictors of toddler language outcomes. *Folia Phoniatrica et Logopaedica*, 72(6), 442–453.

Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy*, 15(4), 362–391.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19(5), 515–523.

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *Oxford handbook of developmental and comparative neuroscience* (pp. 708–729). Oxford University Press.

Gros-Louis, J., & Miller, J. L. (2018). From 'ah' to 'bah' : Social feedback loops for speech sounds at key points of developmental transition. *Journal of Child Language*, 45(3), 807–825.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, 30(6), 509–516.

Ha, S., Kim, H., & Lee, Y. (2022). Predicting speech and language development at 18-20 months from automated language ENvironment analysis measures. *Communication Sciences & Disorders*, 27(2), 384–393.

Hamrick, L. R., Seidl, A., & Tonnsen, B. L. (2019). Acoustic properties of early vocalizations in infants with fragile X syndrome. *Autism Research*, 12(11), 1663–1679.

Hong, Y. (2017). Parental verbal responsiveness during prelinguistic vocal development: Variability and association with language outcomes (Unpublished master' s thesis). University of Iowa.

Jhang, Y., & Oller, D. K. (2017). Emergence of functional flexibility in infant vocalizations of the first 3 months. *Frontiers in Psychology*, 8, Article 300. https://doi.org/10.3389/fpsyg.2017.00300

Jung, J., & Houston, D. (2020). The relationship between the onset of canonical syllables and speech perception skills in children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 63(2), 393–

Keren-Portnoy, T., Majorano, M., & Vihman, M. M. (2009). From phonetics to phonology: The emergence of first words in Italian. *Journal of Child Language*, 36(2), 235–267.

Kim, H., & Ha, S. (2022). Relation between early vocalizations and words. *Communication Sciences & Disorders*, 27(1), 1–13.

Kuchirko, Y., Tafuro, L., & Tamis LeMonda, C. S. (2018). Becoming a communicative partner: Infant contingent responsiveness to maternal language and gestures. *Infancy*, 23(4), 558–576.

Lahrouchi, M., & Kern, S. (2018). From babbling to first words in Tashlhiyt language acquisition: Longitudinal two-case studies. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique*, 63(4), 493–526.

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

Lam-Cassettari, C., Peter, V., & Antoniou, M. (2021). Babies detect when the timing is right: Evidence from event-related potentials to a contingent mother-infant conversation. *Developmental Cognitive Neuroscience*, 48, Article 100923. https://doi.org/10.1016/j.dcn.2021.100923

Lang, S., Willmes, K., Marschik, P. B., Zhang, D., & Fox-Boyer, A. (2021). Prelexical phonetic and early lexical development in German-acquiring infants: Canonical babbling and first spoken words. *Clinical Linguistics & Phonetics*, 35(2), 185–200.

Lieberman, M., Lohmander, A., & Gustavsson, L. (2019). Parents' contingent responses in communication with 10-month-old children in a clinical group with typical or late babbling. *Clinical Linguistics & Phonetics*, 33(10-11), 1050–1062.

Lohmander, A., Holm, K., Eriksson, S., & Lieberman, M. (2017). Observation method identifies that a lack of canonical babbling can indicate future speech and language problems. *Acta Paediatrica*, 106(6), 935–943.

Long, H. L., Bowman, D. D., Yoo, H., Burkhardt-Reed, M. M., Bene, E. R., & Oller, D. K. (2020). Social and endogenous infant vocalizations. 15(8), Article e0224956. https://doi.org/10.1371/journal.pone.0224956

Long, H. L., Ramsay, G., Griebel, U., Bene, E. R., Bowman, D. D., Burkhardt-Reed, M. M., & Oller, D. K. (2022). Perspectives on the origin of language: Infants vocalize most during independent vocal play but produce their most speech-like vocalizations during turn taking. *Plos One*, 17(12), Article e0279395. https://doi.org/10.1371/journal.pone.0279395

Lopez, L. D., Walle, E. A., Pretzer, G. M., & Warlaumont, A. S. (2020). Adult responses to infant prelinguistic vocalizations are associated with infant vocabulary: A home observation study. *Plos One*, 15(11), Article e0242232. https://doi.org/10.1371/journal.pone.0242232

Lorenzini, I., & Nazzi, T. (2020). Consonant production and early word-form processing are linked at 14 months of age (Abstract). In L. Taschenberger (Ed.), *Book of abstracts of 2nd workshop on speech perception and production across the lifespan (SPPL2020)* (pp. 53–55). London, UK.

Lyakso, E. E., Frolova, O. V., & Grigorev, A. S. (2014). Infant vocalizations at the first year of life predict speech development at 2-7 years: Longitudinal study. *Psychology*, 5(12), 1433–1445.

Majorano, M., Vihman, M. M., & DePaolis, R. A. (2014). The relationship between infants' production experience and their processing of speech. *Language Learning and Development*, 10(2), 179–204.

Maltman, N., DaWalt, L. S., Hong, J., & Mailick, M. (2021). Brief report: Socioeconomic factors associated with minimally verbal status in individuals with ASD. *Journal of Autism and Developmental Disorders*, 51(6),

Masek, L. R., McMillan, B. T., Paterson, S. J., Tamis-LeMonda, C. S., Golinkoff, R. M., & Hirsh-Pasek, K. (2021). Where language meets attention: How contingent interactions promote learning. *Developmental Review*, 60, Article 100961. https://doi.org/10.1016/j.dr.2021.100961

Mason, G. M., Kirkpatrick, F., Schwade, J. A., & Goldstein, M. H. (2019). The

role of dyadic coordination in organizing visual attention in 5-month-old infants. *Infancy*, 24(2), 162–186.

McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical development: A productivity approach. *Journal of Speech, Language, and Hearing Research*, 44(3), 670–684.

McDaniel, J., & Schuele, C. M. (2021). When will he talk? An evidence-based tutorial for measuring progress toward use of spoken words in preverbal children with autism spectrum disorder. *American Journal of Speech-Language Pathology*, 30(1), 1–18.

McDaniel, J., Slaboch, K. D. A., & Yoder, P. (2018). A meta-analysis of the association between vocalizations and expressive language in children with autism spectrum disorder. *Research in Developmental Disabilities*, 72, 202–213.

McDaniel, J., Woynaroski, T., Keceli-Kaysili, B., Watson, L. R., & Yoder, P. (2019). Vocal communication with canonical syllables predicts later expressive language skills in preschool-aged children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 62(10), 3826–3833.

McDaniel, J., Yoder, P., Estes, A., & Rogers, S. J. (2020a). Predicting expressive language from early vocalizations in young children with autism spectrum disorder: Which vocal measure is best? *Journal of Speech, Language, and Hearing Research*, 63(5), 1509–1520.

McDaniel, J., Yoder, P., Estes, A., & Rogers, S. J. (2020b). Validity of vocal communication and vocal complexity in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 50(1),

McDaniel, J., Yoder, P., & Watson, L. R. (2017). A path model of expressive vocabulary skills in initially preverbal preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 47,

McGillion, M., Herbert, J. S., Pine, J., Vihman, M., dePaolis, R., Keren-Portony, T., & Matthews, D. (2017). What paves the way to conventional language? The predictive value of babble, pointing, and socioeconomic status. *Child Development*, 88(1), 156–166.

Morgan, L., & Wren, Y. E. (2018). A systematic review of the literature on early vocalizations and babbling patterns in young children. *Communication Disorders Quarterly*, 40(1), 3–14.

Murillo, E., Camacho, L., & Montero, I. (2021). Multimodal communication in children with autism spectrum disorder and different linguistic development. *Journal of Autism and Developmental Disorders*, 51(5),

Nevill, R., Hedley, D., Uljarević, M., Sahin, E., Zadek, J., Butter, E., & Mulick, J. A. (2019). Language profiles in young children with autism spectrum disorder: A community sample using multiple assessment instruments. *Autism*, 23(1), 141–153.

Newman, R. S., Rowe, M. L., & Bernstein R. N. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, 43(5), 1158–1173.

Newman, S. A. (2016). Origination, variation, and conservation of animal body plan development. *Cell Biology and Molecular Medicine Reviews*, 2(3), 130–162.

Northrup, J. B., & Iverson, J. M. (2015). Vocal coordination during early parent-infant interactions predicts language outcome in infant siblings of children with autism spectrum disorder. *Infancy*, 20(5), 523–547.

Nyman, A., Strömbergsson, S., & Lohmander, A. (2021). Canonical babbling ratio–concurrent and predictive evaluation of the 0.15 criterion. *Journal of Communication Disorders*, 94, Article 106164. https://doi.org/10.1016/j.jcomdis.2021.106164

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Erlbaum.

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, 110(16), 6318–6323.

Oller, D. K., Caskey, M., Yoo, H., Bene, E. R., Jhang, Y., Lee, C.-C., ···Vohr, B. (2019). Preterm and full term infant vocalization origin language. *Scientific Reports*, 9(1), Article https://doi.org/10.1038/s41598-019-51352-0

Oller, D. K., Ramsay, G., Bene, E., Long, H. L., & Griebel, U. (2021). Protophones, the precursors to speech, dominate the human infant vocal landscape. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 376(1836), Article 20200255. https://doi.org/10.1098/rstb.2020.0255

Overby, M., Belardi, K., & Schreiber, J. (2020). A retrospective video analysis of canonical babbling and volubility in infants later diagnosed with childhood apraxia of speech. *Clinical Linguistics & Phonetics*, 34(7), 634–

Persson, A., Marklund, U., Lohmander, A., & Flynn, T. (2022). Expressive vocabulary development in children with moderate hearing loss–the impact of auditory variables and early consonant production. *Clinical Linguistics & Phonetics*, 36(6), 547–564.

Perszyk, D. R., & Waxman, S. R. (2018). Linking language and cognition in infancy. *Annual Review of Psychology*, 69, 231–250.

Pretzer, G. M., Lopez, L. D., Walle, E. A., & Warlaumont, A. S. (2019). Infant-adult vocal interaction dynamics depend on infant vocal type, child-directedness of adult speech, and timeframe. *Infant Behavior and Development*, 57, Article 101325. https://doi.org/10.1016/j.infbeh.2019.04.007

Roche, L., Zhang, D., Bartl-Pokorny, K. D., Pokorny, F. B., Schuller, B. W., Esposito, G., ···Marschik, P. B. (2018). Early vocal development in autism spectrum disorder, Rett syndrome, and fragile X syndrome: Insights from studies using retrospective video analysis. *Advances in Neurodevelopmental Disorders*, 2(1), 49–61.

Rovee-Collier, C. (1995). Time windows in cognitive development. *Developmental Psychology*, 31(2), 147–169.

Rubin, M. L. (2021). Delayed babbling at 10 months: Observation, detection and a two-year follow-up (Unpublished doctorial dissertation). Karolinska Institutet, Sweden.

Saul, J., & Norbury, C. (2020). Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment? *Autism*, 24(5), 1217–1231.

Schoen Simmons, E. (2021). Vocalization. In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders* (pp. 5150–5151). Springer International Publishing.

Schwab, J. F., & Lew-Williams, C. (2016). Repetition across successive sentences facilitates young children' s word learning. *Developmental Psychology*, 52(6), 879–886.

Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research*, 6(6), 468–478.

Talbott, M. R. (2014). Autism risk status and maternal behavior: Impacts on infant language and communication development from 6 to 36 months of age (Unpublished doctorial dissertation). Boston University.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, 23(2), 121–126.

Thurm, A., Manwaring, S. S., Swineford, L., & Farmer, C. (2015). Longitudinal study of symptom severity and language in minimally verbal children with autism. *Journal of Child Psychology and Psychiatry*, 56(1),

Tomasello, M. (2008). *Origins of human communication.* Cambridge, MA: MIT Press.

van der Feest, S. V. H., Yi, H., & Davis, B. L. (2020, November). From babbling to first words: Phonological or lexical selection? Paper presented at the proceedings of the 45th Boston University Conference on Language Development, Boston, US.

Van Egeren, L. A., Barratt, M. S., & Roach, M. A. (2001). Mother-infant responsiveness: Timing, mutual regulation, and interactional context. *Developmental Psychology*, 37(5), 684–697.

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A re-assessment of the continuity issue. *Language*, 61(2), 397–445.

Vilain, A., Dole, M., Loevenbruck, H., Pascalis, O., & Schwartz, J. L. (2019). The role of production abilities in the perception of consonant category infants. *Developmental Science*, 22(6), Article e12830. https://doi.org/10.1111/desc.12830

Wang, Y., Williams, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of LENA™ automated measures for child language development. *Developmental Review*, 57, Article 100921. https://doi.org/10.1016/j.dr.2020.100921

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, 25(7), 1314–1324.

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic predictors of language outcome at 3 years of age. *Journal of Speech, Language, and Hearing Research*, 49(6), 1224–1237.

Weismer, S. E., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders*, 40, 1259–1273.

Werwach, A., Mürbe, D., Schaadt, G., & Männel, C. (2021). Infants' vocalizations at 6 months predict their productive vocabulary at one year. *Infant Behavior and Development*, 64, Article 101588. https://doi.org/10.1016/j.infbeh.2021.101588

Wetherby, A. M., Watt, N., Morgan, L., & Shumway, S. (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *Journal of Autism and Developmental Disorders*, 37(5), 960–975.

Woynaroski, T., Oller, D. K., Keceli-Kaysili, B., Xu, D., Richards, J. A., Gilkerson, J., ···Yoder, P. (2017). The stability and validity of automated vocal analysis in preverbal preschoolers with autism spectrum disorder. *Autism Research*, 10(3), 508–519.

Woynaroski, T., Watson, L., Gardner, E., Newsom, C. R., Keceli-Kaysili, B., & Yoder, P. J. (2016). Early predictors of growth in diversity of key consonants used in communication in initially preverbal children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 46, 1013–1024.

Wu, Z., & Gros-Louis, J. (2017). The value of vocalizing: 10-month-olds' vocal usage relates to language outcomes at 15 months. *Infancy*, 22(1), 23–41.

Yankowitz, L. D., Petrulla, V., Plate, S., Tunc, B., Guthrie, W., Meera, S. S., ··· Pruett, J. R. (2022). Infants later diagnosed with autism have lower canonical

babbling ratios in the first year of life. *Molecular Autism*, 13(28), Advance online publication. https://doi.org/10.1186/s13229-022-00503-8

Yoder, P., Watson, L. R., & Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 45, 1254–1270.

Yoo, H., Bowman, D. A., & Oller, D. K. (2018). The origin of protoconversation: An examination of caregiver responses to cry and speech-like vocalizations. *Frontiers in Psychology*, 9, Article 1510. https://doi.org/10.3389/fpsyg.2018.01510

The predicting effect of speech-like vocalizations on language development in young children and its explanations

LIU Min1, 2, LIU Qiaoyun1,2, CHEN Siqi2, XU Zhijia2

(1East China Normal University Maternity and Infant Health Hospital, Shanghai 200062, China) (2Department of Rehabilitation Sciences, Faculty of Education, East China Normal University, Shanghai 200062, China)

Abstract: Speech-like vocalizations are sounds that resemble adult speech and are the precursor for subsequent language development. Studies have shown that the frequency of speech-like vocalizations, canonical syllable ratio, consonant diversity, and communicative vocalizations in young children predict expressive language, but factors that predict receptive language development remain unclear. Additionally, findings investigating whether babble onset predicts word onset are also mixed. To a certain extent, the predictive relationship between speech-like vocalizations and language development can be explained by three mechanisms: speech-like vocalizations provide the basis for language production, speech-like vocalizations create an optimal learning state for language learning, and speech-like vocalizations promote social behavior. Future research may consider exploring the causal relationship between speech-like vocalizations and language development, moderating effects of related factors on speech-like vocalizations in predicting language development, uniquely valuable speech-like vocalizations for children with language disorders, and the dynamic interaction between speech-like vocalizations and social responses.

Key words: infants, toddlers, speech-like vocalizations, prediction, explanation, language development, language disorder

Note: Figure translations are in progress. See original paper for figures.