

Supplemental Material S1. The Vocal Development Landmarks Interview (VDLI): Clinical Version.

Description of the VDLI Items

The VDLI contains a warm-up section and 18 items querying parents about target behaviors expected to develop between 6–21 months of age. Items are organized into three subscales: Precanonical (7 items), Canonical (5 items), and Word (6 items).¹ To set the context for the range and types of items included in the VDLI and the respective subscales, a summary of expected vocal development landmarks and related citations is provided in Table S1.

The VDLI includes items beginning with the stage called *expansion and increasing volitional control* (see Table S1), which occurs in typically developing infants between 3–8 months of age. These items are organized into the Precanonical subscale, which is inclusive of infant vocalizations that precede true consonant–vowel (CV) syllable productions. The Precanonical subscale comprises seven behaviors related to pitch variation, vowel-like sounds (single, mixed, and reduplicated with glides), marginal syllables, vocal imitations, and range of vowels produced. As shown in Table S1, canonical syllable forms are typically emerging between 5–10 months of age. The advanced stage of vocal development (9–18 months) comprises a period of increasingly complex babble (reduplicated, variegated, jargon forms), expansion of the consonant inventory, and emergence of first words. The Canonical subscale includes five items targeting canonical syllables (single, reduplicated, and variegated), jargon, and range of consonants produced. Later accomplishments from the advanced stage of vocal development are included in the Word subscale. The Word subscale has six items examining word imitations, syllable closure in words, range of consonants and vowels produced, true word production, and two-word combinations.

Table S1. Summary of major vocal milestones associated with age in infants and toddlers.

Age range	Summary of vocal or verbal characteristics	Selected references
0–2 months	Vegetative (burps, coughs, sneezes); reflexive (cries, laughs, squeals, grunts); speech-like productions are rare; may produce primitive “quasivowels”	Ertmer & Iyer, 2010; Oller, 2000
1–4 months	Emerging but minimal control of vocal tract; shift from neutral, open-mouthed vocalizations to increasingly controlled and varied vocalizations (loudness, pitch); vowel-like (vocants); grunts (quasi-resonant nuclei); consonant-like sounds (raspberries, clicks, glottal stops)	Nathani et al., 2006; Oller, 2000; Stoel-Gammon, 2011
3–8 months	Stage of expansion with increasing volitional productions; high-pitched squeals, growls, ingressesives (produced on inspiration); vocal turn-taking; adult-like vowels; marginal babbles (primitive attempts at consonant–vowel combination that are not timed like adult syllables)	Ertmer & Iyer, 2010; Oller, 2000

¹The VDLI also contains a section with four items designed to assess vocal quality, screening for atypical high pitch, growls with harsh quality, glottal stops, and ingressesives vocalizations. The vocal quality section is not included in this paper, as validation work on this aspect of the scale is not complete.

Age range	Summary of vocal or verbal characteristics	Selected references
5–10 months	Coordination of neurophysiological–motor, tactile, perceptual, and auditory systems; canonical syllables emerge (i.e., rapid, adult-like transitions between the consonant and vowel); reduplicated [bababa] and variegated [badaba] strings of babble emerge; consonant inventories expand	Ertmer & Iyer, 2010; Oller, 2000; Oller et al., 1999; Stoel-Gammon & Otomo, 1986; von Hapsburg & Davis, 2006
9–18 months	Advanced stage of vocal development: these advanced vocalizations precede and intermingle with true words; increasing consonant inventory and production of complex syllable shapes (CVC, VC, CCVC); stable productions and consistent set of CV productions setting stage for word forms; babbling (e.g., reduplicated, variegated, jargon) may predominate until 16–17 months; first word productions emerge (10–13 months); word combinations occur when the lexicon size is about 50 words	Bates et al., 1995; Ertmer & Iyer, 2010; McCune & Vihman, 2001; Nathani et al., 2006; Oller, 2000; Robb et al., 1994; Stoel-Gammon, 1998

Note. CV = consonant–vowel; CVC = consonant–vowel–consonant; VC = vowel–consonant; CCVC = consonant–consonant–vowel–consonant.

VDLI items included in the Precanonical, Canonical, and Word subscales are summarized in Table S2. These specific items were selected from multiple sources including (1) the extant literature on infant vocal development, (2) existing vocal development hierarchies, and tools, and (3) existing parent-report scales targeting vocal development reviewed above. Many items overlap with those included in a research tool, the Stark Assessment of Early Vocal Development–Revised (SAEVD-R; Nathani, Ertmer, & Stark, 2006). An item targeting bilabial trills was included in the Precanonical section based on advisement from Dr. Carol Stoel-Gammon that such productions provide practice that precedes consonant production. Other items, such as imitation of vocalizations and words, consonant inventories, and syllable closure in word production, were selected based on their inclusion in other studies measuring vocal development (Lyytinen, Poikkeus, Leiwo, Ahonen, & Lyytinen, 1996; Proctor, 1989; Watt, Wetherby, & Shumway, 2006).

Development of the current version was based on our earlier research with the Vocal Development Landmarks Interview–Experimental Version (VDLI-E; Ambrose, Thomas, & Moeller, 2016) and input on the overall approach of the scale itself from researchers involved in the study of infant vocal development (Drs. Kim Oller, David Ertmer, and Carol Stoel-Gammon). This resulted in adding the warm-up section, replacing a few redundant items with additional items at the word level to avoid ceiling effects on that subscale, and making major modifications to the examiner manual and scoring procedures.

Table S2. Items (and item numbers) included in the Precanonical, Canonical, and Word production subscales, along with descriptions of the target for each item, response type, and question formats.

Subscale	Vocal behavior targeted	Description of target	Response	Format
Precanonical	Pitch variation (PC 2-1) ^a	Pitch contours that rise, fall, or are varied (vs. stable, nonvaried pitch)	Frequency	Paired comparison
	Mixed vowels (PC 2-2)	Two or more well-formed vowels in a row or glide + vowel (vs. single vowel)	Frequency	Paired comparison
	Bilabial trills or fricatives (PC 2-3)	Raspberries: sounds produced by vibration of lips or lips and tongue (raspberries)	Frequency	Multiple exemplars
	Marginal syllables (PC 2-4)	Approximates syllable, but timing is slow and syllable not well formed (vs. vowel or glide + vowel)	Frequency	Paired comparison
	Reduplicated glides + vowel (PC 2-5)	Repetitive string of /j, w, h/ + vowel(s) (vs. vowels only)	Frequency	Paired comparison
	Imitations of vocalizations (PC 2-6)	Clear attempts to imitate vocal behaviors/words	Frequency	Multiple exemplars
	Variety of vowels produced (PC 2-7)	Count of vowels produced in isolation or with glides	Inventory	Multiple exemplars
Canonical	Single canonical syllables (CB 3-1)	Clear well-formed syllables such as /ba, gi, uk/ (vs. marginal syllable)	Frequency	Paired comparison
	Reduplicated canonical sequences (CB 3-2)	Repetitive strings of well-formed syllables such as /bababa/ (vs. single canonical form)	Frequency	Paired comparison
	Variegated canonical sequences (CB 3-3)	Repetitive strings with varied vowels/consonants (vs. reduplicated canonical sequences)	Frequency	Paired comparison
	Jargon-like utterances (CB 3-4)	Strings of variegated syllables with changing intonation (vs. variegated canonical sequences)	Frequency	Paired comparison

Subscale	Vocal behavior targeted	Description of target	Response	Format
	Variety of consonants produced (CB 3-5)	Count of consonants in babble	Inventory	Open ended
Word	Imitation single words (WP 4-1)	Child imitates adult words and accuracy is judged	Accuracy	Comparison scale
	Closed syllables (WP 4-2)	Final consonant is present on words (vs. open syllable at end of word)	Correctness	Paired comparison
	Variety of consonants produced in words (WP 4-3a)	Count of consonants produced consistently in words	Inventory	Open ended
	Variety of vowels produced in words (WP 4-3b)	Count of vowels produced consistently in words	Inventory	Open ended
	Variety of words used (WP 4-4)	Number of spontaneous words produced	Inventory	Multiple exemplars
	Variety of word combinations used (WP 4-5)	Number of spontaneous true word combinations (vs. frozen forms like <i>all-done</i> or <i>bye-bye</i>)	Inventory	Paired comparison

Note. Parenthetical comments include the contrastive pairs that are presented when items use a paired comparison format.

Vocal Development Landmarks Interview Methods

The VDLI is a parent interview that is conducted with the use of a series of digital slides, many of which contain audio files of authentic infant vocalizations. The audio files are presented to model target behaviors, sometimes in contrasting pairs, as a way of ensuring that parents understand the target behaviors in question.

Most of the audio vocal samples used in the VDLI were taken from recordings collected during previously reported longitudinal studies (Bass-Ringdahl, 2010; Moeller et al., 2007). During collection of the original recordings, infants were fitted with vest that held a wireless lavalier microphone (Shure Model LX1-V). The microphone was positioned on the chest to maintain a consistent microphone-to-mouth distance of approximately 2 inches. Audio and video recordings were collected with a Panasonic Professional AGDVC10P Mini-DV camcorder or a JVC SR-VS30 Mini-DV/S-VHS video record deck. A microphone mixer (Shure Model LX-4) controlled the audio levels routed to the camera/deck (Moeller et al., 2007). Two consultants in infant vocal development (Drs. Carol Stoel-Gammon and David Ertmer) reviewed the audio files and provided feedback. Minor adjustments were made based on their feedback, including

purposely recording of a few new exemplars using local typically developing toddlers. These children wore a digital recorder positioned on the chest at a consistent microphone-to-mouth distance. Audio files were extracted and prior to their inclusion in the VDLI, all audio levels were equalized using Final Cut Pro software.

The VDLI was designed to be inclusive of the developmental age range of 6 to 21 months. Prior to beginning the interview, a brief explanation of its purpose is given. Parents are informed that (1) the goal is to gain an understanding of their daily observations of their child's sound-making and speech behaviors, (2) they will be listening to examples of infants making various sounds and talking, (3) their task is to think about the sounds their child produces on a typical day and consider whether their child's productions are similar to the ones played, (4) they will hear audio files from real children, and that children develop speech at varying rates, (5) they will see child photos on the slides that are intended to add interest but do not suggest that children of any certain age produced these vocalizations, and (6) they are allowed to listen to the audio examples as often as desired.

Warm-up section. Parents are presented with open-ended warm-up questions before they are presented with items that contribute to VDLI scoring. The purpose of the warm-up section is to orient parents to the procedures and to gain a broad estimate of the child's current vocal developmental stage. This section was added following advisement from Dr. Kim Oller, who indicated that open-ended questions guide examiners in focusing their inquiries within a developmentally appropriate range. The first open-ended question was, "What are some examples of the sounds or words your child is currently saying?" If the parent reported the production of words, the examiner followed-up by asking the parent to provide an estimate of the number of words their child was currently producing (i.e., "Can you give me an estimate of how many words your child is able to say on his/her own, without copying what you say?"). Based on the parental response, the examiner makes an early determination of whether the child is likely to be at the precanonical, canonical, or word stage. The examiner is given the flexibility to ask follow-up questions if the global stage is unclear from the parental response. The warm-up discussion—and, particularly, the estimate of words produced—leads to decisions (similar to basal rules) about the starting section for VDLI questioning. Decision rules that guide the administration starting point are illustrated in Figure S1. This process avoids asking parents of advanced children who are producing words to report on vocalizations that no longer provide relevant information about their child's developmental phonological level or asking parents of children at the precanonical level about word production, which may be several months off developmentally. Once the basal starting level is determined, the interviewer proceeds with administering the appropriate VDLI subscales.

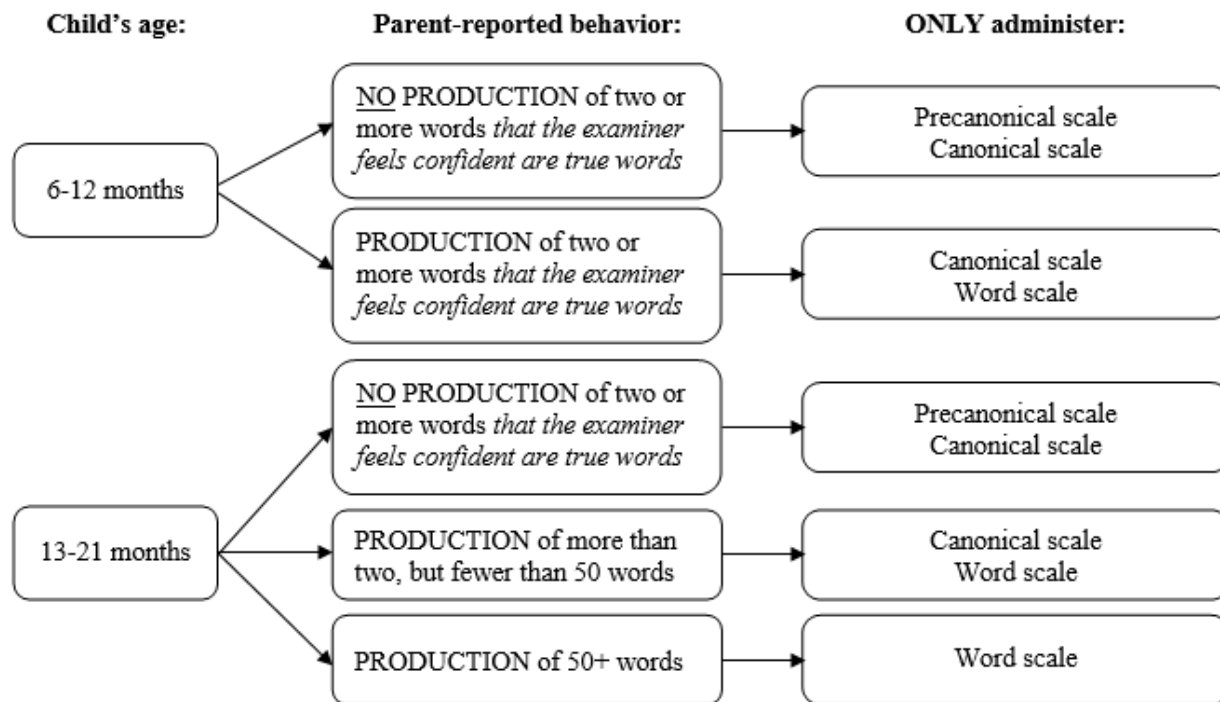
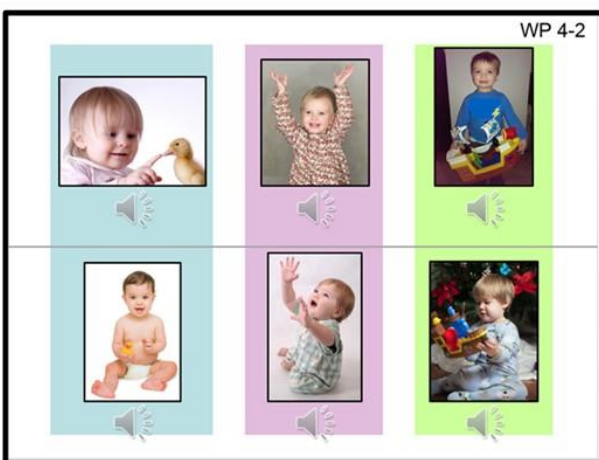


Figure S1. Rules for Vocal Development Landmarks Interview subscale administration.

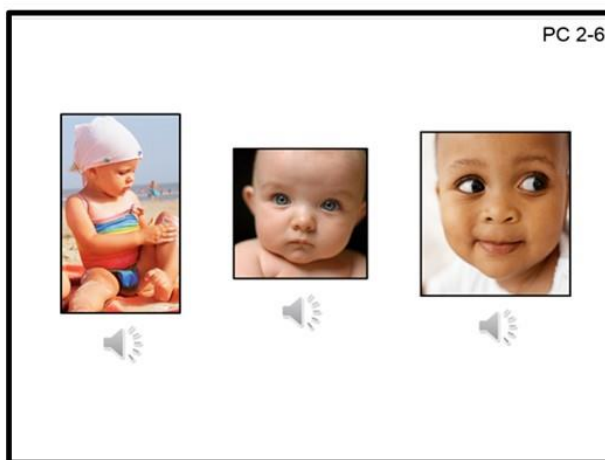
VDLI question formats. The VDLI incorporates four question formats: (1) paired-comparison, (2) multiple exemplars, (3) comparison scale, and (4) open-ended. The first three formats include use of audio files to guide understanding of the target, and the final format uses static images accompanied by open-ended questions. Figure S2 provides illustrations of items that represent each format. The paired-comparison format was designed to juxtapose target vocal behaviors that represent nuanced developmental steps that could be conflated by parents but are important to distinguish developmentally (e.g., glide + vowel vs. true canonical syllable). As shown in Figure S2, item WP 4-2 explores whether the child is producing words with closed syllables. The three audio examples presented in the top row include words produced with final consonants, whereas the three on the bottom row are the same words produced with open syllables (e.g., *boat* vs. *bo*). Parents are asked to consider if their child typically produces words that are more like the children in the top row (closed) or the bottom row (open) or both about equally. Once the selection is made, parents are then asked to rate the frequency of closed syllable productions on a 4-point Likert scale from never to frequently, using a visual scale provided on the following slide. For items where the target is fairly straightforward to model, three to five exemplars are played (see Figure S2, PC 2-6). A minimum of three exemplars was adopted to reduce the chance that parents think they are being asked if their child produces the specific vocal behavior presented in the audio file. For example, when asking about the child's ability to produce a variety of vowels, multiple audio files presenting different vowels prompt the parents to generalize to the category "vowels," rather than a specific vowel. The use of audio files also ensures that parents understand what is meant by "vowel."

A comparison scale is used on a single item to query parents about the closeness of the child's word imitations to the adult model. As shown in Figure S2, item WP 4-1, the audio examples and visual scale provide an opportunity for the parent to compare children imitating

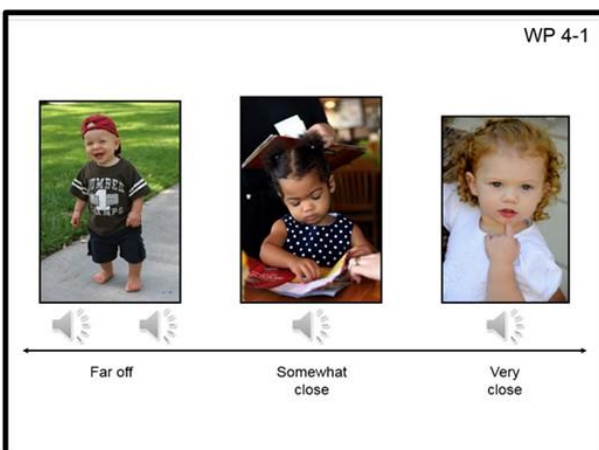
adult models. In the examples, each child is either far off, somewhat close, or very close to the models presented in the audio clips. The parent uses the scale to judge the accuracy of their child's word imitation attempts. The final question format is open ended (see Figure S2, CB 3-5). Static pictures are provided for interest only. At this point in the interview, parents have heard multiple examples of consonants and vowels, and they are asked to comment on the specific consonants and vowels they regularly hear in their child's productions (consonants in babble, consonants and vowels in words).



Paired comparison: This item compares words produced with closed syllables (duck, up, boat) in the top row to those with open syllable productions (duh, uh, bo) in the bottom row.



Multiple exemplars: This item models three examples of infants imitating a parent vocalization.



Comparison scale: This item presents examples of word imitations on a visual scale ranging from far off, somewhat close, and very close to the model.



Open-ended questions: In this item, static photos are presented on the slide and serve as a placeholder for open-ended discussion.

Figure S2. Illustrations and examples of the four Vocal Development Landmarks Interview question format types.

VDLI Scoring

Items in the Precanonical, Canonical, and Word subscales are scored according to one of four methods, as described in Table S3. For frequency-based responses, point scoring is aligned with the ratings on a 4-point Likert scale: *never* (0 points), *rarely* (1 point), *sometimes* (2 points), and *frequently* (3 points). Imitation accuracy scoring is similar (0–3 points), with higher scores aligned with close matches to the model and lower scores reflecting poorer or no imitations. For reports related to consonant, vowel, or word inventories, scores also range from 0–3, with scoring depending on the number of behavior types reported, as summarized in Table S3. For example, for item PC 2-7, production of at least four different vowels at the precanonical level is required to receive the highest score (3), with fewer or no true vowels scored in the 0–2 range. Total scores for each subscale are summed, divided by the total number of possible points (21 for precanonical, 15 for canonical, and 18 for word), and multiplied by 100, yielding a percent score for each subscale. Children whose parents reported in the warm-up that their children were producing two or more words independently were deemed beyond the precanonical stage and, thus, received a score of 100% on that subscale. Likewise, parents who reported that their children were producing 50 or more words independently were deemed beyond the canonical stage and received a score of 100% on that subscale. These ceiling-like rules prevented children from having artificially low scores as a result of replacing developmentally early vocalizations types with more advanced vocalizations. If a child was reported to use fewer than three words, the word scale was not administered due to ceiling rules and the child received a 0% score for that subscale. The total VDLI score was calculated by averaging the percent scores for each of the three subscales. Additional resources related to the VDLI (interviewer and scoring forms) are provided in supplemental online materials for readers interested in greater detail about the VDLI.

Table S3. Vocal Development Landmarks Interview response types and scoring schemes.

Response type	Description	Scoring approach
Frequency rating	Judge how often the child produces the behavior using a 4-point Likert scale	Never = [0] Rarely = [1] Sometimes = [2] Frequently = [3]
Inventory report	Report number of different vowels, consonants, or words the child produces	Varies by item: Vowels (PC 2-7): none = [0]; 1–2 = [1]; 3 = [2]; 4+ = [3] Consonants (CB 3-5): none = [0]; 1–3 = [1]; 4–6 = [2]; 7+ = [3] Consonants (WP 4-3a): 1-2 = [0], 3–5 = [1]; 6–8 = [2]; 9+ = [3] Vowels (WP 4-3b): 1 = [0]; 2–3 = [1]; 4–5 = [2]; 6–7 = [3] Words (WP 4-4): none = [0]; 1–4 = [1]; 5–19 = [2]; 20+ = [3] Word Combination (WP 4-5): none = [0], 1–3 = [1]; 4–9 = [2]; 10+ = [3]
Accuracy rating	Judge accuracy of imitations	No imitation = [0] Far off = [1] Somewhat close = [2] Very close = [3]

Note. PC = Precanonical; CB = Canonical; WP = Word production. Bracketed values signify the point score assigned. For inventory report, nonbracketed values represent the number of reported behaviors.

For readers interested in accessing a beta version of the VDLI and its associated administration manual, contact Dr. Sophie Ambrose at sophie.ambrose@boystown.org.

References

- Ambrose, S. E., Thomas, A., & Moeller, M. P. (2016). Assessing vocal development in infants and toddlers who are hard of hearing: A parent-report tool. *Journal of Deaf Studies and Deaf Education*, 21(3), 237–248. <http://doi.org/10.1093/deafed/enw027>
- Bass-Ringdahl, S. M. (2010). The relationship of audibility and the development of canonical babbling in young children with hearing impairment. *Journal of Deaf Studies and Deaf Education*, 15(3), 287–310. <http://doi.org/10.1093/deafed/enq013>
- Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 96–151). Oxford, United Kingdom: Basil Blackwell.
- Ertmer, D. J., & Iyer, S. N. (2010). 24 prelinguistic vocalizations in infants and toddlers with hearing loss: Identifying and stimulating auditory-guided speech development. In M. Marschark & P. E. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (pp. 360–375). New York, NY: Oxford University Press.
- Lyytinen, P., Poikkeus, A., Leiwo, M., Ahonen, T., & Lyytinen, H. (1996). Parents as informants of their child's vocal and early language development. *Early Child Development and Care*, 126(1), 15–25. <http://doi.org/10.1080/0300443961260102>
- 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. [https://doi.org/10.1044/1092-4388\(2001/054\)](https://doi.org/10.1044/1092-4388(2001/054))
- Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., ... Stelmachowicz, P. (2007). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I - phonetic development. *Ear and Hearing*, 28(5), 605–627. <http://doi.org/10.1097/AUD.0b013e31812564ab>
- Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics & Phonetics*, 20(5), 351–369. <https://doi.org/10.1080/02699200500211451>
- Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, 32(4), 223–246.
- Proctor, A. (1989). Stages of noncry vocal development in infancy: A protocol for assessment. *Topics in Language Disorders*, 10(1), 26–42.
- Robb, M. P., Bauer, H. R., & Tyler, A. A. (1994). A quantitative analysis of the single-word stage. *First Language*, 14(4), 037–048.
- Stoel-Gammon, C. (1998). Sounds and words in early language acquisition: The relationship between lexical and phonological development. In R. Paul (Ed.), *Exploring the speech-language connection* (8th ed., pp. 25–52). Baltimore, MD: Paul H Brookes Publishing Company.
- Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, 38(1), 1–34. <http://doi.org/10.1017/S0305000910000425>
- Stoel-Gammon, C., & Otomo, K. (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders*, 51(1), 33–41.
- von Hapsburg, D., & Davis, B. L. (2006). Auditory sensitivity and the prelinguistic vocalizations of early-amplified infants. *Journal of Speech, Language, and Hearing Research*, 49(4), 809–822. [http://doi.org/10.1044/1092-4388\(2006/057\)](http://doi.org/10.1044/1092-4388(2006/057))
- 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. [http://doi.org/10.1044/1092-4388\(2006/088\)](http://doi.org/10.1044/1092-4388(2006/088))