Advancing Toward What Will Be: Speech Development in Infancy and Early Childhood

Jennell C. Vick

Address:
11635 Euclid Avenue
Cleveland, Ohio 44106
USA

Email:
jvick@chsc.org

Acquiring speech, the most finely coordinated of human skills, is a momentous accomplishment of early development.

How It's Done

"You listen. And then you make noises. And then you shape the noises into words." This, an account of speech development, comes from my seven-year-old daughter, Fiona, who then added, "You just have to keep trying." Neither Fiona nor her five-year-old sister, Annabel, recall a time before they could speak, and both were equally perplexed by my question about how people learn to speak. But, amazingly, the explanation Fiona provided was fairly accurate and easy to understand. With just her short lifetime to practice, she articulates with the near precision and intelligibility of an adult, coordinating the movements of some 70 muscles and 10 different body parts to shape her breath and voice into her carefully articulated message (de Jong, 2018).

The process of acquiring speech is, indeed, a momentous achievement, with a time course that begins long before birth and continues well into the school-age years. Decades of research have helped elucidate the many systems and processes that must be in place for speech to emerge as well as some of the obstacles or constraints that limit the speed of development. From birth, it seems the scaffolding has already been put in place for mature language despite the seemingly limited abilities and functions of the neonate. At every turn on the journey, it seems there is a singular focus on advancing toward what will be, with occasional staggering leaps forward in development that bear little resemblance to the skills or abilities from just a day or week prior. At the earliest stages, parents play a critical role. Together, the baby and the parent exchange and learn together to advance toward the common goal of sharing a conversation. The earliest words represent the attainment of countless other skills and capacities, including the ability to recognize that things have names (in English, most first words are nouns), to perceive the sounds in the environment as meaningful, to participate in the basic turn taking of verbal social exchanges, and, finally, to coordinate the movements of a series of structures in an intentional way to produce something like "mama," "ba," or "da."

Speech and language seem like inseparable concepts, but an account of speech acquisition requires an understanding of how they diverge. Language, the words we use to express ideas and the associated rules for putting them together, is a catalyst for speech, which is just one of a multitude of ways that language can be expressed. Speech, on the other hand, is how we verbalize the sounds, words, and ideas and, more specifically, how the sounds of a language are generated with the lips and tongue. While language and speech develop concurrently, relying on each other to progress, it is noticeable during development how they sometimes diverge. Many parents notice, for instance, that a toddler can understand much more than it can express.

©2018 Acoustical Society of America. All rights reserved.

Do You Hear What I Hear?

The first words spoken represent the interplay of speech and language. The infant's skill as a linguist long predates the emergence of these first communicative utterances. Fiona's remark that learning to speak somehow requires the baby to "listen" is accurate. Because hearing is the first of the senses to emerge during fetal development, the listening begins about halfway through gestation (Lecanuet and Schaal, 1996). Somewhere around 30 weeks, the fetus begins reacting to voices and sounds in the environment (Shahidulla and Hepper, 1994). The ability to discriminate speech sounds also begins in utero, with neonates preferring sounds and voices to which they were exposed before birth (DeCasper and Fifer, 1980). While the sounds in the outside world are filtered through the noise of the womb, unlike the depiction in **Figure 1**, fetuses are still able to learn to discriminate vowel sounds. In one study, mothers regularly played audio samples that included many repetitions of "tahtahtah" and "tahtohtah," starting at 28-weeks gestation. Shortly after birth, these infants were able to discriminate these two samples, whereas infants who did not hear the samples in utero could not (Partanen et al., 2013). In the first months of life, babies are equally skilled at discriminating sounds and patterns from all languages, but later, between 6 and 12 months of age, babies become specialists in the sounds of their own language, likely due to reorganization of how they store and eventually imitate the sounds they hear (e.g., Best and McRoberts, 2003). This early specialization is particularly important, associated with greater spoken language abilities at two years (Kuhl et al., 2005).

When the baby arrives in the world, the full sensory experience of language, speech, and communication extends and enhances the listening the baby did in utero. The familiarity of voices and the rhythms and tones of the language get paired with eye contact, touch, and the nurturing love provided by the primary caregivers. In addition, the baby is ready to play, in a sense, mimicking the actions of adults. As young as a few hours old, babies are able to imitate the actions of adults, including tongue protrusion, mouth opening, and lip rounding (Meltzoff and Moore, 1977). This early skill foreshadows what will be in the later joint attention of play and the dialogue of verbal communication.

This work of listening to and learning the language around them is made easier for babies because of the exaggerated speech used by caregivers to address infants. Infant-directed speech is used by caregivers in many of the world's cultures (Kuhl et al., 1997) and includes simplified vocabulary, slower rate, larger mouth movements (Green et al., 2010), greater pitch variability, and clear articulation (Kuhl et al., 1997). It facilitates language learning in the baby because it increases the baby's attention to the speech of the caregiver in a positive emotional context, as depicted in **Figure 2**, enhancing the social give-and-take that will later become critical for language production by the child (Tamis-LeMonda et al., 2014). In addition, this speaking style calls attention to the contrasts and distinctions of the native language, such as vowel sounds, stress patterns, and the beginnings and endings of words (Golinkoff et al., 2015). The face-to-face nature of early interactions, as depicted in **Figure 2**, plays an important role as well. Perhaps most important, use of infant-directed speech is a critical element of parental responsiveness, a positive predictor of later vocabulary size and on-time achievement of language milestones (Tamis-LeMonda, 2014).

Figure 1. *Listening to speech begins long before birth. Photo from istockphoto.com.*

Taking Control

Fiona had every advantage of early language learning, including a mother who was savvy with responsiveness and infant-directed speech. Despite this, her first true word was not uttered until she was a year and a half old. Individual paths are quite common in the early stages of speech development because different skills develop at different rates. When the development of one element limits progress in the achievement of a skill, it is a constraint. In development, constraints are presumably created by less developed processes or systems that do not mature at the same rate as other requisite processes required for a skill. In Fiona's case, the timely emergence of her speech was likely constrained by having insufficient control and coordination of the systems and structures required for speech, including the speech articulators (the tongue, lips, and jaw). This is not uncommon among children who are otherwise typically developing. Among two year olds, about one in five do not yet combine words as would be expected by this age (Zubrick et al., 2007). The development of speech motor control is the most likely constraint for babies and toddlers as they work toward producing their first words (Moore, 2004).

Speech motor control refers to the integration of inputs from the brain to control the activation of muscles and structures to produce the skilled movements required for speech. The inputs must coordinate exhaled breath, voice, the movements of the articulators, and even the stiffness or laxness of the cheeks; speech requires conscious neural control and coordination that is unparalleled in the repertoire of skilled movements. This is big work for a baby! This advanced control emerges gradually over the course of infancy and childhood, with adult-like control and stability not achieved until adolescence (Smith and Zelaznik, 2004).

Learning this control for speech production occurs gradually through the building on skills that are either innate to the newborn or develop in early infancy (MacNeilage and Davis, 2000). Any parent will tell you that right from the start, a baby can express itself and that it does so loudly and frequently. At birth, an infant is able to produce a cry, but soon gooing and babbling emerge, which are important precursors to early speech development (Oller, 1980). These behaviors follow a predictable pattern, beginning with a stage of nonresonant vowel-like sounds in the first two months. In this, the "phonation" stage, Fiona was able to produce long vowel-like sounds and she seemed to exchange them with me in a dialogue. In this sample (see Multimedia File 1 at acousticstoday.org/mm-vick), when Fiona was two months old, it is clear that she is working to control her voice and seems to be moving her jaw a small amount, which has only a subtle impact on shaping the voice. Following this, the "gooing" stage emerges at two to three months when the baby combines the vocalizations with guttural consonant-like sounds, the most primitive articulation. The gooing stage reflects the infant learning to use the tongue and lips to articulate the voice into something like speech sounds, although very primitive, as can be seen here when Fiona was three months old (see Multimedia File 2 at acousticstoday.org/mm-vick). At four to six months of age, the "expansion"

Figure 2. *Infant-directed speech is common to most of the world's languages and enhances language learning. Photo from istockphoto.com.*

stage emerges and includes clearer vowel sounds, squeals, yells, raspberries, and other vocal play sounds. This stage can be really noisy (see Multimedia File 3 at acousticstoday.org/mm-vick), but it is an important precursor to the clearly formed syllables that are about to emerge.

What we think of as babbling, the combination of consonants and vowels repeated over and over (e.g., "babababa"), tends to emerge between 6 and 10 months of age, right around the time that the baby becomes a specialist in discriminating the sounds of their own language over those of other languages. Later, this babbling becomes "variegated" with varying syllable shapes and sounds (e.g., "badabada"), just prior to the emergence of the first words. These vocalizations herald the emergence of true syllable structure, exciting for both the linguist and speech therapist moms to witness. Understandably, many parents begin shaping these productions into "words" by associating them with objects or people that are important, repeating the syllables and pointing and, ultimately, using the syllables in responsive exchanges with the baby, as seen in this video from when Fiona was six months old (see Multimedia File 4 at acousticstoday.org/mm-vick). In it, she produces beautiful strings of "mamamamama," but I reply by asking her to say "Dada," shaping and expanding Fiona's vocal play.

For the baby, these babbling strings are made in many contexts, most of which are not communicative at all. It seems, instead, that the baby uses these productions to explore and learn the movements and their associated acoustic consequences. The learning that occurs is the creation of linkages between the movements with the sensations: big movements are easy to feel and make loud, hard sounds and small movements and whispers make similar sounds, but they are quiet. In all cases, the baby is learning to "shape the sound," as Fiona put it.

Frank Guenther (e.g., Tourville and Guenther, 2010) proposed a robust neurocomputational model of speech sound learning that explains that this exploration and practice by the baby is an essential step toward achieving mature speech. Called Directions Into Velocities of Articulators (DIVA), the model explains that early listening helps to build auditory targets and that the repetitive, cyclic babbling strings are what allow the baby to create mappings in their own brain about how to achieve these targets. These connections are then used to help the baby learn what combinations of movements and sensations produce which sounds (target maps) and which movements and sensations do not achieve the target sound (error maps). These target and error maps are then used throughout the development of speech and language to build the repertoire of productions and even in adulthood when these skills have fully matured.

The small units, or syllables, produced during babble are eventually layered over the pitch patterns of the native language so that the babbling of a baby in the United States sounds "English," but the babbling of a baby in Japan sounds "Japanese." This reflects the interface between language and speech (Smith, 2006) and the multilevel layering of processes and skills required to achieve mature speech production. For Fiona, the more complex babbling with alternating syllables did not emerge until after her first birthday. In this video from when she was 14 months old (see Multimedia File 5 at acousticstoday.org/mm-vick), you can clearly hear her wild and very loud combinations of many different syllable shapes and structures. Neither her pitch nor the loudness of her voice was modulated in this series and she was using these vocalizations to work on coordination and control for the articulators.

At the same stage, Fiona also produced vocalizations that were communicative but that lacked consonants or the cyclic repetition of babbling. She commonly used single syllables that sounded like words but were composed of a single vowel, something like "uh." With these productions, she carefully modulated the pitch of her voice to communicate questions (rising pitch) and declaratives (rising then falling pitch), and she coordinated these vocalizations with facial and hand gestures. This is illustrated in a sequence captured on video (see Multimedia File 6 at acousticstoday.org/mm-vick) when Fiona was still 14 months old. Fiona seemed to work on speech articulation in parallel (but not simultaneously) with her work on vocal modulation and prosody. She followed her own, highly individualized path.

As a baby progresses through these stages of vocalizations, it is achieving greater mastery of the coordination of the structures that articulate speech, with a presumed focus on learning to independently control these structures. In the earliest months of infancy, only the voice is controlled, but the guttural sounds of the gooing stage add movement of the lower jaw to bring the tongue to the roof of the mouth, which generates consonant-type sounds. Eventually, babbling coordinates movements of the jaw, lips, tongue, and voice, with increasing differentiation of these structures. Finally, in mature speech production, not just the structures but discrete segments of these structures must be controlled independently. Think, for instance, of the way the tongue moves to produce a "t" versus a "k" sound. In the former, the front of the tongue is elevated to shape the voice, whereas in the latter, the back of the tongue is elevated (Edwards, 1992). This independent control, called differentiation, is particularly important for mature speech to be produced accurately (Gibbon, 1999).

Doing More with Less

Only recently has technology allowed us to capture and measure the earliest stages of speech motor control. In one approach, Green and Nip (2010) quantified the movements produced in early speech development using line-of-sight motion capture technology. Tiny reflective markers are placed on the lips, face, and jaw of children (**Figure 3, *left***) and the 3-dimensional motion of the sensors is captured by special cameras and software. The software assembles the tracked sensors into a 3-dimensional model, as seen in **Figure 3, *right***. This approach allows for analysis of the movements that contribute to the earliest vocalizations and speech (e.g., Green et al., 2000). For a syllable like "ba," produced commonly in both babbling and early words, the child must bring its lips together. This can be accomplished with a variety of strategies. For instance, the lips could be held passive and the jaw can bring them together. In adults, the jaw does some of the work, but it is handily assisted by bringing the lower lip to the upper lip with independent movements of these structures. The motion capture techniques provide clarity about the contribution of each structure to lip closure and measure the coordination of the movements.

For instance, the control of the jaw and the lip control required to produce the "ba" sound develops gradually over the early

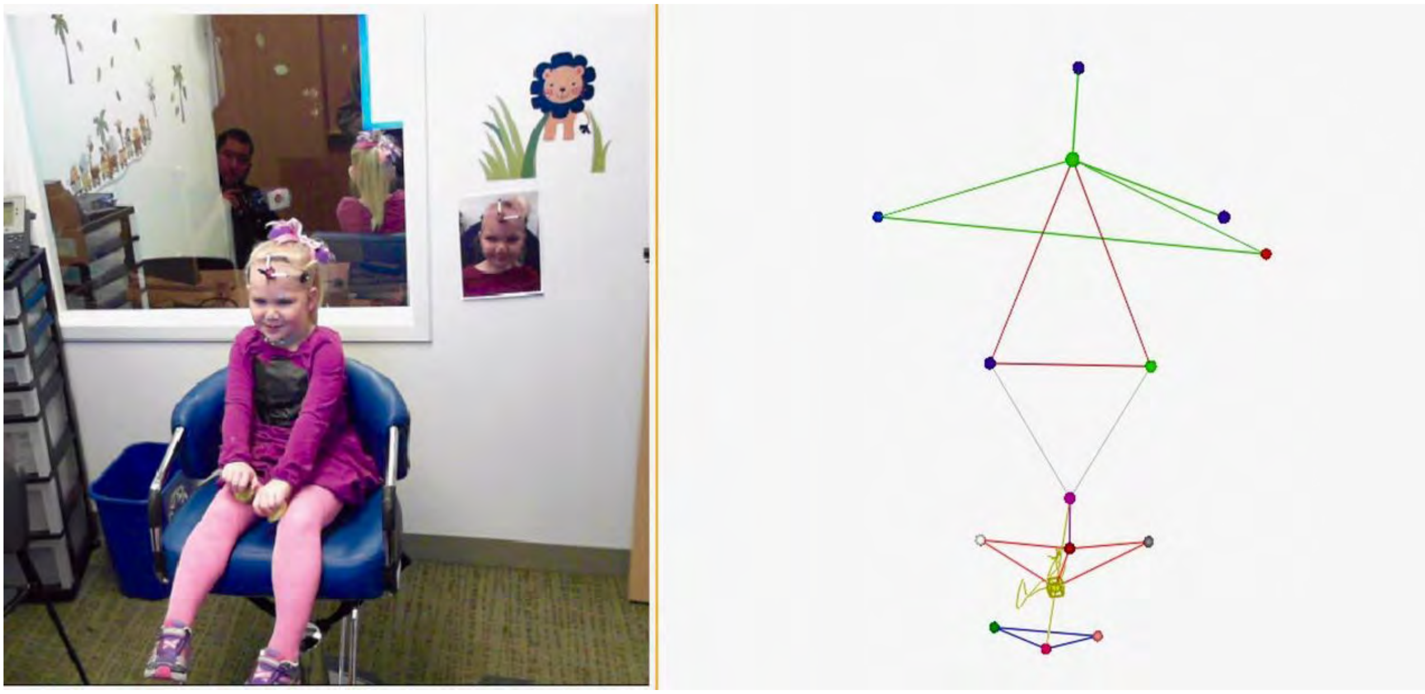

Figure 3. *Noninvasive reflective motion capture markers are applied on Fiona's face, lips, and jaw (**left**). The movement of the markers are tracked in three dimensions, and the resulting model (**right**) can be analyzed in various ways to better understand speech production and development. Photo from Jennell Vick, 2014.*

years of childhood. At 12 months of age, movement of the lower jaw is the primary driver that brings the lips together for the "b" sound. By two years of age, the lower lip and upper lip start to contribute in equal measure to each other, but the jaw is still the primary driver. At six years of age, children use both the lower lip and jaw equally to achieve lip closure for the "b," with less movement from the upper lip, quite similar to the strategy used by adults (Green et al., 2000). In addition to the lack of differentiation in structures in early vocal productions, the movement of the jaw is much more ballistic, lacking the fine control observed in older children (Green et al., 2002). It is clear that the earliest vocalizations are accomplished by simplifying the more complex movements produced by adults to generate sounds that approximate the adult versions. The infant and toddler are, then, reducing the degrees of freedom of the system using the limited control that they have mastered (Gick et al., 2008).

All of this control develops uniquely in children because of the highly individualized constraints within each child. For example, tongue control and coordination develops later than jaw control and many children seem to babble predominantly with jaw-driven segments like "ba" and "ma," but other children babble more complex sounds like "k" and "g," as reported by Stoel-Gammon and Cooper (1984). Regardless, all babies face the same challenge of attempting to mimic the sounds they hear from adults with much different equipment. That is, the size and relative proportion of the speech-producing structures are vastly different in infants and young children compared with adults. Moreover, the structures are growing rapidly and nonuniformly throughout the period of speech development (Vorperian et al., 2005). Acoustically, the output from the growing speech production system is continually changing during this period so that the same configuration at one stage will produce a different sound than at subsequent stages (Menard et al., 2004). This challenge creates a tension between what the child is able to produce with limited control, which is far from perfect, and the desire to produce target sounds like the adult model (McAllister Byun et al., 2016).

As the child builds maps within the brain that associate its own ability to produce a sound that is something like "ba" with the adult word "ball," it creates a foundation on which its early words will build. These earliest syllable shapes, produced during the babbling stage, facilitate the eventual production of words that contain these very sounds (Vihman, 2013). The motoric and anatomic constraints of the baby's system are the most likely determinants of these early productions and preferences, not the speech sounds the baby hears most frequently from its caregiver (Majorano et al., 2014). The baby's own babbling provides the most frequent input to the auditory target map-building system in the brain (Tourville and Guenther, 2010).

The First Word

Considering, now, the skills, limitations, and babbling habits of a baby at 12 months, we come to the moment of the first word. The problem, of course, is that there is no "moment" but instead a period where words begin to emerge that can easily be confused with productions that are not real words. Between 9 and 16 months of age, fewer than 10% of vocalizations are "true words," based on criteria including whether the production is used multiple times, whether the mother identifies it as a word, if there is similarity to the adult form, and not using the vocalization in other contexts (Vihman and McCune, 1994). For instance, "ba" could be considered a true word if used regularly to represent a single concept (e.g., to refer to bottles but not other things), if mom recognizes it as the word the child uses to refer to the concept, and if it is produced similarly from instance to instance. A candidate first word for Fiona early on would have been "mama," which she seemed to use to refer to me, but she also vocalized "mama" in numerous other contexts and as a commonly used babbling string. With these stringent criteria, Fiona did not produce any true words until around 18 months. She soon started combining words, which allowed her to catch up with the verbal achievements of typical peers.

As those words and phrases started to come together, Fiona began expressing her ideas. At 18 to 24 months of age, the typical 2-word combination will be in the form of agent + action. Given the egocentricity of toddlers at this stage, many of these combinations refer to the toddler itself, and it is common for children at this age to refer to themselves in the third person. For Fiona, then, the typical phrase might be "Fiona jump." Recall, though, that the toddler must work with a still-maturing speech motor control system, balancing the desire to be understood with the inability to truly mimic the adult forms of words. Early words, limited by the motor constraints of development, tend to exclude sounds, forms, or complexities that the immature speech motor system has yet to acquire. In the face of a burgeoning vocabulary and desperate desire to express herself, the young talker will, instead, modify the adult form of a production to fit the "templates" that it has already mastered (e.g., Majorano et al., 2014).

Fiona's name requires an incredible amount of control and coordination to produce. For example, because of the lack of independent control for the lower and upper lip at two years of age, the adult version of the "f" sound was not possible. To achieve this sound, the lower lip is curled back sufficiently for the upper teeth to approximate it while an airstream is directed through the small opening. Furthermore, most children have difficulty combining multiple gestures into a single sequence, as in the two vowels in the middle of Fiona's name. Fiona's typical productions at this stage went consonant-vowel-consonant-vowel (CVCV), like "mama." The pattern to produce "Fiona," CVVCV, is much more complex.

Faced with the challenge of her own name in these early two-word combinations, Fiona substituted the word "Sahmi," which capitalized on sounds that she was already able to produce, "s," "m," and the two vowel sounds. As different as it seems from the adult target, it preserved many of the target elements, including the "ah," from the end of "Fiona," the "ee" sound (although not in the accurate position), the nasality of the "n" but produced as an "m," and the sibilance of the "f," produced as an "s." It also fit her template of CVCV, much simpler than what was required by the target form. Importantly, Fiona was able to produce this version consistently so that those in her environment were able to learn she used it to refer to herself.

Predictable Errors on the Road to Mastery

Without the context of decades of developmental phonology research, Fiona calling herself "Sahmi" might seem like a quirk. Instead, this was a predictable occurrence, similar to the child-specific phonological factors accounted for by the A-Map model (McAllister Byun et al., 2016). According to this model, both developing children and adults balance the pressures of producing the accurate form of a speech sound or unit with producing the sound or unit in way that is achievable (i.e., to be precise). This accurate/precise tradeoff provides a simple explanation of the way that children speak in the toddler and preschool years. In the case of "Sahmi," Fiona's system recognized that her attempts to realize the adult target would, by necessity, be inaccurate and unintelligible because of the motoric limitations of her system. She weighted the target map on the side of precision rather than accuracy to improve her ability to be understood by those around her. Her solution to the precision/accuracy tradeoff was very unique, but pervasive individual variability is common at this stage of development. At the same time, the errors that are produced at these early stages are predictable across children, with stereotyped substitutions, omissions, and distortions of certain target sounds being common. Speech sounds that emerge later in childhood, like the "r," "k," and "l," are typically produced in error (Rose and Inkelas, 2011).

How, then, did Fiona go from referring to herself as "Sahmi" to "Fiona" in the course of a year? More precisely, how does speech production evolve as the child eliminates the constraints imposed by an immature speech motor control system? As she said, she had to "just keep trying." The trying and exploration begets a path that is cyclic and nonmonotonic (Green and Nip, 2010), meaning that children do not seamlessly transition from an inaccurate form to an accurate form. Instead, speech acquisition occurs in cycles characterized by periods of rapid acceleration and deceleration, with inflection points that alternate between weightings on "precise" and "accurate" goals (McAllister Byun et al., 2016). The progress advances toward what will be rather than building slowly from the enhancements of the current state. Fiona transitioned from a relatively stable period where the mastered, but inaccurate, "Sahmi" production was produced with precision. Working toward the more complex, "Fiona" production required a period of relative instability, or variability, where the productions were neither accurate nor precise as she worked toward a closer approximation of the adult target. For Fiona, this period of instability was characterized by her referring to herself with productions that began to sound something like "Feena." A period of relative stability and precision followed, with "Feena" being closer to accurate but still missing the mark. Producing the accurate "Fiona" required another period of instability followed, finally, by an accurate and precise target that she still uses to this day.

This cycling through periods of stability and instability with different levels of precision and accuracy has been supported empirically as well using the motion capture technology to quantify the stability and instability in measures of movement variability in a group of over 60 preschool-aged children (Vick et al., 2012). The measures from this large sample were used to subgroup the children. The result was that the performance of each child fell into one of three groups whose characteristics support the anecdotal evidence that the stages of later speech development follow a stair-stepping pattern. Specifically, like Fiona in her "Sahmi" stage, children start produce targets that are inaccurate but reliably produced. In an effort to push toward accuracy, a period of instability emerges, characterized by greater accuracy but much greater variability. About a third of the children in this study were in this stage. Finally, about one-third of the children were in a stage with both accuracy and precision, characterized by productions that were both correct and less variable.

When It Doesn't Go as Planned

The development of speech and the constraints of motor control continue, with refinement and learning occurring throughout childhood and into adolescence (Smith and Zelaznik, 2004). Mastery of all the sounds of American English can take up to eight years in a typical, monolingual child (Edwards, 1992). Speech sound production errors are common in the early years of speech production; over 15% of three-year-old children have speech sound disorders with no known cause (i.e., not attributable to hearing loss or neurological issues; Campbell et al., 2003). The prevalence of this disorder decreases by school age, with nearly 4% of 6-year-old children exhibiting speech sound disorders (Shriberg and Kwiatkowski, 1994). The vast majority of preschool children with speech sound disorders do not seem to have atypical speech motor control (Vick et al., 2014), which suggests that these ongoing errors are part of the individualized constraints and strategies employed by children. For example, Fiona's five-year-old sister, Annabel, was a comparatively precocious talker, producing her first words before her first birthday. Her first productions were likely facilitated by her responsive and facilitating sister (see Multimedia File 7 at acousticstoday.org/mm-vick), but her substitution of the "t" and "d" sounds for "k" and "g" certainly were not emulations of how Fiona produces these sounds. Sure enough, as Annabel approached her sixth birthday, these errors were resolved with the help of speech therapy.

Speech sound disorders that persist beyond the preschool years are of great concern because they are associated with later challenges in reading, spelling, and literacy (Lewis et al., 2015). When speech errors persist into later childhood (1% of 8-year-old children experience this; Wren et al., 2013), there is additional risk of social isolation, bullying from peers, and behavioral challenges that stem directly from the speech errors (e.g., Lindsay et al., 2002), all of which create significant challenges with quality of life. The underlying cause of these persisting errors likely lies in the constraints in speech motor control in early development that continue into the later years (Flipsen, 2003).

Conclusions

Any parent or grandparent can relate to the fascinating nature of speech development in infancy and early childhood. Combined with the catalysts of early language and cognitive development, it can be astonishing what children say and how they say it. Despite great advances in measuring the acquisition of speech, our understanding of how this complex skill develops still lacks clarity; it seems that as our knowledge advances, so, too, do our questions. For example. we are only beginning to understand how to identify preschool-aged children who are most at risk for speech sound disorders that persist into later childhood. If we could somehow sort through the 15% of 3-year-olds with speech sound errors to find the 1% who would face later challenges, we could provide intervention that might prevent the later associated problems that seem to extend into adolescence and adulthood (Lewis et al., 2015). Motion capture measurement is proving to be quite promising in this regard (e.g., Vick et al., 2014), and ongoing work seeks to combine these measures with candidate genetic and hormonal biomarkers (Anthoni et al., 2012) to predict which preschool children with speech sound disorders continue to be affected in later childhood. To be sure, because of the highly individualized trajectories that children follow on the road to mature speech production, large datasets and advanced subgroup discovery algorithms will be key to illuminating what characterizes typical and pathological developmental trajectories and how best to remediate speech sound disorders when they occur.

References

Anthoni, H., Sucheston, L. E., Lewis, B. A., Tapia-Paez, I., Fan, X., Zucchelli, M., Taipale, M., Stein, C. M., Hokkanen, M.-E., Castrén, E., Pennington, B. F., Smith, S. D., Olson, R. K., Tomblin, J. B., Schulte-Körne, G., Nöthen, M., Schumacher, J., Müller-Myhsok, B., Hoffmann, P., Gilger, L. W., Hynd, G. W., Nopola-Hemmi, J., Leppanen, P. H. T., Lyytinen, H., Schoumans, J., Nordenskjöld, M., Spencer, J., Stanic, D., Boon, W. C., Simpson, E., Mäkelä, S., Gustafsson, J.-Å., Peyrard-Janvid, M., Iyengar, S., and Kere, J. (2012). The aromatase gene *CYP19A1*: Several genetic and functional lines of evidence supporting a role in reading, speech and language. *Behavior Genetics* 42, 509-527.

Best, C. C., and McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech* 46, 183-216.

Campbell, T. F., Dollaghan, C. A., Rockette, H. E., Paradise, J. L., Feldman, H. M., Shriberg, L. D., Sabo, D. L,. and Kurs-Lasky, M. (2003). Risk factors for speech delay of unknown origin in 3-year-old children. *Child Development 74*, 346-357.

DeCasper, A. J., and Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science* 208, 1174-1176.

de Jong, K. (2018). Sensitivity to foreign accent. *Acoustics Today* 14(2), 50-57.

Edwards, H. T. (1992). *Applied Phonetics: The Sounds of American English*. Singular Publishing Group, Inc., San Diego, CA.

Flipsen, P. (2003). Articulation rate and speech-sound normalization failure. *Journal of Speech, Language, and Hearing Research* 46, 724-737.

Gibbon, F. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research* 42, 382-397.

Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., and Wilson, I. (2008). A motor differentiation model for liquid substitutions in children's speech. *Proceedings of Meetings on Acoustics* 1(1), 060003. https://doi.org/10.1121/1.2951481.

Golinkoff, R. M., Can, D. D., Soderstrom, M., and Hirsh-Pasek, K. (2015). (Baby) talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science* 24, 339-344.

Green, J. R., Moore, C. A., Higashikawa, M., and Steeve, R. W. (2000). The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research* 43, 239-255.

Green, J. R., Moore, C. A., and Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research* 45, 66-79.

Green, J. R., and Nip, I. B. (2010). Some organizational principles in early speech development. In B. Maassen and P. van Lieshout (Eds.), *Speech Motor Control: New Developments in Basic and Applied Research*. Oxford University Press, Oxford, UK, pp. 172-188.

Green, J. R., Nip, I. S., Wilson, E. M., Mefferd, A. S., and Yunusova, Y. (2010). Lip movement exaggerations during infant-directed speech. *Journal of Speech, Language, and Hearing Research* 53(6), 1529-1542.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684-686.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., and Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development* 1(3-4), 237-264.

Lecanuet, J. P., and Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 68, 1-23.

Lewis, B. A., Freebairn, L., Tag, J., Ciesla, A. A., Iyengar, S.,K., Stein, C. M., and Taylor, H. G., (2015). Adolescent outcomes of children with early speech sound disorders with and without language impairment. *American Journal of Speech-Language Pathology* 24, 150-163.

Lindsay, G., Dockrell, J., Letchford, B., and Mackie, C. (2002). Self esteem of children with specific speech and language difficulties. *Child Language Teaching and Therapy* 18, 125-143.

MacNeilage, P. F., and Davis, B. L. (2000). Deriving speech from nonspeech: A view from ontogeny. *Phonetica* 57, 284-296.

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

McAllister Byun, T. M., Inkelas, S., and Rose, Y. (2016). The A-map model: Articulatory reliability in child-specific phonology. *Language* 92, 141-178. https://doi.org/10.1353/lan.2016.0000.

Meltzoff, A. N., and Moore, M. K. (1977). Imitation of facial and manual gestures by human meonates. *Science* 198, 75-78.

Menard, L., Schwartz, J., and Boe, L. (2004). Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized French vowels from birth to adulthood. *Journal of Speech, Language, and Hearing Research* 47, 1059-1080.

Moore, C. A. (2004). Physiologic development of speech production. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, and W. Hultijn (Eds.), *Speech Motor Control in Normal and Disordered Speech*. Oxford University Press, Oxford, UK, p. 191-209.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J. Kavanagh, and C. Ferguson (Eds.), *Child Phonology, Vol. 1: Production*. Academic Press, New York, pp. 93-112.

Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., and Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences of the United States of America* 110, 15145-15150.

Rose, Y., and Inkelas, S. (2011). The interpretation of phonological patterns in first language acquisition. In C. J. Ewen, E. Hume, M. van Oostendorp, and K. Rice (Eds.). *The Blackwell Companion to Phonology*. Wiley-Blackwell, Malden, MA, pp. 2414-2438.

Shahidullah, S., and Hepper, P. G. (1994). Frequency discrimination by the fetus. *Early Human Development* 36, 13-26.

Shriberg, L.D., and Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech, Language, and Hearing Research* 37, 1100-1126.

Smith, A. (2006). Speech motor development: Integrating muscles, movements,

and linguistic units. *Journal of Communication Disorders* 30, 331-349.
Smith, A., and Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology* 45, 22-33.
Stoel-Gammon, C., and Cooper, J. A. (1984). Patterns of early lexical and phonological development. *Journal of Child Language* 11, 247-271.
Tamis-LeMonda, C., Kuchirko, Y., and Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science* 23, 121-126.
Tourville, J. A., and Guenther, F. H. (2010). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes* 26, 952-981. https://doi.org/10.1080/01690960903498424.
Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J. R., Abdi, H., Rusiewicz, H. L., Venkatesh, L., and Moore, C. A., (2012). Distinct developmental profiles in typical speech acquisition. *Journal of Neurophysiology* 107, 2885-2900.
Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J .R., Truemper, K., Rusiewicz, H. L., and Moore, C. A. (2014). Data-driven subclassification of speech sound disorders in preschool children. *Journal of Speech, Language, and Hearing Research* 57, 2033-2050.
Vihman, M.M. (2013). *Phonological Development: The First Two Years*. John Wiley and Sons, Oxford, UK.
Vihman, M. M., and McCune, L. (1994). When is a word a word? *Journal of Child Language* 21, 517-542.
Vorperian H. K., Kent, R. D., Lindstrom, M. J., Kalina, C.M., Gentry, L. R., and Yandell, B.,S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America* 117, 338-350. https://doi.org/10.1121/1.1835958.
Wren, Y., McLeod, S., White, P., Miller, L. L., and Roulstone, S. (2013). Speech characteristics of 8-year-old children: Findings from a prospective population study. *Journal of Communication Disorders* 46, 53-69.
Zubrick, S. R., Taylor, C. L., Rice, M. L., and Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research* 50, 1562-1592.

BioSketch

Jennell C. Vick, PhD, is the executive director of the Cleveland Hearing and Speech Center, a century-old nonprofit whose mission is to ensure that all people can communicate effectively. Her research focuses on implementation science in communication disorders, the impact of developmental communication disorders on the intergenerational poverty cycle, and improving outcomes in speech-language pathology, audiology, and vocational rehabilitation for adults who are D/deaf. She also studies the movements that children and adults use to produce spoken and signed languages using different mediums for motion capture as well as the effectiveness of visual biofeedback for the treatment of residual speech errors.

NGC® TESTING SERVICES
ACOUSTICAL • FIRE • STRUCTURAL • ANALYTICAL™
Over 50 Years of Test Services
NVLAP Accredited (Code 200291-0)
IAS Accredited (Lab Code 216)
ISO/IEC 17025 Compliant
Test Capabilities Include:
E492 IIC, E2179 Delta IIC E90 STC
C423 NRC, E1414 CAC, E1111 AC
Sound Power, Pipe Lagging, SAE J1400
E84 Flame Spread, E119 Fire Endurance
NGC Testing Services • 1650 Military Road
Buffalo, NY 14217 • Tel: 716-873-9750
ngctestingservices.com