Human Speech Sound Development

Kaitlyn Holtz

Duquesne University
Human Speech Sound Development

The language process begins with speech perception and evolves into expressive and receptive elements. The progression of language supports language acquisition theories of nature and nurture. Although these components establish language, each may independently vary to one’s situation, client, and practice.

Chapter One

Language structure comes from the simplest unit of speech known as a phoneme. By themselves, these units carry no meaning; but once arranged in specific combinations, they create morphemes that form words, phrases, and sentences. Speech perception allows an individual to take acoustic signals and interpret them into meaningful, structural, and pragmatic units. The ability to recognize different phonemes is vital to language development, especially during an infant’s first year. During infancy, a child is unable to reproduce all adult-like acoustic sounds and identifies sounds from adult productions. Observations and tests reveal known information on speech perception, as well as methods used to study infant categorization, discrimination, and preferences of lingual sequences.

Until 50 years ago, measuring an infant’s speech perception was limited because there was no explicit way to monitor their awareness. This changed once researchers began basing tests around visible infant capabilities including their sucking reflex, transfixion on objects, and head turn. When observing a child’s ability to discriminate speech sounds, the standard layout of a procedure started with habituation, normalizing the child to a stimulus, which was then followed with a test that determined an infant’s ability to categorize speech.

Sucking data was collected by recording the child’s baseline sucking reflex and then incorporating a habituation stimulus while the child remained above the baseline. Data was also
collected by using a child’s eye gaze. A child was shown a picture until they focused on it. Then the habituation stimuli was produced until the child broke focus. The visually reinforced head turn, VRHT, did not use habituation and was not infant controlled. VHRT took place in two stages: training, where the infant learned to turn towards a visual reinforcer when sound change occurred, and testing, where a person showed the child an object while in the presence of the reinforcer and sound stimulus. Certain age groups were more suited for the types of data collection such as high amplitude sucking extending up to four months and VRHT requiring an age over four months. The visual fixation was the method of choice because it could be used across a range of ages.

Vowels and consonants devise the two main phonetic categories. Vowels are created by the free flow of air through the vocal tract, while consonants constrict the airway to make a sound; both vowels and consonants have slight differences in manner, placement, and vary from language to language. Research found that infants were fairly good at discriminating among vowels and consonants, even when speakers were different genders or produced a vowel in an alternative pitch. The distinctions that children recognized included voice onset time, place, and manner of the generated sound.

Researchers also explored categorical perception, how variations of sounds fall on a continuum. This continuum has a categorical boundary determined by the midpoint of voice onset time, VOT. For example a VOT of 50 milliseconds may be recognized as /ba/ half of the time and “pa” the other half of the time. Additionally, versions of speech that lied on the same side on the continuum were more difficult to differentiate between due to the VOT. Researchers concluded that infants were able to perceive categories of consonants that differed in VOT, placement, and manner.
Although speech is an auditory phenomenon, visual cues make an impact on a person’s interpretation of speech. In one example, infants paid more attention to a video with matching visuals and sound. This showed an early awareness of visual and auditory information in children. In another study, infants watched a video that presented a speaker saying a short phrase, like /va/, with a different consonant vowel combination, like /da/, dubbed over what the speaker was actually producing. This experiment tested out the McGurk Effect and revealed that infants had slower reactions to the video. This allowed researchers to conclude that children, like adults, heard an utterance that was not /va/ or /da/.

Although studies documented that children were able to discriminate between various speech sounds, some adults and older children had trouble pinpointing differences between vowels or specific consonants. This led researchers to believe that a child’s environment and exposure had an effect on a person’s ability to discriminate between non-native phonemes. A study was conducted on infants whose native language was English; each group was shown a native phonetic contrast and two non-native contrasts. The earliest group, six to 8-month-old, found all three of the contrasts, the eight to 10-month-old group was less perceptive of the contrasts, and the 10 to 12-month-old group only recognized the native contrasts. This supported the belief that infants lose their innate ability to differentiate between sounds that are not prevalent in their own language.

Non-native vowels studies, although less frequent than consonant studies, took place. Results showed that vowel perception happened earlier than consonant perception and that there was directional asymmetry in vowel perception. This meant the order in which vowels were learned could make a difference when testing, since space in the oral cavity tends to be more natural in some vowel productions as opposed to others. One theoretical model known as PAM,
perceptual assimilation model, explained how speakers incorporate non-native speech into their own language. This model showed that children at the end of their first year will go from perceiving language on a general level to a specific level, focusing on certain rules of a specific language.

Another interest of researchers was whether children preferred listening to certain sounds. In a preference procedure, a child could select two knobs that would emit different sounds. During this task the infant chose to play the recording of the mother’s voice over musical episodes. Another study tested whether a child was more responsive to infant-directed speech, IDS, or adult-directed speech, ADS. This study showed that even when IDS occurred in the child’s non-native language, the participant still gravitated towards IDS. Another study assessed whether prosodic characteristics in speech affected the listener and showed that children prefer to listen to rhythmic patterns similar to their native language. Data also showed it is not until after six months that children start to prefer their native language or words they are already accustomed to, as opposed to other forms of unfamiliar speech.

Speech perception builds the base of language acquisition, which is vital to both speech and basic communication. This ability takes auditory or visual signals and transforms them into meaningful messages that allow an individual to form thoughts and develop relationships. Additionally, speech perception of infants is malleable. If a child receives minimal exposure to a native language, he or she may not pick up on distinct sounds that coincide with his or her chosen language.

Chapter Two

In regards to speech sound development, the debate between nurture and nature is not new. Some clinicians believe children learn language through their environment, while others
believe that language is genetic. The idea of nurture surrounds Behaviorist theory. This is the idea that an infant develops language through reinforcement, exposure, and imitation. Whereas, Nativist theory surrounds the concept that infants inherit genes that allow them to acquire language later on.

In the video, “Why Do We Talk?” researchers observe the emergence of language in children and provide examples that support both innate and learned theories. In one case, a language expert works with an autistic man, Christopher Taylor. Although Taylor is dependent on others for everyday activities, he can speak 23 languages. To test Taylor’s ability to learn a language, a man speaking Nahuatl, a language spoken in central Mexico, was brought in to introduce Taylor to a new language. In 10 minutes, Taylor could memorize all the words that he had been taught.

This case demonstrated how language in an individual is innate. Taylor struggled with tasks correlated to learned functions and behaviors. Even though Taylor was autistic, he was still capable of learning multiple languages that varied in form, content, and use. This suggested that all humans have a genetic component for language regardless of other mental difficulties.

Another idea from, “Why Do We Talk?” involved a study on animals. Researchers looked at the vocal tract of certain mammals when they made noise. Analysts discovered that the animals’ larynges descended when sound was emitted and that the vocal tracts moved dynamically. These anatomical findings showed that animals were able to produce speech. Since mammals’ inability to produce speech was not due to vocal anatomy, the difference between humans and animals was believed to lie in the human brain. To test this idea, researchers attempted to teach a monkey how to speak but failed.
DNA provides instructions for the specific order that allows for the development of a fetus. Differences in this code can cause mutations or alter brain development. If the brain does not form properly, language in an infant may not be correctly processed. The mammal study suggests that language is innate, since language is coded into human processing and not other mammal processing. This was seen when primates were unable to comprehend speech. Their inability to express learned linguistic patterns demonstrated how an innate piece is needed to interpret and produce language. This example supported the idea that a genetic component, unique to humans, is obtained from one’s genetic code.

Genetics played a major role in the story of the KE family. A gene passed down through a couple generations affected 15 of the 30 family members. These individuals all had a speech disorder that affected their expressive language but did not inhibit their thought processes. Research was conducted on the family and another child with similar characteristics. This led to the finding of a broken piece on chromosome 7. Furthermore, a mutation in the DNA of the KE family was identified, as well as a gene that controlled speech. This gene was identified as FOXP2 and was linked to motor movements around the mouth area.

A gene that directly correlated to speech was discovered, supporting innate speech sound production. The family members were not made to produce certain sounds. Even though the members of the KE family were taught language, they did not have the biological components to execute specific productions.

The BBC video also supported the idea of the learned side of language. In one case, a one-day-old baby was monitored using electrodes. These electrodes were able to monitor brain activity in and out of consciousness. During the experiment, three voice recordings containing the same sequence of words were played with a different person speaking in each recording. The
voices included a computer, the baby’s mother, and a stranger. The findings of this test found that the baby attended the most to her mother’s voice.

This experiment demonstrated how babies were able to perceive speech sounds early on. When in the womb, a fetus’s environment and exposure to certain sounds created neural connections in the brain. These connections were reinforced by the frequency of which a particular sound was heard. Here, the voice of the mother was what the child was most exposed to, so the child developed a learned connection to the particular sounds produced.

The idea of speech and language acquisition was discussed in a Nova episode called, “Secrets of a Wild Child.” This episode focused on a young girl, Genie Wiley. Wiley was known as a feral child due to the life she was subject to for 13 years. Her father was abusive and had left her in solitary confinement for extended periods of time, strapped her to a potty chair, and forced her into silence. Additionally, Wiley’s family did not engage in conversation with her. The closest she got to communication was when her dad would bark or growl at her. By the time Wiley was found, her habits no longer matched typically developing children of her age group.

Researchers quickly jumped to examine Genie Wiley and saw an opportunity to understand the development of language. Some initial observations about Genie included that she had an abnormal walk, could not properly chew, recognized only a couple words, and remained mostly silent. Those studying Wiley decided to start exposing her to the world that she had been isolated from. Wiley was brought out on excursions that allowed for her brain to be stimulated by new noises, smells, feelings, tastes, and sights. Researchers also began teaching Wiley language, where she was able to quickly learn words. She continued to explore and take interest in people, things, and the environment around her.
Although Wiley grew in both language and life skills, she was unable to properly form sentences. The messages she attempted to convey were clear, however, they lacked proper syntax. Genie Wiley’s story supported Behaviorist theory. Since Wiley was deprived of language during critical periods of language learning and development, she was never able to fully grasp all aspects of language.

With all this research in mind, speech sound development requires both innate and learned experiences. In order to form clear semantic, syntactic, pragmatic, morphological, and phonological thoughts, a child must have the correct genetic sequences to process language. Without a base to comprehend and express language, the child will not compare to the level of a typically developing child. It is also necessary for a person to be in an environment where he or she has a chance to use his or her senses to explore the world. Even if a child has a solid genetic base for language, without the appropriate language stimulation during early periods of development, it will be too late for an individual to fully develop speech sounds.

Chapter Three

There are general characteristics associated with infant vocal production. Some of these characteristics that occur during the 0 to 12 months range include engaging in non-reflexive and reflexive utterances, showing pre-linguistic expression, producing mostly vowels, assimilating lip and jaw movements, and beginning articulator coordination. These characteristics can be further categorized into four main stages. These stages include phonation, primitive articulation, expansion, and canonical babbling. In a YouTube video that can be found at https://youtu.be/CRD9RcZM1ho, the child exhibits characteristics in each of the four stages (iamalittlepeanut, 2008). The child in the video interacts with a cupboard promoting natural exploration and leading to sporadic sound productions.
Stage one, phonation, extends up until one-month-old. During this stage, a child has no formal language. The few speech sounds the infant produces include reflexive vegetative sounds as well as grunts. This stage is also marked by quasi-resonant nuclei as well as closed mouth sounds. The infant in the video is in a later developed stage than this one because the sounds she produces are clearly vowels and made with an open mouth. The infant does hum to herself with a closed mouth, but her understanding of how to create vowels places her out of stage one.

Stage two, primitive articulation, occurs during months two to three. Infants begin to experiment with sounds like cooing, gooing, laughing, and grunting. The girl in the video engages in audible laughs and grunts to herself when opening and closing the cupboard. These laughs and grunts are responses to the loud noises that occur when the cupboard is opened and closed. The grunting that takes place is also in response to the child’s size when she opens the cabinet. These grunts are unplanned by the child and occur subconsciously due to the weight of the door. Although the young girl fits the stage two profile, she has other speech sounds that are not included in this period.

Expansion, stage three, takes place from months four to six where infants gain greater control over laryngeal and oral systems and begin marginal babbling. The child in the video is at a higher level of babbling than marginal babbling, since she uses CV structures more complicated than CV and VC. However, the infant plays with other distinct characteristics found at this stage including pitch, loudness, vibrants, and intonation. In one instance, the child demonstrates using loudness and pitch together. She strokes the exterior of the cupboard and emits a loud squeal that goes from a low frequency to a high frequency. She also demonstrates intonation throughout her babbling where her voice gets higher at the end of her utterance.
From about 6 to 12 months, a child is considered to fall into the canonical babbling stage. This period is characterized by two types of babbling: reduplicated and variegated. Reduplicated babbling is the repeated change of alike syllables such as a CVCV sequence. The child demonstrates examples of reduplicated babbling on multiple occasions when she says /baba/ and /mama/. Variegated babbling involves a different sequence of multiple vowels and is seen in the video when the infant utters sequences such as /babamaba/. This stage is also recognized with certain consonant classes dominating over others. The girl favors bilabials when she produces /b/, /w/, and /m/ phonemes. Additionally, the infant explores the idea of lip coupling when she blows raspberries and when she properly executes the bilabial sounds /b/, w/, and /m/.

The child’s language is mostly composed of vowels. The vowels she has mastered include /a/, /o/, /ʌ/, /æ/, /ɔ/. The infant tends to use vowels that are backed in the mid to low range. She only produces one rounded vowel, probably due to her inability to fully control her lip movements. This child also produces multiple consonants including /b/, /m/, /w/, /d/, /h/. These occur during sequences of variegated babbling as well as individual instances. These consonants occur before the vowels, which is typical for a child. The child surpasses the 15-month stage of initial consonants mastered and falls into the 18-month stage. Here, the child should be producing six initial consonants including /b/, /d/, /m/, /n/, /h/, /w/. However, the child produces all the consonants in the 18-month stage except for /n/. In the 18-month stage the final consonant /t/ should be produced, but she does utilize any final consonants.

Although verbal productions are important when determining an infant’s stage of vocal development, nonverbal cues are vital in making sure that the child is on track with communication. In the YouTube video, the infant is attentive to the sounds around her, particularly the ones made by the two other women in the room. When the person behind the
camera laughs, the child turns to make eye contact with her; this also goes for the other lady in the room. At the end of the video, the front door of the house is opened and emits a creaking noise. This immediately prompts the child to lock eyes with the men entering and to observe them until she loses interest. As she watches the men, she makes obvious tongue thrusts and movements outside of her mouth. This could be seen as the child attempting to control her tongue movements or mirroring the talking that she sees. Earlier in the video, the girl demonstrates a successful mirrored task. The older woman in the room coughs while the child watches, and the young girl imitates the cough.

The child in the Youtube video appears between 6 to 12 months old. This young girl demonstrates physical signs in her age group including the near ability to walk as well as linguistic signs. Although the child produces more initial consonants than this stage dictates, she has yet to move into the linguistic stage and form any true words.

**Conclusion**

Speech-language pathologists must understand how language varies. This applies to theories, development, speech errors, disorders, and so on. Clients possess varying backgrounds, strengths, weaknesses, and impairments. Numerous approaches must be considered when developing a treatment plan specific to each client.
References

British Broadcasting Corporation (Producer). Why do we talk? Available from

