

The Relationship between Phoneme Production and Perception in Speech-Impaired and Typically-Developing Children

Karin Stromswold & Aliza Lichtenstein

One of the central questions that Eric Lenneberg raised in his seminal book, *Biological Foundations of Language* is: What is the relationship between language comprehension and language production? This paper reviews Lenneberg's case study of a child with congenital anarthria and then presents the results of two studies that investigate the relationship between phoneme perception and production. The first study investigates the phoneme identification skills of a child with developmental apraxia who, like the anarthric child studied by Lenneberg, had essentially no speech yet had no difficulty understanding speech. The second study investigates the extent to which 28 typically-developing children's ability to identify phonemes is related to their ability to produce phonemes. The results of both studies support Lenneberg's conclusion that children's ability to perceive speech is not dependent on their ability to produce speech. Thus, Lenneberg's original case study and the two studies presented in this paper argue against gestural theories of speech perception such as the Motor Theory.

Keywords: developmental apraxia; phoneme; production; perception; Motor Theory

1. Introduction

One of the hallmarks of much of Lenneberg's work and especially his seminal book *Biological Foundations of Language* is the importance he placed on the study of language acquisition by special populations and the insights that such populations can provide about the biological bases of language and language acquisition. A case in point is the production and perception of speech. It is well established that the articulatory gestures used to produce phonemes vary depending on the speaker, the situation in which speech occurs, and the phonological environment in which the phonemes appear. The result is that the acoustic realization of phonemes

We are grateful to all of the children who participated in the studies, and the parents, daycare providers and teachers who facilitated us in testing them. We thank Paul deLacy and Shigeto Kawahara for their advice on the design and implementation of the second experiment, and anonymous reviewers for their helpful comments and suggestions. These studies could not have been conducted without support from the Merck Foundation and the National Science Foundation (BCS-9875168; BCS-0446850).



varies tremendously both within and across speakers. By studying children with Down Syndrome, deafness and speech impairments, Lenneberg sought answers to three related questions that continue to haunt developmental psycholinguists. The first is how children learn to produce the articulatory gestures needed to produce phonemes. The second is how children (and by extension adults) perceive speech despite the acoustic variability associated with phonemes. The third question is, what is the relationship between the development of speech production and speech comprehension.

At the time Lenneberg was writing, behaviorist theories of language development held sway, with many researchers positing that children learned to talk by listening to their own babble and successively modulated their speech to match that of the people around them. According to gestural theories of speech perception such as the Analysis by Synthesis theory (Halle & Stevens 1962) and the Motor Theory (e.g., Liberman & Mattingly 1985, 1989), speech perception is to a greater or lesser extent parasitic on speech production. Proponents of the Motor Theory, for example, argue that there is a set of invariant motor commands (gestural scores) that underlie phoneme production and perception, and identification of the gestural scores associated with phonemes form the basis of speech perception (e.g., Liberman & Mattingly 1985, Liberman & Mattingly 1989, Liberman & Whalen 2000, Galantucci, Fowler & Turvey 2006).

As Lenneberg (1967) succinctly put it:

It is a fundamental assumption [of such theories] that responding is prior, in a sense, to understanding. However, there is a type of childhood abnormality that contradicts this assumption. These are children with inborn disability to coordinate their muscles of the vocal tract sufficiently to produce intelligible speech. The disturbance is seen in varying degree ranging from mild impediment to congenital anarthria.

(Lenneberg 1967: 305)

Over the course of the 1960s, Lenneberg conducted an in-depth study of a child with congenital anarthria who had no intelligible speech, yet had no difficulty understanding language. The child was born at 38 weeks gestation and small for gestational age. In addition to his expressive language disorder, the child had dysmorphic features (bilateral club feet, a hair lip, bilateral simian palmar creases, strabismus), "soft" neurological signs (e.g., difficulty distinguishing left from right), and mildly depressed IQ (between 70–85), all suggestive of a syndromic disorder. As an infant, the child reportedly cried normally, but never babbled. Throughout childhood, his vocal productions were extremely limited with occasional grunts that accompanied the gestures he used to communicate, and vocalizations that sounded like "Swiss yodeling" when he played.

Despite having profoundly impaired speech, Lenneberg reported that the child had no difficulty understanding what was said to him in either normal social settings or in experimental contexts. Lenneberg argued that this child's intact comprehension of language argued against theories that posited that children "learn" to speak by listening to their own babble and modulating their speech. Lenneberg further argued that,

since his own sounds are demonstrated to be objectively very different from those of the adults, the child must have some peculiar way of determining or recognizing similarities in the presence of diversifications.

(Lenneberg 1962: 126)

Since Lenneberg's landmark case study, a number of studies have investigated the relationship between speech perception and production in children with speech disorders. A study of children with cerebral palsy who were anarthric or dysarthric (a less severe form of anarthria) revealed that such children were just as good as control subjects at detecting whether the name of a picture was spoken correctly or altered by a single phoneme, indicating that they had no difficulty discriminating between phonemic contrasts that they could not produce (Bishop, Brown & Robson 1990). Researchers have also investigated the speech perception of children with developmental verbal dyspraxia (which is also referred to as childhood apraxia of speech, congenital apraxia or simply dyspraxia). As is the case with anarthria and dysarthria, dyspraxia affects all aspects of speech, with the most severely dyspraxic children having no intelligible speech. However, whereas dysarthria is a neuromotor disorder, dyspraxia is believed to be a motor-speech planning disorder, the hallmark of which is difficulty coordinating and executing the purposeful articulatory movements necessary for speech (see Hall, Jordan & Robin 1993).

Some studies suggest that dyspraxic children have intact phoneme perception despite their profound speech impairments. For instance, Hoit-Dalgaard et al. (1983) found no significant relationship between phoneme perception and production of voice onset time (VOT) in dyspraxic children. Groenen et al. (1996) conducted a study of Dutch-speaking dyspraxic children's ability to perceive and produce synthetically-produced minimal pairs of words that differed in place of articulation. In an identification task, dyspraxic children's identification function was equally as sharp as typically-developing children's, indicating that the phonetic processing of the two groups was equally consistent. However, in a discrimination task, the dyspraxic children had lower scores than the typically-developing children. Furthermore, the frequency with which the dyspraxic children made place of articulation errors was correlated with their scores on a place of articulation discrimination task (Groenen et al. 1996).

In contrast with Groenen et al.'s findings that dyspraxic children performed poorly on a phoneme discrimination task but not on a phoneme identification task, Sussman, Marquardt, Doyle & Knapp (2002) reported that all of the three dyspraxic children in their study performed aberrantly on a phoneme identification task. Marion, Sussman & Marquardt (1993) assessed four dyspraxic children's phonological awareness through a series of tasks that assessed the children's ability to produce and identify rhyming words. In striking contrast to age- and sex-matched controls who performed at or near ceiling on all tasks, the dyspraxic children were not only incapable of producing rhyming words, but they were at or near chance level on all the identification tasks including a simple rhyme recognition task. However, rather than attributing the dyspraxic children's poor performance on the perceptual tasks as being caused by their inability to produce rhymes as a motor theorist might, Marion et al. (1993) attributed the dyspraxic children's poor performance on both types of tasks as reflecting an underlying deficit in phoneme representation.

In summary, although the results of studies of the speech perception abilities of dyspraxic children are mixed, probably reflecting differences among studies in the criteria used to diagnose dyspraxia and the tests used to assess speech perception (Hall et al. 1993), at least some dyspraxic children appear to perceive speech normally. Furthermore, it is possible that abstract phonological deficits underlie both the production and the perceptual impairments exhibited by some dyspraxic children (Marion et al. 1993), or that perceptual deficits are the cause of dyspraxic children's impaired speech production.

Other studies have investigated the speech perception abilities of children with more circumspect articulatory impairments that affect the ability to produce particular phonemes. Some of these studies have failed to find a relationship between the ability to produce phonemes and the ability to perceive them. For example, Rvachew & Grawberg (2006) found that preschool children's articulatory accuracy was not related to their phonological awareness. In another study of preschool children with phonological impairments, Bird & Bishop (1992) found that all 14 children were able to discriminate between phonemic contrasts that they could not produce, with 7 of the 14 children performing near perfectly on the discrimination task. Similarly, Thyer & Dodd (1996) found no differences in auditory processing in children with impaired speech. In a study comparing the categorical perception abilities of children who did and did not have speech sound disorders, Johnson et al. (2011) found no difference between the groups in the sharpness or location of categorical boundary for synthetic stop-vowel (da/ta) syllables that varied in voice onset time (da/ta), but they did find marginally significant group differences for synthetic fricative-vowel (su/ju) syllables that varied in frequency of the friction noise.

In contrast with the studies mentioned above, some studies have found a correlation between phoneme perception and production in children with speech sound disorders. For example, Marquart & Saxman (1972) found a significant correlation between how often children with speech sound disorders misarticulated words and how often they misperceived words. Rvachew et al. (2003) found that misarticulating children have poorer phonemic perception of both correctly articulated words (e.g., *lake*) and incorrectly articulated words (e.g., *lake* as *wake*). In categorical speech perception studies with synthetically produced /r/ and /w/ tokens, children who frequently mispronounced /r/ as /w/ (e.g., saying *rabbit* as *wabbit*) had less clear categorical boundaries for /r/ and /w/ than children who did not mispronounce /r/ (Monnin & Huntington 1974, Hoffman et al. 1985, Ohde & Sharfe 1988). Less sharp categorical boundaries during perception tasks have also been found for other contrasts, such as /s-ts/ contrast in coda position (Raaymakers & Crul 1988) and fricatives (Rvachew & Jamieson 1989), a finding that Rvachew & Jamieson attributed to some misarticulating children having an underlying deficit in speech perception.

Given that impairments in speech perception are likely to result in impairments in speech production (e.g., as is evident in the impaired speech of most children with substantial hearing impairments), the mere correlation of speech production and speech perception abilities does not provide evidence for gestural theories of speech perception that posit that speech perception is parasitic on speech production. On the other hand, if even some children with impaired speech pro-

duction nonetheless have normal speech perception abilities, this argues *against* the primacy of speech production. We sought to further elucidate the relationship between speech perception and speech production in two studies.

2. Study 1—Case Study of Phoneme Identification in a Profoundly Dyspraxic Child

The first study investigated the speech perception abilities of a profoundly dyspraxic child who – like Lenneberg’s anarthric child – had no intelligible speech, yet appeared to understand everything that was said to him. In a phoneme identification task, we found that despite being unable to speak, the child had no difficulty understanding and discriminating among words that differed in phonemically minimal ways (e.g., *wake*, *lake* and *rake*), even when these words were said out of context.

2.1. Case History

2.1.1. Medical History

Review of the child’s medical records revealed that his prenatal course was unremarkable except for a mild case of polyhydramnios (a condition sometimes seen with oral motor problems) and a cesarean section delivery for breech presentation at 41.5 weeks gestation. Notably, he had no history of seizures, head injury, anoxic insult, or otitis media. All developmental milestones were reportedly achieved at the normal age, with the exception of an expressive language disorder first noted by his parents at 12 months of age and his pediatrician at 18 months of age. In contrast to the child Lenneberg studied, the child had no other delays or abnormal findings aside from his expressive language disorder. Specifically, he had no dysmorphic features, exhibited none of the “soft” neurological signs frequently observed in children with mild developmental disabilities, had no sign of cranial nerve damage, and had no difficulty producing simple rapid voluntary movements of the mouth or hands. He also had no history of excessive drooling or the sorts of feeding problems often associated with oral motor problems. Brainstem auditory evoked response potentials and audiometric examination revealed normal hearing bilaterally. Electroencephalography (EEG) and computed tomography (CT) scans (performed without contrast agent) were also reportedly normal.

2.1.2. Psychological Testing

At age 2;4 (years;months), the child’s performance on the Bayley Scales of Infant Development (Bayley 1969) was reportedly age-appropriate for all areas except for delays noted in language and fine motor skills. At 2;8, his performance on the Stanford-Binet Scale IV (Terman & Merrill 1960) and the Merrill-Palmer Psychomotor Scale (Stutsman 1981) were age appropriate and his performance on concrete problem solving were at the late 4-year-old level, suggesting average or above average intelligence and normal fine motor skills. The clinical psychologist who evaluated him at that time described him as a “pleasant, well-organized and independent little boy.”

2.1.3. *Language History*

The child's mother reported that his speech and expressive language development was markedly different from that of her five older children. He never babbled or cooed, but began to use points and gestures to communicate at or before a year of age. Despite having no expressive speech, his parents, therapists and doctors reported that he had no difficulty understanding what was said to him.

The child's language was formally evaluated for the first time when he was 2;4. According to the speech pathologist's report, he made no linguistic sounds, and communicated through points and gestures with an occasional grunt and high pitch squeal. His receptive language was at the early 2-year level and his expressive language was at the 6- to 12-month level on the Reynell Developmental Language Scales (RDLS, Reynell & Huntley 1971). Based on his history, vocalizations (or lack thereof) and RDLS scores, he was given the clinical diagnosis of developmental verbal dyspraxia.

The RDLS was repeated when the child was 2;8 at which time he scored at the 2;3 level on receptive section and at the 12-month level on the expressive section. At 2;8, AS phonological development was formally evaluated. According to the speech pathologist's report, his speech was grossly impaired at both the segmental and suprasegmental level: his vocal repertoire consisted of three sounds that were "consonant-like" (most closely resembling [d], [r] and [m]) and 2 or 3 sounds that were "vowel-like" ([u], [o] and possibly [i]), and these sounds were only used as isolated vowels and in simple consonant-vowel combinations. During the course of the evaluation, he produced only a handful of linguistic or nonlinguistic vocalizations, and he did not produce any vocalizations more complex than a single syllable, nor did he produce any intelligible words.

2.2. *Phoneme Identification Task*

2.2.1. *Stimuli*

At age 3;5, the child's ability to identify phonemes was assessed by having him point to pictures that depicted words that differed from one another in phonemically minimal ways (e.g., *van* and *fan*; *coat* and *goat*; *deer* and *tear*). Forty-four words were chosen because they were easy to depict, frequent, and were phonologically minimally distinct from other words on the list (see Appendix A). Of the 44 words, 9 had one phonological foil (e.g., *van* only had the foil *fan*), 15 had two phonological foils (e.g., *wake* had *lake* and *rake*), 9 had three phonological foils (e.g., *door* had *four*, *sore* and *shore*), 8 had four phonological foils (e.g., *wrap* had *cap*, *lap*, *map*, and *rat*), two words had 5 phonological foils (e.g., *mat* had *bat*, *cat*, *hat*, *map* and *rat*), and one word had 6 phonological foils (*cat* had *bat*, *cap*, *coat*, *hat*, *mat*, and *rat*). All of the words had at least one phonological foil that differed only in onset position (e.g., *fan* and *van*), 5 had at least one phonological foil that differed only in the vowel (e.g., *coat* and *cat*), and 8 had at least one phonological foil that differed only in coda position (e.g., *map* and *mat*). Some words differed from their phonological foils by only a single articulatory feature. For example, *goat* and *coat* differed only in voicing, *feet*, *seat* and *sheet* differed only in place of articulation, and *sea* and *tea* differed only in manner of articulation, whereas other words differed from one an-

other in more than one phonetic feature (e.g., *cat, hat, rat, mat*), or by the addition of a phoneme (e.g., *sea* and *seat, sore* and *store, ear* and *tear*).

2.2.2. Procedure

Forty-four colored pictures were placed in random order in front of the child. As each picture was laid out, the experimenter said the word depicted by the picture. Once all of the pictures were displayed, the child was told. "See these cards. We're going to play a game – I'm going to say a word and I want you to look very carefully and find the picture that matches what I say." The words were then read in random order. Words were said live, and if the child did not respond, the word was repeated up to two times. Each trial took approximately 1 minute, and the entire task took approximately 1 hour to complete. During the task, the child gestured, but made no attempt to say any of the words.

2.2.3. Results

For 42 of the 44 trials (93%), the child correctly chose the picture that matched the word. Even if we assume that, for each trial, the child selected randomly from the target word and a single phonological foil word (i.e., $p = .5$ for each trial), it is extremely unlikely that the child did this well by chance alone (cumulative binomial $p < .000001$).¹ Successful performance on a phoneme identification task requires not just the ability to perceive relatively subtle phonemic differences, but also knowledge of the meanings of words being tested and the ability to interpret the pictures correctly. Consider the child's two mistakes: for *hall* he chose the "door" picture (of a partially opened door) and for *sore* he pointed to the "tear" picture (of an eye with a tear). The semantic similarity – and the lack of phonological similarity – between the target words and the words he chose suggest that these errors reflect limitations in his picture identification skills or vocabulary, rather than his phoneme perception skills.

3. Study 2—Phoneme Identification and Production in Typically-Developing Children

Lenneberg's original study of an anarthric child and the case study of a dyspraxic child presented in the first study demonstrate that normal phoneme comprehension is possible even when phoneme production is profoundly impaired. In a second study, we investigated the relationship between phoneme comprehension and phoneme production in preschool-aged children who were typically developing. In a phoneme identification task, the children chose the picture that matched target words from among four pictures, and in a phoneme production task, the same group of children said the target words used in the phoneme identification task.

¹ Arguably, a more realistic hypothesis is that, the child selected randomly from among the target word and the foil words that differed for the target word by a single phoneme. Because each word had on average 2.6 phonologically minimal foils, on average, the probability of guessing correctly on a trial is .278 (1/3.6). Thus, selecting the correct picture for even 19 target words by chance alone is unlikely (binomial cumulative $p < .05$).

3.1. Methods

3.1.1. Participants

Twenty-eight (16 males and 12 females) monolingual, English-speaking children (mean age = 4.15; range 3.0–5.25) participated. All children were typically developing, with no history of speech, hearing, language or other impairment that might influence language development or interfere with their ability to perform the tasks. In addition, all of the children performed at age-appropriate levels on the Denver Articulation Screening Examination (DASE; Drumwright 1971, Drumwright et al. 1973).

3.1.2. Stimuli

There were 45 target words in the phoneme identification task (see Appendix B). Each target word was grouped with three distractor words that differed minimally from the target word to form a phonological minimal quartet. Quartets were designed to assess consonants in both onset and coda position because the acoustic features that distinguish between phonemes often differ depending on whether the consonants are onsets or codas (e.g., voice onset time affects perception of voicing for oral stops in onset position, whereas the duration of the preceding vowel affects perception of voicing for oral stops in coda position), and because children sometimes mispronounce the same phoneme differently in onset and coda position. For example, children tend to voice unvoiced consonants in onset position (e.g., mispronouncing *park* as *bark*) and de-voice consonants in coda position (e.g., mispronouncing *pig* as *pick*). Quartets were also designed to assess consonants in both consonant clusters and non-consonant clusters because the acoustics and articulation of consonants differ in clusters and non-clusters.

Of the 45 quartets, 30 assessed phoneme perception in onset position (e.g., target *rip* and distractors *lip*, *whip*, *zip*) and 15 assessed phoneme perception in coda position (e.g., target *pig* and distractors *pick*, *pin*, *pit*). In 30 quartets, the target word's onset or coda was a consonant cluster (target *snail* and distractors *sail*, *nail*, and *mail*) and in 15 quartets the target word was a non-cluster (e.g., target *buzz* and distractors *bug*, *bus*, *bud*). All target and distractor words were depictable, high frequency, monosyllabic words that are acquired at a young age.

Taken as a group, the 45 target words and their distractors assessed children's perception and production of consonants that differed in voicing, manner of articulation and place of articulation. We selected and grouped target and distractor words to maximize our ability to assess children's phoneme production and perception for common childhood speech errors (Sanders, 1972; Grunwell 1987) such as consonant cluster reduction (saying *snail* as *sail*), fronting errors in which the place of articulation of a phoneme is substituted (e.g., saying *crash* as *trash*), voicing errors (e.g., saying *park* as *bark*), stopping errors in which a fricative is said as an oral stop (saying *toe* for *sew*), gliding errors in which liquids are produced or perceived as glides (e.g., saying *wake* for *rake*) and /r/-/l/ substitutions (e.g., saying *rip* for *lip* or vice versa).

Phonological quartets were designed so that a single quartet targeted multiple aspects of phoneme perception and production. Consider the phonological

quartet *tree*, *tea*, *dee* (the letter “D”), and *three*. *Tree* and *tea* differ in that *tree* has a complex onset, *tea* and *dee* differ in that *tea* has an unvoiced onset, and *tree* and *three* differ in that *tree* begins with a stop consonant rather than a fricative. Thus, this one quartet potentially provides information about the production and perception of consonant clusters, voicing, and stopping. Because we used phonological quartets rather than phonological pairs in the phoneme identification task, there are 3 target-distractor word pairs and 2 distractor-distractor pairs for each trial, for a total of 225 word pairs (45 items \times 5 word pairs per item). Sixty-five word pairs differed in a way that targeted consonant cluster reduction, 28 word pairs targeted fronting errors, 25 word pairs targeted voicing errors, 16 word pairs targeted stopping errors, 7 word pairs targeted gliding, and 7 word pairs targeted /r/-/l/ substitution (see Appendix C).

3.1.3. Recordings

A native monolingual English-speaking woman who was naïve to the nature of the experiment and received no guidance regarding the pronunciation of the words said the instructions and target words used in the phoneme identification and perception tasks and in the DASE. Two experimental phonologists deemed that she had a typical New Jersey accent with no evidence of articulatory problems, and that she spoke clearly, but did not hyper-articulate. Words were recorded in a sound attenuated booth using a head-mounted *Shure Microphone* attached to a *Roland Edirol R09 Solid State Recorder* that recorded stimuli in 16 bit, 44.1kHz.wav format.

To avoid list intonation, each target word was inserted in the carrier sentence *say the word [...], twice*. The carrier sentence ended with the word *twice* in order to avoid phrase-final lengthening or creakiness in the target word. Target words were extracted from the carrier sentences using Praat (version 5.0.4, Boersma & Weenink, 2008). Each target word was recorded 9 times, and the best example of each target word was chosen using the criteria of naturalness, clarity, least background noise, and least aspiration. The amplitude of each word was then adjusted to a mean of 70 dB. Five monolingual English speakers with no background in linguistics and no knowledge of the experiment judged the target words to be natural sounding, clear, and similar to one another.

3.1.4. Experimental Apparatus

Visual stimuli were presented on a 15” Macbook Pro computer screen and audio stimuli were presented via *Sennheiser HD 202* headphones. Psyscope was used to control presentation of stimuli and record data, with the experimenter initiating trials and marking selections via an external keypad attached to the laptop. Children’s eye movements were recorded with the laptop’s built-in camera and the entire session was video-recorded.

3.1.5. Phoneme Identification Task

In a phoneme identification task, children listened to target words and selected the picture that matched each target word from among four pictures that depicted the target word and its phonological foils. For example, in one experimental trial,

children heard the target word *snail* and viewed pictures of a snail, a nail, a sail and mail (see figure 1). Prior to doing the experimental trials, children were given practice trials with phonologically distinct words (e.g., *star*, *bird*, *fork*, and *cheese*). If a child got a practice trial wrong or pointed to more than one picture, s/he was corrected. For the experimental trials, no feedback was given and, if a child pointed to more than one picture, his or her first selection was counted.



Figure 1: Sample Trial in Phoneme Identification Task. (Written words are presented for explanatory purposes only and did not appear in the task presentation.)

All trials began with a cartoon character appearing at the center of the screen for 2500 msec. After the cartoon fixation target disappeared, the screen went blank and the experimenter asked the child if s/he was ready. When the child was focused on the screen, the experimenter initiated a trial and the target word was played simultaneously with the four pictures. Stimuli were presented in pseudo-random order. The target picture did not appear in same quadrant more than 2 times in a row, the target segment was in onset position no more than 4 times in a row, and the target segment was a consonant cluster no more than 4 times in a row. In addition, onset and coda target segments and consonant cluster and non-cluster target segments occurred equally often in the first and second half of the list. Half of the participants received the items in the original order, and half received the items in the reverse order.

3.1.6. Standardized Articulation Assessment

After the children finished the phoneme identification task, we assessed their articulatory abilities using the DASE (Drumwright 1971; Drumwright et al. 1973), a test in which children repeat 22 words that contain 30 target phonemes. We used the same equipment and procedures for the DASE as we did for the phoneme production task.

3.1.7. Phoneme Production Task

Once children had completed the DASE and taken a short break, they repeated the 45 target words from the phoneme identification task. For each of the 45 target words, children viewed a picture of the target word while listening to the instruction “say the word [___].” (In order to ensure that the target words were acoustically identical in the identification and production tasks, the production task audio instructions were extracted from the same recordings that were used in the phoneme identification task.) Using pictures in the production task helped ensure that children’s errors were true mispronunciations and not the result of misunderstanding the target word (e.g., mistakenly saying *wake* rather than *lake* because they misperceived the target word as *wake*), and using a repetition task rather than a picture-labeling task ensured that children said the same words that were used in the identification task (e.g., they didn’t call a lake a pond).

For each trial, the audio instruction began at the same time that the matching picture appeared in the center of the computer screen. The picture remained on the screen until the child said the word. When the child finished saying the word, the experimenter pressed a key and the screen went blank. When the child indicated s/he was ready for the next trial, the experimenter pressed a key and the next trial began. Items were presented in the same order in the phoneme production task as they were in the phoneme identification task, and the apparatus that was used in the phoneme production task was the same as in the phoneme identification task except that, in addition to wearing headphones, children wore a head-mounted microphone that was attached to an Edirol recorder.

3.2. Results

3.2.1. Phoneme Identification

Because each target word had three phonological distractors, chance performance rate is 25%. Overall, children correctly identified the target word 63% of the time, and all children performed at significantly better than chance level. As expected, children’s performance was significantly correlated with their age ($r = .52, p = .005$). Accuracy and reaction times were analyzed using ANOVAs with subject as a random variable. There was no significant main effect of sex on either accuracy or reaction time (RT) regardless of whether incorrect trials were included or excluded (both F ’s < 1) and, thus, sex was excluded from all subsequent analyses.

Because there were no liquids in coda position, in this and subsequent onset/coda analyses, items in which the target segment contained a liquid were eliminated. When these items were excluded, children correctly identified significantly more onset target words than coda target words (73% and 54% respectively) by both non-parametric tests ($\chi^2 = 34.57, p < .00005$) and parametric tests with subject as a random variable ($F(1, 27) = 28.96, p < .0005, \eta_p^2 = .518$). Children were also significantly more accurate for target words without consonant clusters than target words with consonant clusters (67% and 61%, respectively) by both non-parametric tests ($\chi^2 = 5.52, p = .019$) and parametric tests with subject as a random variable ($F(1, 27) = 8.355, p = 0.008, \eta_p^2 = .236$).

Consistent with onset/coda accuracy results, when all trials were included, children were faster on onset target words than coda target words (2680 ms and 2991 ms respectively, $F(1, 27) = 8.52, p = 0.007, \eta_p^2 = .240$), but this difference was not significant when incorrect trials were eliminated (2489 and 2634 respectively, $F(1, 27) = 2.795, p = 0.11$). Also consistent with accuracy results, children were faster for target words that lacked consonant clusters compared to those that had consonant clusters, regardless of whether incorrect trials were included (2655 ms and 2985 ms respectively, $F(1, 27) = 25.89, p < .0005, \eta_p^2 = .490$) or excluded (2409 ms and 2766 ms respectively, $F(1, 27) = 18.64, p = .003, \eta_p^2 = .408$) from the analyses.

3.2.2. Phoneme Production

For the production task, if children said the target segment correctly, the item was scored correct even if they mispronounced other parts of the target word (e.g., if a child said *late* for *lake* and the onset was the target segment, the child received credit for having said the item correctly). Overall, children said the target segment correctly 85% of the time. As was the case for the phoneme identification task, children's accuracy rates on the phoneme production task were significantly correlated with their age ($r = .47, p = .01$). There was no significant difference in accuracy rate for boys and girls (83% and 88% correct respectively, $F < 1$), and sex was eliminated from all subsequent analyses. When trials with target words containing a liquid were excluded, children were not significantly worse at producing onset versus coda targets (87% accuracy for both) by either non-parametric or parametric tests ($\chi^2 = 0.43; F(1, 27) < 1$). Children were, however, significantly less accurate at producing consonant cluster targets than non-cluster targets (82% and 91%, respectively) by both non-parametric tests ($\chi^2 = 15.15, p < .00005$) and parametric tests ($F(1, 27) = 7.35, p = .01, \eta_p^2 = .214$).

3.2.3. The Relationship between Phoneme Identification and Phoneme Production

A multiple regression analysis of children's phoneme identification accuracy rates was conducted with age and phoneme production accuracy as predictors. This analysis revealed that children's age was a significant predictor of identification accuracy ($\beta = .428, t(25) = 2.25, p = .033$), but their phoneme production accuracy rate was not ($\beta = .185, t(25) = 0.97, p = .340$), with the overall fit of the model being fairly good ($R^2 = .292$). We next analyzed each child's data separately to determine whether individual children tended to misidentify and mispronounce the same target items. These analyses revealed Spearman's r of between $-.16$ and $.31$ (mean $r = .04$), with only one child's correlation coefficient being significant at the $p = .05$ level ($r = .31, p = .038$).

The above analyses simply address the question of whether children misidentify and mispronounce the same target items. A more precise question is whether children misperceived items in the same way that they mispronounce them. Collapsing across children, of the 1,260 trials (28 participants \times 45 target words), there were only 87 cases in which a child misidentifying and mispronounced the same target word. Of these 87 instances, in 78 cases children misidentified and mispronounced a target word in different ways (e.g., misidentifying the target word *flight* as being *light* and mispronouncing *flight* as *fight*). There were only 9 cases

in which a child misidentified and mispronounced a target word in the same way (e.g., misidentifying and mispronouncing the target word *robe* as *rope*).

Even though the children very rarely mispronounced and misidentified target words in the same way, it is possible that the same types of phonological processes underlie both their misidentification and mispronunciation errors. For example, a child who reduces fricative consonant clusters might misidentify *skis* as *keys* and mispronounce *spark* as *park*, yet correctly pronounce *skis* and correctly identify *spark*. To investigate whether the same phonological processes underlie children's misidentification and mispronunciation errors, we divided target segment phonemes into three groups based on their manner of articulation (oral stops, approximates and fricatives).²

For every phoneme in target onsets and codas that children either misidentified or mispronounced, we compared the place of articulation (POA), manner of articulation (MOA) and voicing of that target phoneme with the child's erroneous phoneme. In addition to tallying the children's POA, MOA and voicing errors, we also tallied their phoneme deletions and additions. Additions were very rare, with most cases of epenthesis involving the insertion of a vowel within a consonant cluster.³ In most cases, target phonemes and erroneous phonemes differed by a single feature (e.g., POA /s/ ⇒ /f/, MOA /s/ ⇒ /t/, voicing /s/ ⇒ /z/), but occasionally target and erroneous phonemes differed by more than one feature (e.g., /s/ and /d/ differ in both MOA and voicing). In all cases we assumed that children's errors differed minimally from the target segment. For consonant cluster targets, we considered each consonant separately. So, for example, if a child said /tr/ as /w/, we assumed that the child deleted the stop /t/ and said the approximate with the wrong POA. Table 2 provides an example of how errors were coded.

Using this coding procedure, we tallied the types of errors each child made on approximates, fricatives and stops in the identification and production tasks (see Appendix D). Inspection of these tallies suggests that the pattern of errors differed from child to child and that individual children had different patterns of errors on identification and pronunciation tasks. To test statistically whether children misidentified and mispronounced the same types of phonemes, for each child we determined the type of phoneme (approximate, fricative or stop) the child got wrong most often on the identification task and on the production task. Of the 28 children, two children were eliminated from the analysis because

Stop Errors	Approximate Errors
POA: tr ⇒ pr	POA: tr ⇒ tw, tl
MOA: tr ⇒ sr	MOA: tr ⇒ ts
Voicing error: tr ⇒ dr	Deletion: tr ⇒ t
POA + voicing: tr ⇒ br	
POA + MOA: tr ⇒ fr	
MOA + voicing: tr ⇒ zr	
Deletion: tr ⇒ r	
Addition: tr ⇒ tər	

Table 1: Sample error classification for the target cluster /tr/.

² For the purposes of the error analyses, affricates were treated as being composed of a stop followed by a fricative (tʃ = t+f, dʒ = d+z). Although there were 3 target segments that contained nasals (*snail*, *smell*, *crunch*), because children only made a handful of mistakes involving nasals, we chose to exclude them from our error analyses.

³ For the purposes of the analyses, epenthetic phonemes were classified with the phoneme that preceded them.

they had no production errors and 6 children were eliminated because, for one or both of the tasks, two phoneme types were tied for most common. Of the remaining 20 children, for 6 children the phoneme type that was most frequently mispronounced was also the phoneme type that was most often misidentified. Given that there were 3 types of phonemes, the probability that a child would make the most errors on the same phoneme type in both tasks is .33 by chance alone. Thus, we cannot reject the hypothesis that it was a chance occurrence that, for 6 of 20 children, the same phoneme type was the most commonly misidentified and mispronounced (cumulative binomial $p = .69$).

We next investigated whether children made the same types of errors on the identification and production tasks. The vast majority of children's errors were POA, MOA, voicing or deletion errors, so, for each child, we determined which of these 4 error processes was the most common for each of the two tasks. Two children were eliminated from the analysis because they had no production errors and 4 children were eliminated because for one or both of the tasks, two error processes were tied for most common. Of the remaining 22 children, for 8 children the most common error process on the identification task was the same as the most common error process on the production task. Given that there were 4 possible error processes, the probability by chance alone that the same error process would be the most common on both tasks is 0.25 (1 in 4) for each child. Thus, we cannot reject the hypothesis that it was a chance occurrence that the most common error process in the identification task was the same as the most common error process in the production task for 8 of 22 children (cumulative binomial $p = .16$).

4. Discussion

In 1962, Lenneberg wrote:

Our understanding of human behavior is often greatly enlightened by careful investigations of clinical aberrations and in many instances disease or congenital abnormalities provide conditions that may replace the crucial experiments on children that our superego forbids us to plan and perform. (Lenneberg 1962: 419)

Five years later, in *Biological Foundations of Language*, Lenneberg (1967) went one step further, arguing not just that special populations can provide important insights about language and language acquisition, but that

to ignore or overlook [such cases] is inexcusable as it may result in theories that are flatly contradicted by pertinent facts in pathology.

(Lenneberg 1967: 304)

The intact language comprehension abilities of Lenneberg's anarthric child and the dyspraxic child reported in this paper underscore the importance of studying how special populations use and acquire language. As Lenneberg so elegantly wrote:

The theoretical importance of extreme dissociation between perceptive and productive ability lies in the demonstration that the particular ability which we may properly call "having knowledge of a language" is

not identical with speaking. Since knowledge of a language may be established in the absence of speaking skills, the former must be prior, and, in a sense, simpler than the latter. (Lenneberg 1967: 308)

Contrary to Lenneberg's position, proponents of gestural theories of speech perception such as the Motor Theory argue that speech production is primary, that

the phonetic elements of speech, the true primitives that underlie linguistic communication, are not sounds but rather the articulatory gestures that generate those sounds. (Liberman & Whalen 2000: 188)

The discovery of mirror neurons in nonhuman primates has led to a resurgence of interest in gestural theories of speech perception such as the Motor Theory. Indeed, the results of some neuroimaging studies that show activation of motor areas during speech perception tasks appear to provide support for the Motor Theory. (For a critical review of such studies, see Hickok 2010). However, the fact that adults are able to perceive speech normally despite temporarily-induced impairments of speech production (e.g., Hickok et al. 2008) or acquired neurological insults that permanently impair speech production (e.g., Hickok et al. 2011) puts into question the Motor Theory's claim that articulatory gestures form the basis of adults speech perception (for a review see Stassenko, Garcea & Mahon 2013).⁴

It is logically possible that articulatory gestures/motor areas play a critical role in the development of speech perception, even if they no longer play such a role in speech perception in adults. Lenneberg (1964) argued that the linguistic abilities of children who are profoundly deaf, have Down Syndrome or have anarthria show that

motor skills are neither necessary nor sufficient prerequisites for the development of those psychological skills which seem to be an essential substrate for mature language. (Lenneberg 1964: 127)

Indeed Lenneberg (1962: 423) argued that, for children, "the vocal production of language is dependent upon the understanding of language but not vice versa." Furthermore, the fact that developmental impairments in speech perception almost always result in impairments in speech production means that studies that show children with speech impairments often have impaired speech production are not evidence for gestural theories of speech perception. In contrast, the existence of even a handful of children like Lenneberg's anarthric child and the dyspraxic child presented in the first study who have grossly impaired speech yet intact speech perception argues strongly against gestural theories of speech perception. Consistent with the primacy of speech perception abilities, Rvachew, Nowak & Cloutier (2004) found that training children with phonological expressive delays to attend to phonemic contrasts (by providing them feedback on a task very much like the experimental task used in the second experiment) improved the children's ability to produce these contrasts.

⁴ One might argue, however, that some weaker version of the Motor Theory could accommodate these findings. Perhaps, for example, the gestural scores could be encoded in an abstract manner that does not involve the motor regions per se. However, as noted by Hickok and others, this leaves open the possibility that the underlying encoding is fundamentally perceptual rather than motoric.

According to Liberman & Whalen (2000),

co-articulation creates a complex relationship between the acoustic signal and the phonetic structure [...]. Unraveling that complex relationship between signal and message is the business of the same phonetic module that produced it, for that module incorporates the constraints necessary to process the signal so as to recover the very gestures that were, by their co-articulation, responsible for its apparent complications.

(Liberman & Whalen 2000: 189)

If this were true, not only would we fail to find cases like Lenneberg's anarthric child and the dyspraxic child presented in the first study, we would predict a causal link between speech production and speech perception in typically-developing children. Contrary to this prediction, in the second study, we failed to find any evidence of a relationship between typically-developing children's ability to identify and produce phonemes. A multiple regression analysis revealed that children's accuracy on the phoneme production task was not a significant predictor of their accuracy on the phoneme identification task independent of their age, and analyses of individual children's performance on the items in the identification and production tasks revealed a significant correlation for only one of the 28 children.

Consistent with the results of these regression analyses, error analyses revealed that the children in the second study rarely got the same items wrong on the identification and pronunciation tasks and, when they did, they almost always did so in different ways (e.g., misidentifying the target word *lake* as *rake* and pronouncing it as *wake*). Furthermore, when we analyzed each children's patterns of errors in the two tasks, we found they diverged considerably. First, we found no evidence that children typically made mistakes on the same class of phonemes (approximates, fricatives and stops) in the identification task and the production task. Second, we found no evidence that children made the same types of errors (e.g., voicing errors, POA errors, MOA errors, deletion errors etc.) in the two tasks.

Taken as a whole, many of the developmental studies reviewed in this paper and the results of the two studies presented provide strong evidence against the claim that speech production serves as the developmental backbone of speech perception and against gestural theories of speech perception more generally.

Appendix A: Study 1—Target Words and Phonological Foils

Target Word	Onset substitution	Vowel substitution	Coda substitution	Phoneme addition or deletion
ball	doll, hall, wall			
bat	cat, hat, mat, rat	boat		
boat	coat, goat			
cap	lap, map, wrap		cat	
cat	bat, hat, mat, rat	coat	cap	
cheese	knees			
coat	boat, goat	cat		
deer	tear			ear
dish	fish			
dog	log		doll	
doll	ball, hall, wall		dog	
door	four, shore, sore			
ear				deer, tear
fan	van			
feet	seat, sheet			
fish	dish			
floor				four
four	door, shore, sore			floor
goat	boat, coat			
hall	ball, doll, wall			
hat	bat, cat, mat, rat			
lake	rake, wake			
lap	cap, map, wrap			
log	dog			
mat	bat, cat, hat, rat		map	
map	cap, lap, wrap		mat	
knees	cheese	nose		
nose	toes	knees		
rake	lake, wake			
rat	bat, cat, hat, mat		wrap	
sea	tea			seat
seat	feet, sheet			sea
sheet	feet, seat			
shore	door, four, sore			
sore	door, four, shore			store
store				sore
tea	sea			tree
tear	dear			ear, tea
toes	nose			
tree	sea			tea
van	fan			
wake	lake, rake			
wall	ball, doll, hall			
wrap	cap, lap, map		rat	

Appendix B: Study 2—Phoneme Identification Stimuli Quartets

The target segment is in bold, followed by the three distractor words.

ONSETS	CODAS
SIMPLE C (N = 7)	SIMPLE C (N = 8)
<i>Fricative</i>	<i>Fricative</i>
s -eat sheet street eat	ri- ce write ride rise
s -ing string sting wing	bu- zz bug bus bud
f -at flat rat sat	<i>Stop</i>
<i>Stop</i>	pi- g pick pin pit
t -oe sew row throw	co- ke cone coat comb
g -as grass class glass	ba- d bat bag badge
<i>Approximate</i>	k- id kick kit king
l- ake rake wake ache	wro- te road rope robe
r- ip lip whip zip	ro- be wrote road rope
CC CLUSTER (N = 18)	CC CLUSTER (N = 6)
<i>Fricative-Stop</i>	<i>Fricative-Stop</i>
sp -ark bark park shark	bea- st bees beat beach
sk -is keys seas squeeze	cru- st crushed crunch crutch
<i>Fricative-Nasal</i>	fi- st fish fin fizz
sn -ail sail mail nail	<i>Nasal-Stop</i>
sm -ell shell sell bell	pa- nt pats pans pants
<i>Fricative-Approximate</i>	<i>Affricate (Stop-Fricative)</i>
sl -eep leap weep sweep	ca- tch cash cat cap
fr -y lie tie fly	ba- dge bad bat bag
fl -ight light white write	
<i>Stop-Approximate</i>	
bl -ed red bread bed	
gr -ab crab cab stab	
br -eak bake rake wake	
cr -ash trash cash rash	
tr -ail tail rail whale	
tr -ee tea dee three	
cl -ock block lock rock	
<i>Affricate (Stop-Fricative)</i>	
ch -ase face lace vase	
ch -ew shoe zoo two	
ch -ick sick tick trick	
ch -ip sip tip ship	
CCC CLUSTER (N = 5)	CCC CLUSTER (N = 1)
<i>Fricative-Stop-Approximate</i>	<i>Nasal-Stop-Fricative</i>
scr -eam cream steam stream	cru- nch crust crutch crush
str -ip trip rip tip	
spr -ing swing sing wing	

Appendix C: Study 2—Common Phonological Error Processes Assessed in Phoneme Identification Target-Distractor Word Pairs

Error Type	Onset Distractors	Coda Distractors	Distractors
Fronting	17 (e.g., crash > trash)	11 (e.g., coke > coat)	28
Voicing	12 (e.g., grab > crab)	13 (e.g., rice > rise)	25
Stopping	5 (e.g., fry > tie)	11 (e.g., rice > write)	16
Gliding	7 (e.g., lake > wake)	0	7
/r/-/l/ substitution	7 (e.g., rip > lip)	0	7
Simple C Deletion	2 (e.g., seat > eat)	0	2
Cluster Reduction	51 (e.g., snail > nail)	14 (e.g., beast > beat)	65

Appendix D: Study 2—Classification of Identification and Pronunciation Errors

Participants are listed in order according to the combined number of errors they made in the two tasks, with the child who made the most errors given first.

Age	# Shared Errors (# identical)	Identification Errors ¹	Pronunciation Errors
3.33	12 shared errors (2 identical)	8 stops (4 delete, 2 voice, 1 MOA, 1 POA+voice) 8 fricative (5 delete, 1 POA, 1 voice, 1 POA+MOA) 5 approximates (4 POA, 1 POA+MOA)	23 fricatives (19 delete, 3 MOA, 1 POA) 8 approximates (8 POA) 8 stops (7 delete, 1 voice)
3.75	10 shared errors (1 identical)	13 stops (6 delete, 3 voice, 2 POA, 1 POA+voice, 1 other) 11 fricatives (5 delete, 3 POA, 1 POA+voice, 1 POA+MOA, 1 POA+MOA+voice) 3 approximates (2 POA, 1 delete)	17 approximates (15 delete, 2 POA) 7 fricatives (3 POA+voice, 2 POA, 1 delete, 1 POA+MOA) 4 stops (3 delete, 1 MOA,
3.00	14 shared errors (1 identical)	10 stops (5 delete, 2 MOA, 2 voice, 1 POA+voice) 10 fricatives (4 POA, 4 delete, 2 voice) 6 approximates (5 delete, 1 POA)	9 approximates (8 POA, 1 delete) 8 stops (3 POA, 2 add, 2 voice, 1 delete) 7 fricatives (5 add, 2 delete)
3.33	6 shared errors (0 identical)	10 stops (4 delete, 2 POA+voice, 1 voice, 1 MOA, 1 POA+MOA) 8 fricatives (6 delete, 1 POA, 1 POA+MOA) 5 approximates (3 delete, 1 POA, 1 POA+MOA)	6 approximates (6 POA) 4 fricatives (3 POA, 1 add) 4 stops (2 voice, 1 add, 1 delete)
4.75	4 shared errors (0 identical)	13 stops (6 delete, 2 voice, 2 POA, 1 MOA, 1 POA+voice, 1 POA+MOA) 9 fricatives (6 delete, 2 POA+MOA, 1 POA) 7 approximates (5 POA, 2 delete)	4 fricatives (4 POA) 3 stops (1 delete, 1 MOA, 1 add)
3.42	6 shared errors (0 identical)	12 stops (6 delete, 3 voice, 2 POA, 1 add) 10 fricatives (7 delete, 2 POA, 1 POA+MOA, 1 add) 8 approximates (4 POA, 3 delete, 1 POA+MOA)	7 fricatives (4 POA, 2 add, 1 delete) 2 stops (1 delete, 1 add)
3.50	3 shared errors (0 identical)	14 stops (7 delete, 2 POA, 1 MOA, 1 voice, 1 POA+MOA, 1 add, 1 POA+voice) 14 fricatives (9 delete, 4 POA, 1 voice) 8 approximates (7 POA, 1 POA+MOA)	3 fricative (3 add) 2 stops (2 add)
4.08	4 shared errors (0 identical)	11 stops (6 delete, 3 voice, 1 POA+voice, 1 add, 8 fricatives (3 delete, 2 voice, 2 POA, 1 POA+MOA) 6 approximates (3 delete, 3 POA)	3 fricatives (3 POA) 2 approximates (2 POA) 2 stops (1 delete, 1 add)
3.42	1 shared errors (0 identical)	12 fricatives (8 delete, 3 POA, 1 POA+MOA) 9 stops (6 delete, 1 voice, 1 POA, 1 POA+voice) 4 approximates (4 POA)	3 approximates (POA) 1 fricative (delete)
4.83	4 shared errors (1 identical)	10 stops (3 voice, 3 delete, 1 POA, 1 MOA, 1 POA+voice, 1 add) 7 fricatives (4 delete, 1 voice, 1 POA, 1 POA+MOA) 1 approximates (POA)	9 stops (4 add, 3 delete, 2 voice) 5 fricatives (3 POA, 2 add) 2 approximates (2 POA)
3.75	3 shared errors	11 stops (5 voice, 3 POA, 2 delete, 2 add)	4 fricatives (4 add)

	(0 identical)	10 fricatives (5 delete, 2 POA, 2 POA+MOA, 1 voice) 2 approximates (1 delete, 1 POA)	1 stops (1 add) 1 approximates (1 delete)
3.17	1 shared error (0 identical)	12 fricatives (8 delete, 2 POA+voice, 1 POA, 1 voice) 8 stops (5 delete, 2 POA, 1 POA+MOA, 1 add) 5 Approximate (4 POA, 1 delete)	2 stops (1 voice, 1 delete) 1 Approximate (1 delete)
3.42	3 shared errors (0 identical)	9 fricatives (6 delete, 2 POA, 1 POA+MOA) 7 stops (2 delete, 2 voice, 1 POA, 1 MOA, 1 POA+MOA) 2 approximates (1 delete, 1 POA)	5 fricatives (3 POA, 2 POA+voice) 4 approximates (4 POA) 2 stops (2 add)
4.92	3 shared errors (0 identical)	8 stops (4 delete, 2 voice, 2 POA) 7 fricatives (4 POA, 1 delete, 1 POA+MOA, 1 MOA+voice) 3 approximates (2 delete, 1 POA)	4 approximates (4 POA) 2 stops (add)
4.50	2 shared errors (1 identical)	13 fricatives (9 delete, 3 POA, 1 POA+MOA) 7 stops (6 delete, 1 POA) 6 approximates (3 POA, 1 delete, 1 POA+MOA, 1 POA+MOA+voice)	1 fricative (delete) 1 stops (voice)
3.17	1 shared error (0 identical)	7 stops (4 delete, 2 voice, 1 POA) 5 fricatives (2 delete, 2 POA, 1 POA+MOA) 2 approximates (2 POA)	6 Fricative (6 POA) 1 stops (add)
3.50	2 shared errors (1 identical)	12 stops (5 delete, 4 voice, 2 POA, 1 MOA) 5 fricatives (2 delete, 3 POA) 2 approximates (2 POA)	1 stops (voice) 1 fricative (POA)
5.17	2 shared errors (0 identical)	9 approximates (4 delete, 3 POA, 1 POA+MOA, 1 POA+MOA+voice) 7 fricatives (4 delete, 1 POA, 1 voice, 1 MOA) 6 stops (3 delete, 1 voice 1 POA, other)	2 approximates (2 POA)
3.75	2 shared errors (1 identical)	11 stops (5 voice, 4 delete, 1 POA, 1 MOA) 8 fricatives (6 delete, 1 POA, 1 MOA) 2 Approximate (1 POA, 1 delete)	2 stops (2 delete, 1 voice) 1 fricative (POA) 1 approximates (delete)
4.83	1 shared error (0 identical)	8 stops (3 voice, 3 delete, 2 POA, 1 add) 4 fricatives (3 deleted, 1 POA+MOA) 3 approximates (3 POA)	1 stops (delete)
3.97	0 shared errors (0 identical)	9 stops (3 POA, 3 delete, 2 voice, 1 MOA) 6 fricatives (6 delete) 6 approximates (5 POA, 1 delete)	1 stops (delete)
4.00	2 shared errors (0 identical)	7 stops (3 delete, 2 voice, 2 MOA) 4 fricatives (2 reduc, 1 POA+MOA, 1 add) 1 approximate (POA)	4 approximates (4 POA)
5.25	0 shared errors (0 identical)	7 fricatives (4 delete, 2 POA, 1 POA+MOA) 5 stops (2 delete, 2 voice, 1 MOA) 3 approximates (2 POA, 1 MOA)	1 stops (add)
4.25	0 shared errors (0 identical)	7 fricatives (4 delete, 1 voice, 1 POA, 1 POA+MOA)	2 stops (1 voice, 1 POA+MOA)

		4 stops (2 voice, 1 delete, 1 POA+MOA) 2 approximates (2 POA)	
4.42	1 shared error (1 identical)	6 fricative (4 delete, 1 POA, 1 MOA) 5 stops (2 delete, 2 voice, 1 POA+MOA) 1 approximate (1 POA)	1 stops (1 delete) 1 Fricative (1 delete) 1 Approximate (1 add)
5.17	0 shared errors (0 identical)	7 stops (4 delete, 2 voice, 1 add) 4 fricatives (3 delete, 1 POA) 3 approximates (2 delete, 1 POA)	No errors
4.83	0 shared errors (0 identical)	4 fricative (2 delete, 1 voice, 1 MOA) 4 stops (2 delete, 1 POA, 1 add) 1 approximate (1 POA)	1 stops (1 voice)
5.17	0 shared errors (0 identical)	5 stops (2 voice, 1 delete, 1 POA, 1 MOA) 1 fricative (1 POA) 1 approximate (1 POA)	No errors

ⁱ Error code:

Voice = voicing error (e.g., *robe* => *rope*; *pig* => *big*)
 POA = Place of Articulation error (e.g., *crash* => *trash*, *catch* => *cats*)
 MOA = Manner of Articulation error (e.g., *sing* => *ting*, *trail* => *srail*)
 Delete = deletion error (e.g., *seat* => *eat*; *break* => *rake*; *strip* => *rip*)
 Add = epenthesis error (e.g., *skis* => *sakis*)

References

- Bayley, Nancy. 1969. *Bayley Scales of Infant Development*. New York: Psychological Corporation.
- Bird, Judith, Dorothy V. M. Bishop, & Norman H. Freeman. 1995. Phonological awareness and literacy development in children with expressive phonological impairments. *Journal of Speech and Hearing Research*, 38(2), 446–462.
- Bird, Judith & Dorothy V. M. Bishop. 1992. Perception and awareness of phonemes in phonologically impaired children. *European Journal of Language & Communication Disorders* 27(4), 289–311.
- Bishop, Dorothy V. M., B. Byers Brown, & J. Robson. 1990. A relationship between phoneme discrimination, speech production, and language comprehension in cerebral palsied individuals. *Journal of Speech and Hearing Research* 33(2), 210–219.
- Boersma, Paul & David Weenink. 2008. *Praat: Doing Phonetics by Computer*. Version 5.0.4. Available from <http://www.praat.org>.
- Drumwright, Amelia. 1971. *Denver Articulation Screening Examination*. Denver, CO: University of Colorado Medical Center.
- Drumwright, Amelia, Pearl Van Natta, Bonnie Camp, William Frankenburg, & Hazel Drexler. 1973. The Denver Articulation Screening Exam. *Journal of Speech and Hearing Disorders* 38(1), 3–14.
- Galantucci, Bruno, Carol A. Fowler, M. T. Turvey. 2006. The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review* 13(3), 361–377.
- Groenen, Paul, Ben Maassen, Thom Crul, & Geert Thoonen 1996. The specific relation between perception and production errors for place of articulation in developmental apraxia of speech. *Journal of Speech and Hearing Research*, 39(3), 468–482.
- Grunwell, Pamela. 1987. *Clinical Phonology*. London, England: Croom Helm.
- Hall, Penelope K., Linda S. Jordan, & Donald A. Robin. 1993. *Developmental Apraxia of Speech*. Austin, TX: Pro-Ed.
- Halle, Morris & Kenneth N. Stevens. 1962. Speech recognition: A model and a program for research. *IRE Transactions on Information Theory* 8(2), 155–159.
- Hickok, Gregory. 2010. The role of mirror neurons in speech perception and action word semantics. *Language and Cognitive Processes* 25(6), 749–777.
- Hickok, Gregory, K. Okada, W. Barr, J. Pa, C. Rogalsky, . . . , & A. Grant. 2008. Bilateral capacity for speech sound processing in auditory comprehension: Evidence from Wada procedures. *Brain and Language*, 107(3), 179–184.
- Hickok, Gregory, Maddalena Costanzo, Rita Capasso, Gabriele Miceli. 2011. The role of Broca's area in speech perception: Evidence from aphasia revisited. *Brain and Language* 119(3), 214–20.
- Hoffman, Paul R., R. G. Daniloff, D. Bengoa, & Gordon Schuckers. 1985. Misarticulating and normally articulating children's identification and discrimination of synthetic [r] and [w]. *Journal of Speech & Hearing Disorders* 50(1), 46–53.
- Hoit-Dalgaard, Jeannette, Thomas Murry, Harriet G. Kopp. 1983. Voice onset time production and perception in apraxic subjects. *Brain and Language* 20(2), 329–339.

- Hustad, Katherine C. 2008. The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech, Language, and Hearing Research* 51(3), 562–573.
- Johnson, Erin P., Bruce F. Pennington, Joanna H. Lowenstein, & Susan Nittrouer. 2011. Sensitivity to structure in the speech signal by children with speech sound disorder and reading disability. *Journal of Communication Disorders* 44(3), 294–314.
- Kuijpers, Cecile T. L. 1996. Perception of the voicing contrast by Dutch children and adults. *Journal of Phonetics* 24 (3), 367–382.
- Lenneberg Eric H. 1962. Understanding language without ability to speak: A case report. *Journal of Abnormal and Social Psychology* 65(6), 419–425.
- Lenneberg, Eric H. 1964. Speech as a motor skill with special reference to nonaphasic disorders. *Monographs of the Society for Research in Child Development* 29(1), 115–127.
- Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: Wiley.
- Lieberman, Alvin M. & Ignatius G. Mattingly. 1985. The motor theory of speech perception revised. *Cognition* 21(1), 1–36.
- Lieberman, Alvin M. & Ignatius G. Mattingly. 1989. A specialization for speech perception. *Science* 243(4890), 489–494.
- Lieberman, Alvin M. & Douglas H. Whalen. 2000. On the relationship of speech to language. *Trends in Cognitive Sciences* 4(5), 187–196.
- Lotto, Andrew J., Gregory S. Hickok, & Lori L. Holt. 2009. Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences* 13(3), 110–114.
- Marion, Michelle J., Harvey M. Sussman, & Thomas P. Marquardt. 1993. The perception and production of rhyme in normal and developmentally apraxic children. *Journal of Communication Disorders* 26(3), 129–160.
- Marquardt, Thomas P. & John H. Saxman. 1972. Language comprehension and auditory discrimination in articulation deficient kindergarten children. *Journal of Speech, Language, and Hearing Research* 15(2), 382–389.
- McAllister Byun, Tara, Mark Tiede. 2017. Perception-production relations in later development of American English rhotics. *PLoS ONE* 12(2), e0172022.
- Monnin, Lorraine M. & Dorothy A. Huntington. 1974. Relationship of articulatory defects to speech-sound identification. *Journal of Speech and Hearing Research* 17(3), 352–366.
- Morgan, Rosemarie A. 1984. Auditory discrimination in speech impaired children and normal children as related to age. *British Journal of Disorders of Communication* 18(1), 89–96.
- Ohde, Ralph N. & Donald J. Sharf. 1998. Perceptual categorization and consistency of synthesized /r-w/ continua by adults, normal children and /r/- misarticulating 26 children. *Journal of Speech, and Hearing Research*, 31, 556–568.
- Raaymakers, Emile M. & Thom A. Crul. 1988. Perception and production of the final /s-ts/ contrast in Dutch by misarticulating children. *Journal of Speech and Hearing Disorders* 53(3), 262–270.
- Reynell, Joan, & R. M. C. Huntley. 1971. New scales for the assessment of language development in young children. *Journal of Learning Disabilities* 4(10), 10–18.
- Rvachew, Susan, Michele Nowak, & Genevieve Cloutier. 2004. Effect of phonemic perception training on the speech production and phonological aware-

- ness skills of children with expressive phonological delay. *American Journal of Speech-Language Pathology*, 13, 250–263.
- Rvachew, Susan, & M. Grawberg. 2006. Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research* 49, 74–87.
- Rvachew, Susan, & Donald G. Jamieson. 1989. Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders* 54(2), 193–208.
- Rvachew, Susan, Alyssa Ohberg, Meghann Grawburg, & Joan Heyding. 2003. Phonological awareness and phonemic perception in 4-year-old children with delayed expressive phonology skills. *American Journal of Speech-Language Pathology* 12, 463–471.
- Sander, Eric K. 1972. When are speech sounds learned? *Journal of Speech and Hearing Disorders* 37(1), 55–63.
- Shuster, Linda I. 1998. The perception of correctly and incorrectly produced /r/. *Journal of Speech, Language, and Hearing Research* 41, 941–950.
- Stasenko, Alena, Frank E. Garcea, & Bradford Z. Mahon. 2013. What happens to the motor theory of perception when the motor system is damaged? *Language and Cognition* 5(2-3), 225–238.
- Sussman, Harvey M., Thomas P. Marquardt, J. Doyle, & H. Knapp. 2002. Phonemic integrity and contrastiveness in developmental apraxia of speech. In F. Windsor, N. Hewlett, & L. Kelly (eds.), *Themes in Clinical Linguistics and Phonetics*, 311–326. Mahwah, NJ: Lawrence Erlbaum.
- Stutsman, Rachel. 1981. *Merrill Palmer Scale of Mental Tests*. East Aurora, NY: Slosson Educational Publications.
- Terman, Lewis Madison & Maud Amanda Merrill. 1960. *Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M*. Boston, MA: Houghton-Mifflin.
- Waterson, Natalie. 1987. *Prosodic Phonology: The Theory and Application to Language Acquisition and Speech Processing*. Newcastle upon Tyne: Grevatt & Grevatt.

Karin Stromswold
Rutgers University
Department of Psychology &
Center for Cognitive Science
152 Frelinghuysen Road
Piscataway, NJ 08854
United States of America
karin@rucss.rutgers.edu

Aliza Lichtenstein
Rutgers University
Department of Psychology
152 Frelinghuysen Road
Piscataway, NJ 08854
United States of America
atbh22@gmail.com

