Chapter 4

Lexical and Frequency Effects in Phonological Development

In this chapter, I provide developmental evidence that lexical contrast is a crucial cue in the acquisition of phonological distinctions. I use the Providence Corpus (Demuth et al., 2006) to 1) quantify the extent to which lexical contrast cues are present in the parental input, and 2) demonstrate that lexical contrast cues are predictive of phonological acquisition as measured by child production accuracy. The results show that minimal pair cues are abundant in both parental and child speech. Moreover, minimal pairs are predictive of production accuracy on both the word and the phoneme levels, indicating that the structure of lexical contrast plays an important role in phonological acquisition.

4.1 Background

The idea that phonological representation emerges from generalization over lexical items has been around for a long time. The early work by Ferguson and Farwell (1975) outlines a sketch of a phonological acquisition model where "children learn words from others, construct their own phonologies, and gradually develop phonological awareness." Their conception of acquisition emphasizes "the primacy of lexical items" and "individual variation." Subsequent work on phonological acquisition led to more fully developed conceptual models as well as growing experimental evidence on the interaction between the lexical development and phonological acquisition. There are many ways that lexical or sub-lexical cues can influence the formation of phonological categories, and a detailed study of child production can shed light on the factors that play a role in phonological development.

4.1.1 Development of child production

Over the first year, the infants' vocalizations become increasingly speech-like. Before the onset of referential word use, infants often produce vocalic forms with stable meanings that do not correspond to any adult models, and these forms have been termed "protowords" or "quasi-words" by some researchers (Menn, 1976, 1983). Several studies suggest that protowords contain emergent phonological structures, and language-specific effects can be observed in early vocalizations (Menyuk et al., 1979; Stoel-Gammon and Cooper, 1984; Vihman and Miller, 1986).

For instance, adult listeners are able to discern the differences in vocal productions of 8- and 10-month-old infants acquiring French, Cantonese, and Arabic (de Boysson-Bardies et al., 1984). Additionally, differences in laryngeal articulation has been found in the babbling between infants acquiring English, Bai, and Arabic (Esling, 2012). The babbling vowel space is different for 10-month-old infants acquiring Algerian Arabic, Hong Kong Chinese, London English, and Parisian French (de Boysson-Bardies et al., 1989). The early babbling consonant repertoire appears to be similar across languages, with the majority of the consonants being stops, nasals, and the glide [h] (Locke, 1983), even when the ambient language does not contain /h/ as a phoneme (Vihman, 1992). However, there are significant cross-linguistic differences in consonant babbling that reflect the external linguistic input (de Boysson-Bardies and Vihman, 1991; Vihman et al., 1994). Moreover, Whalen et al. (1991) found significant differences in intonation patterns of reduplicated two- and three-syllable forms between English- and French-learning children between 6-12 months.

A general observation about infants' earliest words is that they tend to be surprisingly close to the adult targets. This observation has been ascribed to a process known as preselection, whereby infants "choose" to produce words that contain sounds which are more similar to their babbling repertoire and avoid words that contain more difficult sounds (e.g., Ferguson and Farwell, 1975; Fikkert and Levelt, 2008; Menn and Vihman, 2011). While this avoidance of difficult sounds has been attributed to potential metalinguistic knowledge about the sounds themselves (Menn, 1983), Vihman (1991) argues for an alternative explanation which she terms the *articulatory filter*. In Vihman's (1991) account, the infants use both bottom-up and top-down knowledge in their learning and production of words. In other words, infants selectively produce certain sound sequences as the result of the interaction and reinforcement between their own familiar articulatory sequences and similar sequences in the input. There is experimental evidence supporting this view (Vihman et al., 2014).

After the accurate initial production, there is usually a drop in the similarity to the adult targets in the child's production. This kind of U-shaped development of linguistic ability has been observed in other domains of first language acquisition, especially in inflectional morphology (Cazden, 1968; Marcus et al., 1992). A child may initially correctly produce irregular forms (e.g., fall → fell), but as they acquire the -ed past tense rule, they will often overgeneralize this rule to irregular forms and produce "falled", leading to a drop in overall accuracy. The drop in word form production accuracy has similarly been attributed to increasing phonological systematicity (e.g., Macken and Ferguson, 1983; Vihman and Velleman, 2000).

4.1.2 Lexical and sub-lexical factors in phonological development

The development of performance accuracy in child production has been the subject of many previous studies. A number of probabilistic and distributional lexical and sub-lexical cues have been found to have some effect on the phonological and word learning. Other factors such as vocabulary size and phonological neighborhood density have also been put forward as explanations for phonological development. Nevertheless, the observation of an effect from any of these factors does not mean that these factors are necessary for the development of a linguistic system. It is, therefore, important to evaluate the relative contributions of these factors in phonological development.

4.1.2.1 Frequency

With the rise in popularity of distributional learning as an explanation for language acquisition, the role of frequency in the linguistic input has received considerable attention in acquisition studies (Ellis, 2002). Experimental and corpus studies have found frequency effects on phonemic and lexical acquisition.

On the phoneme level, the vowel space of infant babbling tends to reflect phoneme frequencies of the ambient language (de Boysson-Bardies and Vihman, 1991). For instance, English learning children's coda production has been found to match English coda frequencies (Zamuner et al., 2005). Also, Ingram (1988a) suggests that English-learning children usually acquire /v/ late because /v/ is not a frequent phoneme in English, while children learning Swedish, Estonian, and Bulgarian acquire this sound earlier because they are more frequent in the lexicon. Additionally, Beckman and Edwards (2010) shows correlation between phoneme frequency and consonant production accuracy for English- and Cantonese-learning children. Edwards and Beckman (2008) looked at production accuracy of word-initial consonants for 2- and 3-year-olds and concluded that language acquisition is influenced by both universal constraints and language specific frequencies.

Frequency effects have also been observed on the level of word learning. While the total frequency of words are not predictive of child production of these words, the frequency of words uttered in isolation in the input is predictive of child word production at a later date (Brent and Siskind, 2001). Furthermore, there is evidence that more frequent words tend to be learned more accurately. Japanese learning children aged 1;5–2;1 are less likely to truncate words that are more frequent in the maternal input (Ota, 2006). Additionally, frequency interacts with positional salience in predicting the child's production of lexical items. For Italian-learning children aged 1;4–1;8, the occurrence of nouns in utterance-final positions in the input predicted the production of nouns, while the occurrence of verbs in utterance-initial positions is correlated with verb production (Longobardi et al., 2015).

Overall, it appears that frequency has some effect on the acquisition of phoneme and word production. However, the effect is not always straightforward and the interaction with other factors sometimes needs to be considered.

4.1.2.2 Phonotactic probability

In phonological analysis, phonotactics refers to the restrictions on the combinations of phonemes. For instance, /kn-, pt-, ps-, sr-/ are not permissible onset clusters for English, but /sp-, tr-, gl-/ are. An English speaker might judge a made-up word like /srum/ to not be a possible word of English, while /spum/ could be a word in English. Phonotactics is part of the speaker's implicit knowledge about the phonology of their language and part of their linguistic competence. Along with the rising interest in applying statistical learning to various acquisition problems, phonotactic probability has been proposed to play a role in lexical acquisition. Unlike phonotactics which describes patterns of possible and impossible sound sequences in discrete terms, phonotactic probability quantifies the likelihood of sound combinations through the frequencies of the co-occurrences of sound sequences in a language. Higher phonotactic ability may facilitate speech processing on the sublexical level (Vitevitch and Luce, 1998, 1999).

There is some evidence that phonotactic probability influences phonological and lexical acquisition, but the results are inconclusive. Infants show preference for sound sequences with higher phonotactic probability at 9 months (Jusczyk et al., 1994), and older children (aged 3;2-6;3) learn words with higher phonotactic probability with more ease (Storkel, 2001). Another study showed that children aged 3;2-8;10 can repeat non-words with frequent phoneme sequences with higher accuracy (Edwards et al., 2004). On the other hand, 4-year-olds are more accurate at learning words with rare sound sequences (Storkel and Lee, 2011). For older children, no effect of phonotactic probability was found in the learning of nonwords by 7-year-olds, while 10- and 13-year-old children had an easier time learning high probability non-words (Storkel and Rogers, 2000).

4.1.2.3 Vocabulary size

The effect of vocabulary size on word learning and phonological tasks is more consistent. Vocabulary size has been found to predict children's performance on word learning and phonological tasks. At 14 months, children with larger vocabulary find it easier to learn minimally contrasting non-words (Werker et al., 2002). For continuous speech processing at 18 and 21 month, children with larger productive vocabulary were more accurate and faster at responding to familiar words (Fernald et al., 2001). At ages 3-5, children with larger vocabulary tend to be more accurate at non-word repetition (Metsala, 1999). Although Edwards et al. (2004) found effects of phonotactic probability, children with larger vocabularies showed less frequency effects. In a subsequent study that included children with specific language impairment, Munson et al. (2005b) found that these children performed similarly as their vocabulary size matched peers, and overall vocabulary size is the best predictor of non-word repetition accuracy. These results suggest that greater vocabulary sizes provide the learner with the opportunity to generalize phonological contrasts over more words, resulting in better phonological awareness and word learning abilities.

4.1.2.4 Minimal pairs and related concepts

In traditional phonological analysis, minimal pairs are used to establish the phonemic inventory of a language. Minimal pairs refer to two words that are distinct in meaning and differ by one phonological unit. As this definition stands, phonological neighborhood density and functional load are very similar concepts, but they tend to be used in different contexts.

The phonological neighbors of a word is defined as the set of words that can be obtained by adding, subtracting, or substituting one segment of this word (Luce, 1986). Calculating the number of minimal pairs and the neighborhood density is methodologically similar. The main difference between them is that in practice, linguists tend to restrict minimal pair analysis to words of the same lengths. For instance, pairs like "cold" and "gold" can be used to establish the contrastiveness of /k/ and /g/, while pairs like "old" vs. "cold" and "old" vs "gold" are rarely used in such analyses even though both pairs technically differ by one segment. Minimal pairs and phonological neighborhoods differ mostly in how they are used rather than how they are identified and calculated. Phonological neighborhoods are commonly used in models and experiments in speech processing, while phonologist and phoneticians use minimal pairs to identify and study properties of specific linguistic contrasts.

Another related concept is functional load, which tends to be used in work on sound change as a measure of the importance of a phonological contrast in the lexicon (Martinet, 1952; Wedel et al., 2013). If a phoneme is used to distinguish many words, it has a high functional load. Methodologically, a phoneme's functional load is often quantified as the number of minimal pairs it distinguishes. In a study on mergers, Wedel et al. (2013) shows that phonemes with lower functional load (i.e., fewer minimal pairs) are more likely to merge than high functional load phonemes.

Even though minimal pairs have been used in phonological analysis for a very long time, there has been few studies on first language acquisition that explicitly look at the interaction between minimal pairs and acquisition results. In second language acquisition, however, minimal pair training has been found to improve both perception and production of second-language phonemic contrasts (Logan et al., 1991; Bradlow et al., 1997; Wang et al., 1999), and minimal pair training results in better discrimination abilities than perceptual training alone (Hayes-Harb, 2007). However, given the differences between first and second language acquisition, these findings do not imply that minimal pairs are also predictive first language acquisition outcomes. The study in this chapter is intended to fill the gap in the general lack of direct study on the role of minimal pairs in first language acquisition.

4.1.3 Quantifying linguistic competence from linguistic performance

Since the primary evidence for this study comes from a corpus of child production data, it is necessary to carefully consider the relationship between linguistic performance and linguistic competence in drawing conclusions from such an analysis. Phonological competence, like any other level of linguistic knowledge, is part of the speaker's I-language, i.e., the internal mental representation of their language (Chomsky, 1986). The obvious challenge to the study of I-language is that barring some exceptional advances in neuroliguistics, it cannot be directly observed. With mature speakers, it is possible to indirectly study the nature of their I-language experimentally or through linguistic tasks such as grammaticality judgments. However, studies aimed at understanding children's developing I-language are limited by practical concerns when working with infants and young children.

The limits of experimental data on early perception has been reviewed in the previous chapters. Essentially, the discrepancy between perceptual discrimination and word learning results show that perceptual discrimination should not be used as the sole evidence for the existence of phonological distinctions in the internal grammatical representation of the child. These studies nevertheless reveal something about the units of perception in early language learning, which are the necessary precursors for adult-like phonological units. There are two interpretations for the disparity between young children's phonetic and phonological performance: 1) Young children are phonologically competent but their performance suffers from non-linguistic factors like cognitive processing demands and motor control skills, and 2) young children have not developed complete phonological competence yet, and hence the poor performance.

If perceptual results offer limited but not conclusive indications of the state of development of a child's language, what about production? This chapter uses production accuracy from the Providence Corpus as a proxy of phonological competence. There are definite concerns with this approach since linguistic competence and linguistic performance are not equivalent, especially for children with developing motor control skills. Observations like the fish-phenomenon calls into question the validity of using child production to measure phonological knowledge. First documented by (Berko and Brown, 1960), the fish-phenomenon describes a situation in which a child misses pronounces a word [fɪs] for "fish", but rejects the pronunciation by an adult when it is repeated back to the child (Smith et al., 1973).

The wrong production of a word or phoneme can be the result of either linguistic performance or competence: It is possible that the child has not arrived at an adult-like representation or has trouble executing the specific sequence of articulatory gestures. However, it would be unreasonable to attribute *consistently accurate* production of word forms and phonemes to mere performance. As discussed earlier, child production tends to follow a U-shaped curve in terms of target-like accuracy. Before reaching phonological competence, we should expect to see variation in production accuracy as the result of the systematicization of the phonological system. Sporadically accurate production is not informative about the child's linguistic competence, only that phonological reorganization is taking place. However, if the production data shows consistent accuracy towards the adult targets, it is possible to draw conclusions about the child's linguistic competence.

4.2 The Providence Corpus

The Providence Corpus consists of recordings and transcripts of spontaneous mother-child interactions for six monolingual English-learning children (Demuth et al., 2006). The recordings made by Katherine Demuth and her research assistants at the Child Language Lab at Brown University in Providence, RI. Data collection occurred between the years 2002-2005, with a total of 364 hours of recorded video and audio data. The recordings were carried out every two weeks and they usually occurred at the homes of the subjects. Each recording session was approximately one hour long. The Providence Corpus was chosen for this study because 1) it contains naturalistic data of both the parental input and the child production, 2) the children recorded in the corpus were in the age range (1-3 years) of interest for phonological acquisition, 3) there is sufficient data for each child, and 4) this corpus has been orthographically transcribed for both parents and children, and 5) phonetic transcription made by trained transcribers is available for all the children. A summary of the children from the corpus is provided in Table 4.1.

To conduct phonological analysis of the Providence Corpus, the existing transcription first needs to be processed into a form suitable for the goals of this study. Different types of transcriptions are available for the parents and the children. The parental speech has been transcribed with standard orthography and marked for part of speech, morphological stem, and position in the utterance. The corpus does not provide phonemic or phonetic transcrip-

Name	Age Range	Sessions	Sex
Alex	1;04.28-3;05.16	51	M
Ethan	0;11.04-2;11.01	50	M
Lily	1;01.02-4;00.02	80	F
Naima	0;11.27-3;10.10	88	F
Violet	1;02.00-3;11.24	51	F
William	1;04.12 - 3;04.18	44	M

Table 4.1: Summary of the information about the children and recordings in the Providence Corpus.

tions of the parental speech. The children's speech has been transcribed for orthography, actual produced phonetic forms, target phonemes, part of speech, morphological stem, and position in the utterance.

4.2.1 Processing of parental speech

For the parental speech, the orthography, phonemic transcription, morphological stem, and part of speech were obtained for analysis in this chapter. Because the parental speech was not transcribed phonemically in the corpus, it is necessary to first obtain phonemic transcriptions before any phonological analysis can occur. Phonemic transcriptions were applied to the orthographic forms of the parental speech from the CMU Pronouncing Dictionary, which covered the majority of the orthographic transcriptions in the corpus. There was a number of frequent words whose phonemic transcriptions were not available from the CMU Dictionary. For relatively more frequent words (>20 occurrences in the entire corpus) such as content words (e.g., lollie, scrumptious, hummus), diminutive forms (blankie, nursie, piggie), proper names (Naima, Eeyore, Mufasa), an additional dictionary was created where the phonemic transcriptions were manually entered. Other words were excluded from this analysis. The excluded words include unintelligible speech (e.g., xxx, www), interjections (e.g., uhoh, uhhuh, tadah), highly reduced forms (e.g., dya, whaddya), and less frequent words (<20 occurrences in the entire corpus). In total, 4911 word types were excluded. Even though this seems to be a large number, most of the words were of the types described above, and 2739 only occurred once in the entire corpus. The transcriptions were converted from Arpabet in the CMU dictionary to the IPA so that it easier to compare with the IPA transcriptions of child production.

Stem-level transcription is available from the corpus. For example, a word with the orthographic representation of "hats" would be transcribed on the stem level as "hat-pl". From stem-level transcription, the root of each word was found by removing the suffixes. Thus, for "hat-pl", the root is simply "hat". Additionally, the corpus transcription has detailed part of speech (POS) tags. For example, the pronoun category "pro" is further divided into subcategories, such as demonstratives ("pro:dem"), relative pronouns ("pro:rel"), and indefinite pronouns ("pro:indef"). From these labels, the larger, more basic POS categories (e.g., just "pro") were obtained for each word.

After each word was processed, exclusions of certain words were applied as follows. Words without POS tags were excluded, and these were almost exclusively space fillers (e.g., um, uh, ooh, aw). Some words labeled with the POS tag "co" (communicator) were also excluded; these include fillers such as "mhm", "huh", "wah". Communicator words like "yeah", "okay", and "please" were kept. Further exclusions based of POS tags include "sing" (8 tokens for when the parent was singing), "none" (2 tokens), "chi" (120 tokens of child-invented forms like "dede", "wa", "balog"), "wplay" (157 tokens, e.g., "phooey", "snip"), "neo" (10 tokens, e.g., "tso", "skinks").

4.2.2 Processing of child production

For each word in the child speech, the child's actual production, target phonemes, morphological stem, and POS categories were obtained directly from the corpus. The transcriptions of each child's actual productions and the target forms were available from the corpus. Words were excluded if they did not include a target or actual transcription from the corpus. Most of the exclusions were unintelligible forms (e.g., xxx, yyy), and only 191 word types were excluded in total. Out of the 191 words, 119 only occurred once, and 175 had frequencies of 5 or less. Although both target phonemic forms and actual phonetic productions were transcribed in the IPA, further processing of the transcriptions was carried out to eliminate internal inconsistencies. For example, /ɚ/ was transcribed in a number of different ways (e.g., ʌɹ, əɹ, ər, ɝ), and these were all standardized to /ɚ/, and affricates were transcribed both as digraphs (e.g., dʒ, tʃ) and as single letters (e.g., ʤ, ʧ), and these were consolidated as single letters. Moreover, diphthongs were represented as single units in the analysis conducted in this chapter.

After the processing of phonemic and phonetic transcriptions, the morphological stem of each word and the basic POS categories were derived in the same way as for the parental speech. Similar to the adult speech, words without POS markers were excluded. These were mostly unintelligible babbling transcribed as "xxx" or "yyy". Several other POS categories excluded from analysis include "fam" (with only one word "zoob"), "L" (116 tokens with no orthographic transcription at all), "neo" (15 tokens of nonwords like "vrap", "tso"), "cm" (557 tokens of words such as "uh", "um", "sssh"), "wplay" (306 tokens, e.g. "zub", "pommy"). Words with "*" marked as the model production or "*" as the actual production were also excluded, and these were for the most part fillers like "um" or sounds like "ss", "wa". An additional 1,692 words excluded are words whose orthographic representations are not found in parental speech. Most of these are low frequency forms. About half of these words (814) only have one occurrence in the entire corpus. These words include more communicator type words like "uhuh", "uhhuh", "tadah", and "mkay". Some of these other words in this excluded group demonstrate overgeneralization by the child, like "falled", and some of the words are the result of the transcriber attempting to transcribe phonetically with orthography; for example, data from one child had "goldipocks", "goldidocks", "goldisocks", and "goldiblocks", each with frequency of 1 or 2.

Some words were missing orthographic representations, but the transcription for stem, model production, and actual production were all available. There are 662 of these items, and most of them (judging from the phonetic transcriptions) are reduced forms of common words, such as "about", "around", "because". Also, in a few cases, the orthography appears to be a phonetic transcription. For instance, the word whose orthography transcribed as "ta" has the stem "to" and the POS "inf", but its actual production was marked as [tɑ]. Some orthographic transcriptions are misspelled, like "gree" for "green". Since most of these words do not decompose further, the stem is used as a substitute for orthography.

4.2.3 Descriptive statistics of the processed data

Table 4.2 summarizes the post-processing data used for analysis in the rest of this chapter, including the total word counts for each of the participants, average word counts per session, the total number of unique orthographic words, and the number of unique stems. There appears to be individual variation in how much each parent and child talked, but some of the difference comes from the fact that some children were recorded until an older age. Naima and Lily have more total sessions because they were recorded weekly rather than every other week. Recordings of Lily, Naima, and Violet were carried out monthly between 3-4 years of age, while the other children were only recorded until they were around 3 years old.

Mothers					
child	sessions	word count	words per session	unique words	unique stems
Alex	51	144518	2833.69	3876	2555
Ethan	50	158607	3172.14	4518	2754
Naima	88	301420	3425.23	6240	3821
Lily	80	340372	4254.65	8370	5111
Violet	51	125525	2461.27	5265	3363
William	42	127042	3024.81	3539	2339
Children					
child	sessions	word count	words per session	unique words	unique stems
Alex	51	40102	786.31	1690	1279
Ethan	50	32057	641.14	2355	1705
Naima	88	112460	1277.95	3745	2446
Lily	79	77064	975.49	3081	2143
Violet	48	29226	608.88	1945	1403
William	44	34201	777.30	1658	1228

Table 4.2: Descriptive statistics of the data used for the analysis in this chapter.

4.3 Quantifying minimal pair cues in first language acquisition

This section of this chapter has a straightforward goal: to quantify the amount of minimal pair cues that exist in child-directed speech as well as child speech. To do so, I provide minimal pair counts with different word exclusion criteria, and I also quantify the amount of minimal pair cues per session and between pairs of phonemes. This section of the chapter is meant to be purely descriptive, and the implications of minimal pair cues will be elaborated in the following section and the general discussion.

4.3.1 Methods

Although the definition of a minimal pair is quite straightforward, in practice, counting minimal pairs is a little bit more complicated especially in the context of language acquisition. The most straightforward examples of minimal pairs are words like "bad" vs. "bed", which differ by one vowel (/bæd/ vs. /bɛd/), and "pat" vs. "bat", which differ by one consonant (/pæt/ vs. /bæt/). However, in conversational speech, many words are inflected. Words like "hide" vs. "hid" differ by one phoneme, and their meanings differ with respect to tense. Being able to use "hide" and "hid" to learn the /ɑɪ/ vs. /ɪ/ distinction requires the child to have some knowledge that tense is a dimension of meaning difference, and young children may not have acquired this distinction early on. Additionally, the meanings of functional words are often rather abstract. For example, is it possible for the child to learn the /ð/ vs. /f/ distinction from "that" vs. "fat"? Additionally, the phoneme /ð/ rarely occurs in content words but is highly frequent in functional words. On the other hand, the abstractness of function words may not be a huge hurdle for learning phonemic categories from them since function words tend to occur in very different syntactic contexts than content words. If the child notices the word "that" consistently occurs in different positions than "fat", perhaps this distinction alone will enable the child to know that there is some difference between "that" vs. "fat".

In order to provide a full picture of minimal pair cues in parent-child interactions, minimal pairs were counted with different degrees of word exclusion. The word exclusions were meant to account for different scenarios where child may or may not be able to access certain word categories for phonological learning. In addition, phonological neighbors are also counted for a comparison with minimal pair measures. I present data from:

1. All the transcribed words from the processed corpus

2. Content words only (nouns, adjectives, verbs, and adverbs)

3. Monomorphemic words (content and functional)

4. Monomorphemic content words (nouns, adjectives, verbs, and adverbs) only

5. Frequent monomorphemic content words ($n > 10$)

6. Phonological neighbors of monomorphemic content words

To count minimal pairs, pairwise string comparisons were carried out between unique phonemic transcription types for all the parents, as well as unique target types for all the children. Two words were determined to be a minimal pair if they were the same length and differed by only one phoneme. For phonological neighbors, two words that differed in length by one are also included in addition to equal-length words. Two words were determined to be phonological neighbors if they differ by one phoneme through substitution, deletion, or addition. Phonological neighbors were only calculated for the monomorphemic content subset of the words.

Like phonological neighbor calculations, the subsequent analysis all used the restrictive monomorphemic content words to quantify minimal pair cues. This set of words were chosen to provide conservative estimates of what lexical contrast cues the child could use. Two measures computed from monomorphemic content words include the average number of minimal pairs each individual child heard and produced in each one hour recording session, as well as the numbers of minimal pairs for each pair of phonemes, both in parental speech and in child production.

4.3.2 Results

4.3.2.1 The effect of word exclusion criteria and method of counting

Figure 4.1 plots the overall numbers of minimal pairs counts for each phoneme according to different exclusion criteria delineated above. Unsurprisingly, the number of unique minimal pair counts decreased as the word exclusion criteria became more and more restrictive. When all word forms were included in minimal pair count, parental speech included 32,648 unique minimal pairs in total, and the children produced 14,077 minimal pairs all combined. The total numbers decreased as word exclusions were applied. Counting only content words, the parents produced 27,919 minimal pairs, and the children produced 11,364. When only monomorphemic words were counted, parental speech had 16,191 minimal pairs, and the children had 8455. When only monomorphemic content words were considered, the counts were 9416 for the parents and 5059 for the children. For more frequent ($n > 10$) monomorphemic content words, the numbers decreased to 3647 words for the parents and 1631 words for the children.

While excluding words by type (content vs. functional), morphological complexity, and frequency reduced the number of minimal pairs, the relative numbers of minimal pairs between phonemes remain roughly the same. For instance, /d/, /k/, /t/ have more minimal pairs relative to other phonemes when all words were used to count minimal pairs, and this trend remained when functional words were excluded, when only root forms were used, when only root forms of content words were used, and when a frequency threshold was applied to the root forms of content words. Phonological neighborhood counts yielded likewise similar results overall; all the counts are slightly above monomorphemic content counts. The major difference is that phonological neighborhood counts included the correspondence of phonemes to null elements.

The observation that different exclusion criteria and counting methods result in similar trends is confirmed by pairwise correlations between the different minimal pair counts and the phonological neighbors count. Table 4.3 is a correlation matrix of all six measures of minimal pair for both the parents and the children, and it shows the correlation is high between different word exclusion conditions, minimal pair and phonological neighborhood counts, and parental and child counts. Because of this trend, different ways of minimal pair counting should have similar predictive power, as long as the method of counting and word type exclusion is systematic.

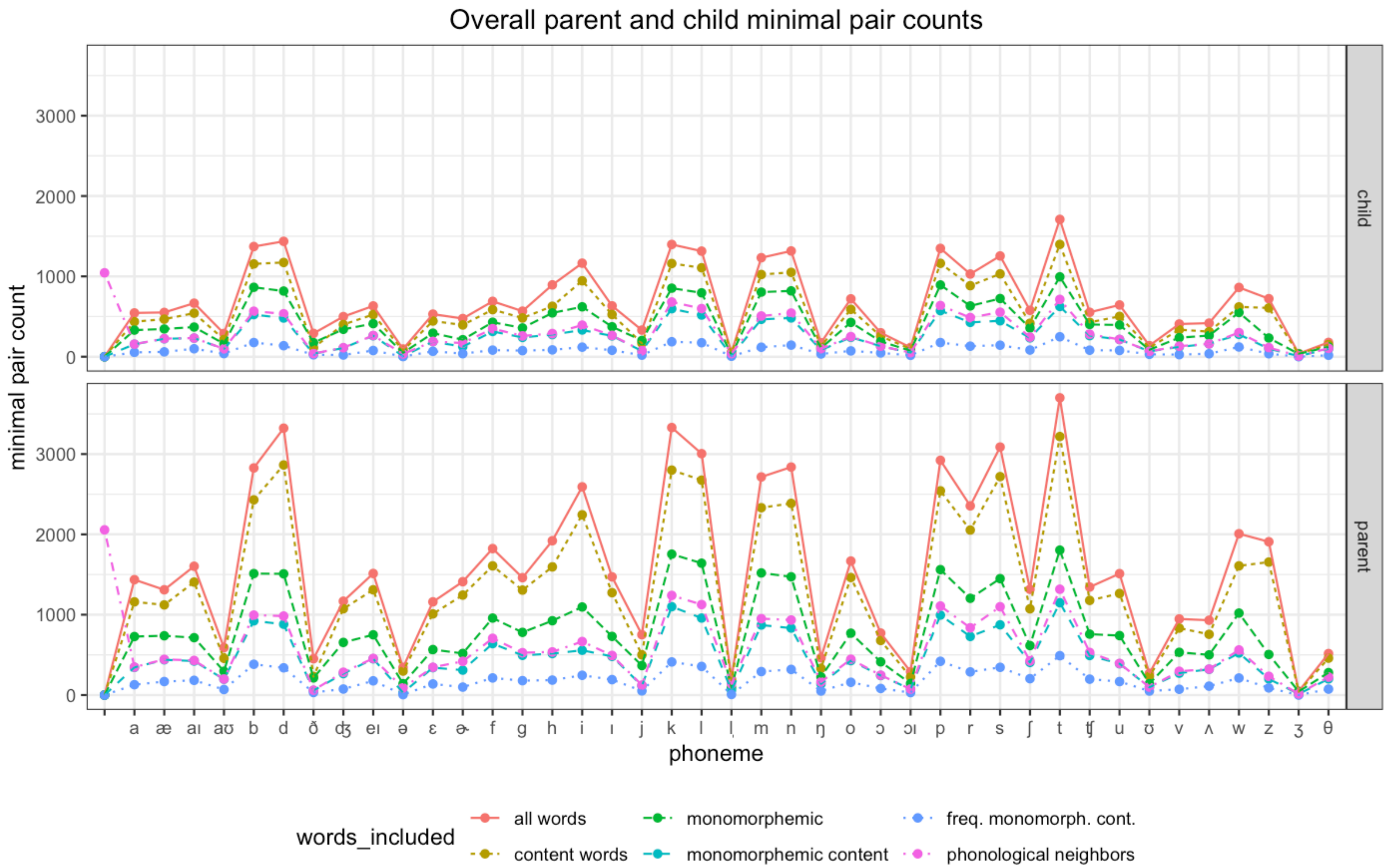


Figure 4.1: Comparison of minimal pair counts with different exclusion criteria.

	P-A	P-C	P-M	P-MC	P-FMC	P-PN	C-A	C-C	C-M	C-MC	C-FMC	C-PN
P-A	1.00	1.00	0.98	0.96	0.96	0.96	0.99	0.99	0.97	0.95	0.94	0.96
P-C	1.00	1.00	0.98	0.96	0.95	0.96	0.99	0.99	0.97	0.95	0.93	0.96
P-M	0.98	0.98	1.00	0.99	0.97	0.97	0.98	0.98	0.99	0.98	0.95	0.98
P-MC	0.96	0.96	0.99	1.00	0.99	0.99	0.96	0.97	0.98	0.99	0.96	0.99
P-FMC	0.96	0.95	0.97	0.99	1.00	1.00	0.96	0.97	0.97	0.99	0.98	0.99
P-PN	0.96	0.96	0.97	0.99	1.00	1.00	0.96	0.97	0.97	0.99	0.98	0.99
C-A	0.99	0.99	0.98	0.96	0.96	0.96	1.00	1.00	0.99	0.96	0.95	0.96
C-C	0.99	0.99	0.98	0.97	0.97	0.97	1.00	1.00	0.98	0.97	0.95	0.97
C-M	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98	1.00	0.98	0.96	0.98
C-MC	0.95	0.95	0.98	0.99	0.99	0.99	0.96	0.97	0.98	1.00	0.97	1.00
C-FMC	0.94	0.93	0.95	0.96	0.98	0.98	0.95	0.95	0.96	0.97	1.00	0.97
C-PN	0.96	0.96	0.98	0.99	0.99	0.99	0.96	0.97	0.98	1.00	0.97	1.00

Table 4.3: Correlations between various minimal count measures and phonological neighbor counts for all the phonemes. Labels are abbreviated for space: "P-" = parental counts, and "C-" = child counts. A = all words, C = content words only, M = monomorphemic words, MC = monomorphemic content words, FMC = frequent monomorphemic content words, PN = phonological neighbors.

4.3.2.2 Minimal pair cues in natural speech

	All Words				Monomorphemic Content			
	Parents		Children		Parents		Children	
child	mean	sd	mean	sd	mean	sd	mean	sd
Alex	504.06	117.70	95.70	81.84	119.49	34.04	26.17	19.37
Ethan	578.24	205.11	121.28	76.09	145.52	51.36	29.57	17.09
Lily	810.08	204.81	177.13	101.32	222.69	77.21	38.72	22.63
Naima	545.92	139.15	168.40	105.79	141.05	46.78	37.80	24.75
Violet	471.75	174.33	116.76	69.02	120.06	64.59	22.12	15.50
William	483.71	228.76	112.30	99.90	110.98	59.99	21.78	23.18

Table 4.4: Means and standard deviations of minimal pair counts for the parents and children.

The above analysis counted minimal pair cues for each participant in the corpus across all their speech data. However, in each hourly recording session, is the child likely to hear many words that are minimal pairs? To answer this question, Individual analysis was carried out for both the parents and children for each session, and the numbers of minimal pairs were quantified per session. Table 4.4 shows the average counts per session using all the words and when only monomorphemic content words were considered. The distributions of minimal pair counts for monomorphemic content words for all the sessions are plotted in Figure 4.2. It appears that minimal pairs are a common occurrence in natural speech. Of

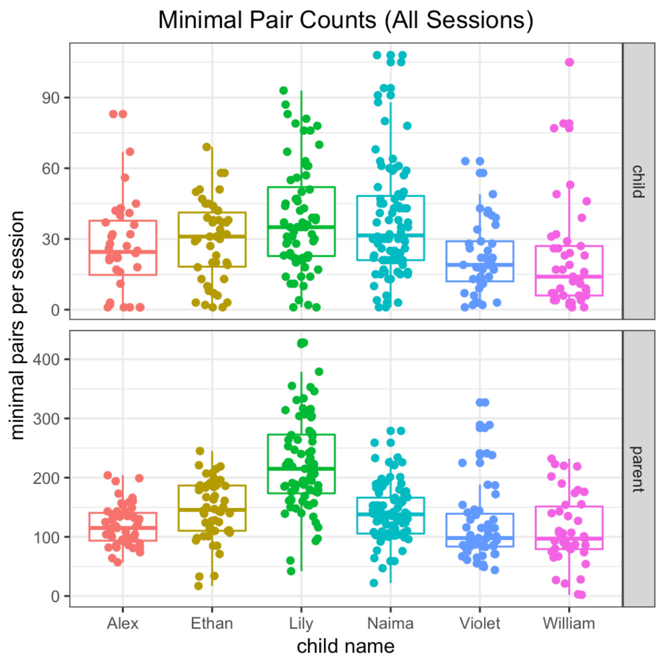
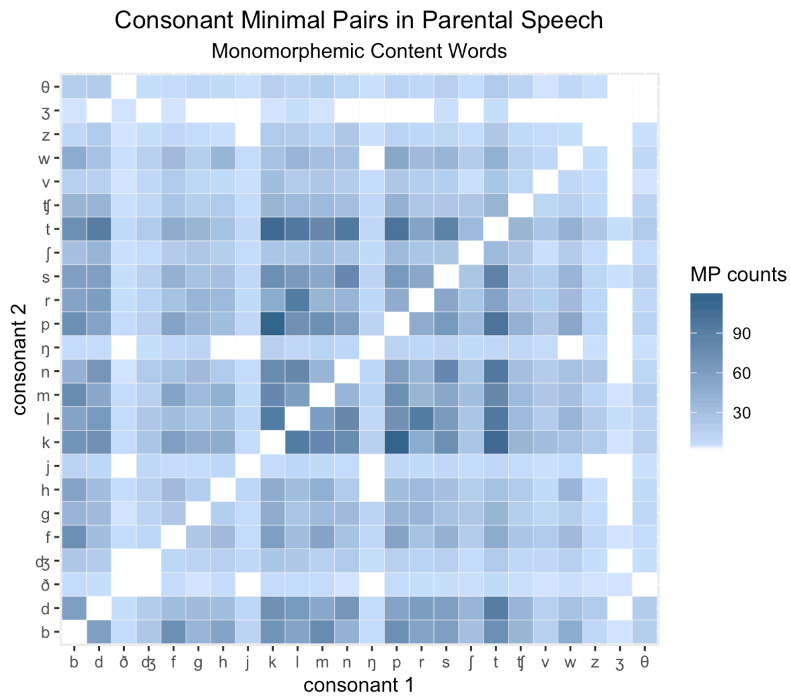


Figure 4.2: Boxplot of the number of minimal pairs in parental and child production for all the sessions.

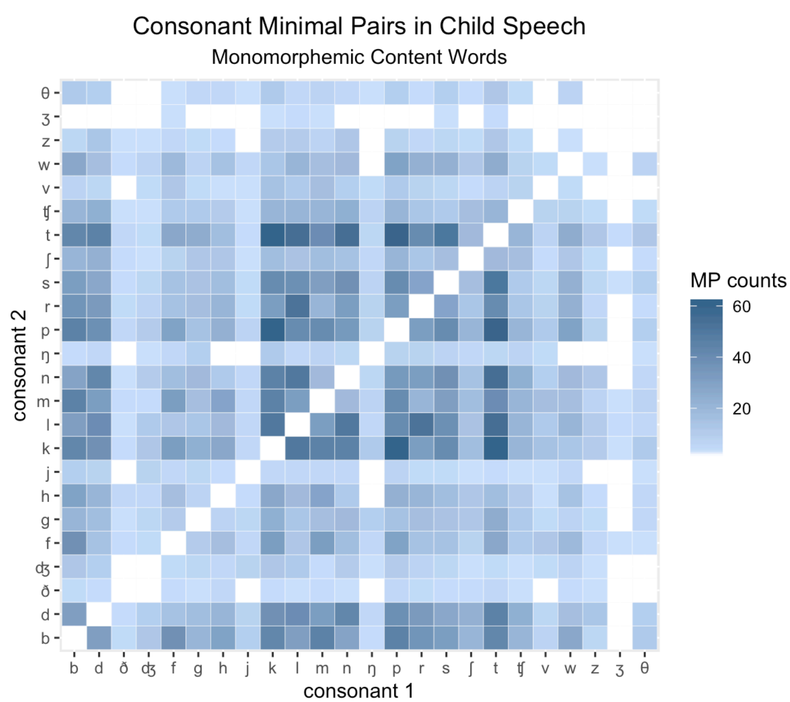
course, inclusion of more words results in high counts, but even when only monomorphemic words were considered, there were well overall 100 minimal pairs per hour of parental speech. It remains a question how much of this information the child can use in learning, but so far, the existence of minimal pair cues has been well demonstrated.

4.3.2.3 Minimal pair cues for pairs of phonemes

Although it is clear that minimal pair cues are abundant in natural speech, the question remains whether minimal contrast between phonemes can be learned from these minimal pairs. I present the results on minimal pair counts for each pairs of phonemes in parental and child speech. I include here only measures from the most restricted data set – the counts from only monomorphemic content words. Figure 4.3a and Figure 4.3b are heatmaps visualizing minimal pair counts for each pair of consonant phonemes for parents and children respectively. Most pairs of consonant phonemes are well represented. For both parents and children, the phoneme /ʒ/ has the fewest number of minimal pairs with other consonants. This is expected since /ʒ/ appears in relatively few words and typically restricted to word-

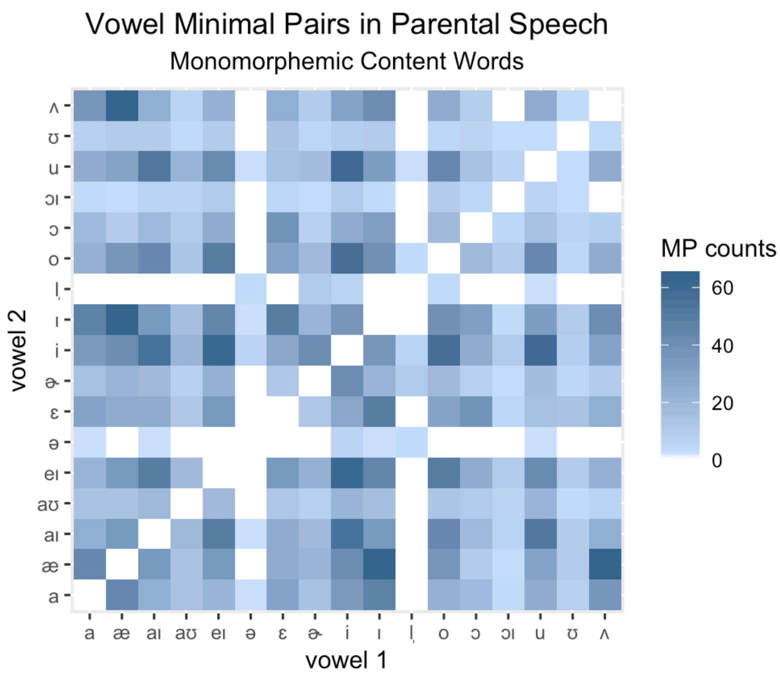


(a) Consonant minimal pair counts in parental speech.

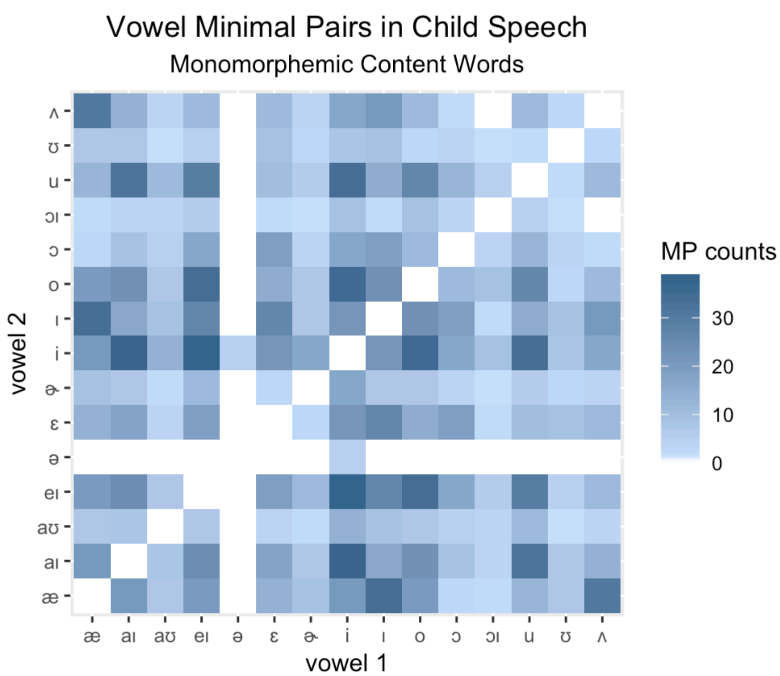


(b) Consonant minimal pair counts in child speech.

Figure 4.3: Unique minimal pair counts for each pair of consonant phonemes from both parental and child speech for monomorphemic content words.



(a) Vowel minimal pair counts in parental speech.



(b) Vowel minimal pair counts in child speech.

Figure 4.4: Unique minimal pair counts for each pair of vowel phonemes from both parental and child speech for monomorphemic content words.

medial contexts. This should not be problematic for acquisition, since the existence of any minimal pair with /ʒ/ should be enough evidence that a contrast needs to be learned. Similarly, vowels are well represented by minimal pairs in parent (Figure 4.4a) and child speech (Figure 4.4b). Of course, when more word categories are included, the minimal pair counts increase for all pairs of consonants and vowels.

4.3.2.4 Word frequency and minimal pairs

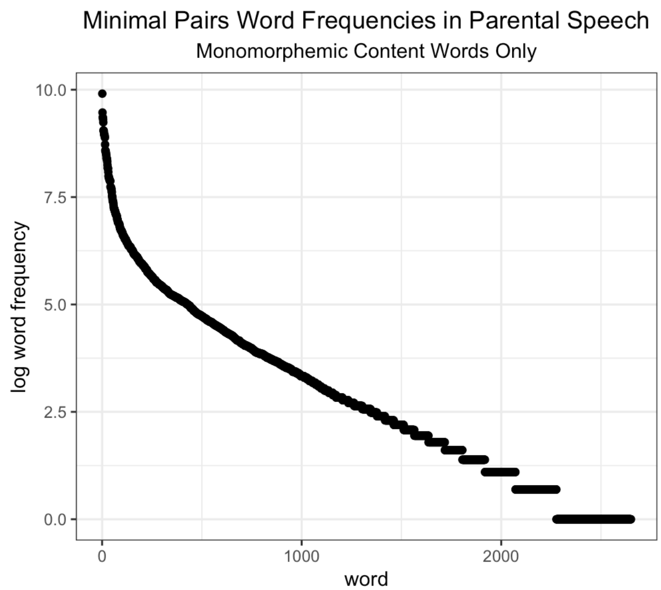


Figure 4.5: The frequencies of words included in the minimal pair count for parental speech. Only monomorphemic content words are included in this frequency count.

So far, I have demonstrated the abundance of minimal pair cues in parental speech and in child production. Nevertheless, the acquisition of phonological categories based on minimal pairs might also be influenced by how frequent these words occur. If the minimal pair of a highly frequent word seldom occurs, the child is likely to acquire the less frequent word and make use of the minimal pair contrast for phonological learning. The log word frequency is plotted in Figure 4.5. Out of 2650 unique words used in the minimal count, 374 occurred only once in the corpus. These words are not necessarily rare words in English, but they are perhaps less common in child-direct speech. Some examples of words that occurred only once include “mulch”, “caution”, “elite”, “feeble”, and the like. However, a large number of minimal pairs are highly frequent. Of the monomorphemic content words from parental speech, 1297 pairs are words that have frequency counts above 50, from which 494 unique phonemic contrasts are represented.

4.4 An evaluation of factors in phonological acquisition

In the above section, I have demonstrated that minimal pairs are abundant in parental and child speech. However, the existence of minimal pair cues does not imply that the child can make use of it in learning phonology. In this section, I provide evidence that minimal pairs have an effect on phonological acquisition. I quantify child production accuracy on the word and phoneme levels and compare the effectiveness of minimal pairs, frequency, and phonotactic probability in predicting production accuracy.

4.4.1 Word level production

This section investigates word level production accuracy and which factors are predictive of word production accuracy. Production accuracy was quantified for each child, and the predictive power of word length, minimal pair counts, and word frequency were examined.

4.4.1.1 Methods

To quantify production accuracy for each child on the word level, the phonetic transcription of the child's actual production was compared to the target phonemic forms of each word. Two accuracy measures were calculated: categorical and gradient. For categorical accuracy, if the actual form is different from the target form by any segment, the production of the word is marked inaccurate; if the actual production matched the target form exactly, the word was marked as accurate. For gradient accuracy, the number of correctly produced segments is divided by the total number of segments of a word. For example, if a child produces [d.ʌ.d.i] for the target form "doggie" /d.ɑ.g.i/, this production would be categorically inaccurate but have a gradient accuracy of $2/4 = 0.5$. The overall accuracy of each word is calculated as the average accuracy over all productions of this word for both categorical accuracy and gradient accuracy.

Minimal pair counts from the previous section are used in visualizations and statistical modeling in this section. In addition, orthographic word frequencies are calculated for each parent and child. Biphone phonotactic probabilities are calculated using Phonological CorpusTools (Hall et al., 2017) with log token frequencies as in the algorithm originally outlined in Vitevitch and Luce (2004). Linear regressions were estimated to test the effects of word length, minimal pair counts, word frequency, and phonotactic probability on production accuracy on the word level (word accuracy$\sim$word length + minimal pairs + phonotactic probability + frequency). Because each measure was independently calculated for each child, separate linear regression models are estimate created for each child.

4.4.1.2 Results

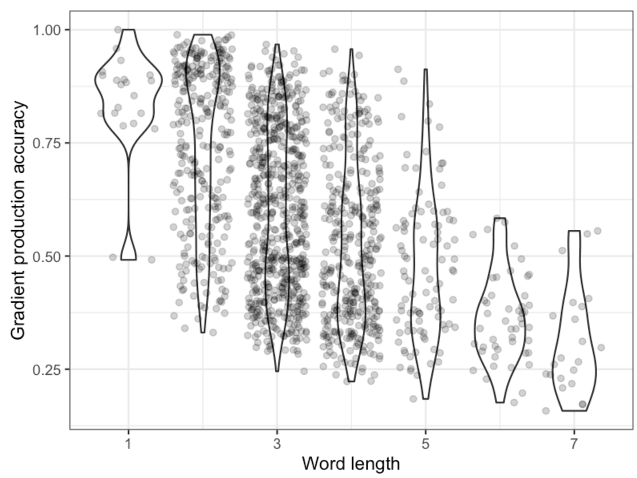
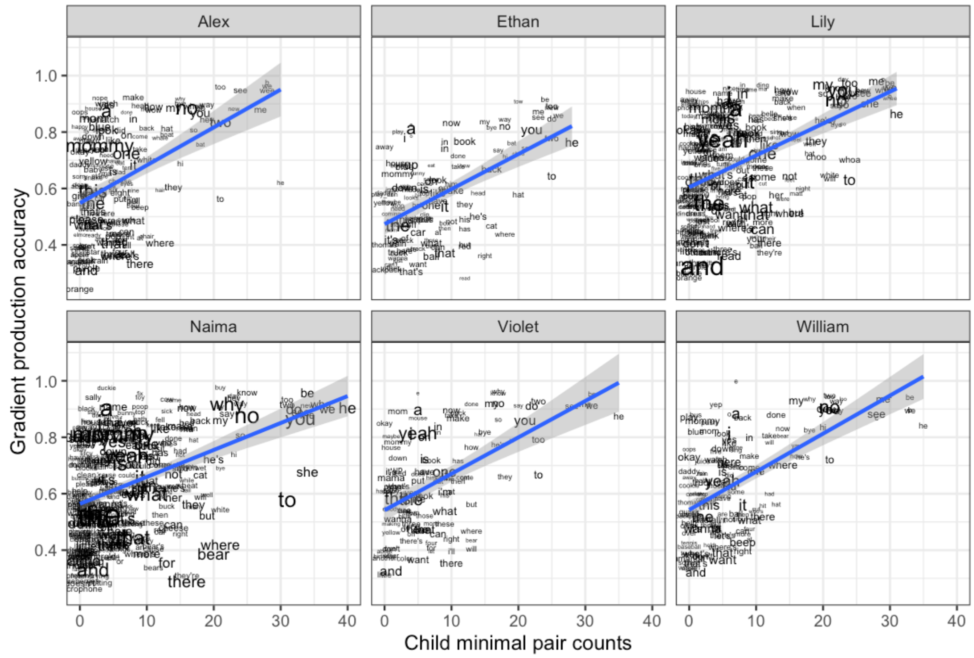
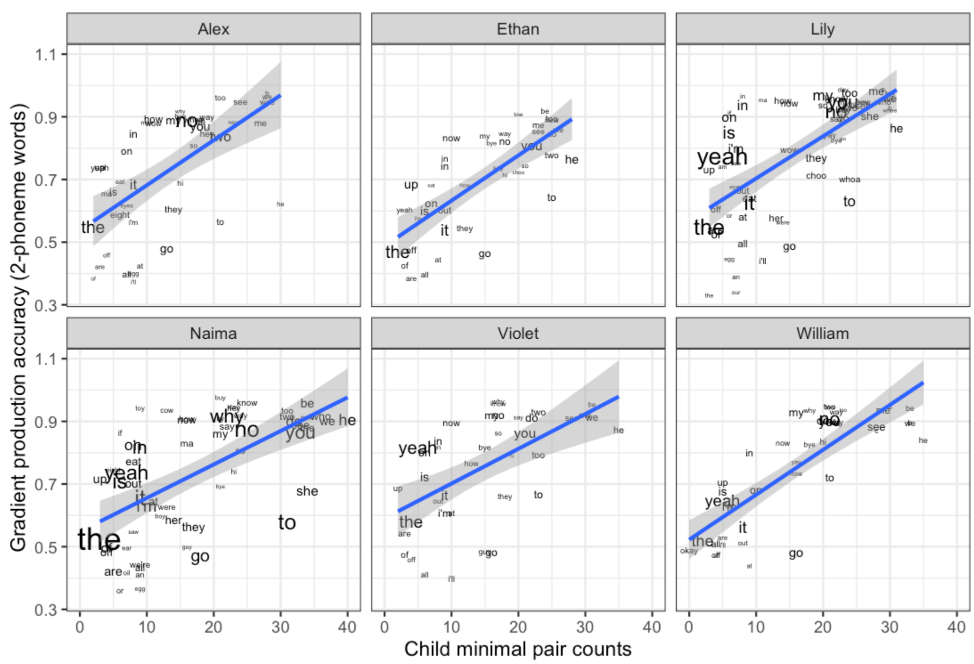


Figure 4.6: Word length and gradient production accuracy. Shorter words tend to be produced more accurately.

Word length. Figure 4.6 shows the effect of word length on production accuracy. Word length is quantified by the number of phonemes in the target forms of each word, and production accuracy is pooled from all six children. Unsurprisingly, shorter words tend to be more accurately produced than longer words. There is a wide range of production accuracy for all lengths of words, but on average, accuracy drops as word length becomes



(a) Child minimal pair counts and gradient word production accuracy for the six children for all words. There is an overall trend that more minimal pairs indicate better production accuracy.



(b) Child minimal pair counts and gradient word production accuracy for the six children for all 2-phoneme words. The trend that more minimal pairs indicate better production accuracy remains.

Figure 4.7: Minimal pairs and word production accuracy.

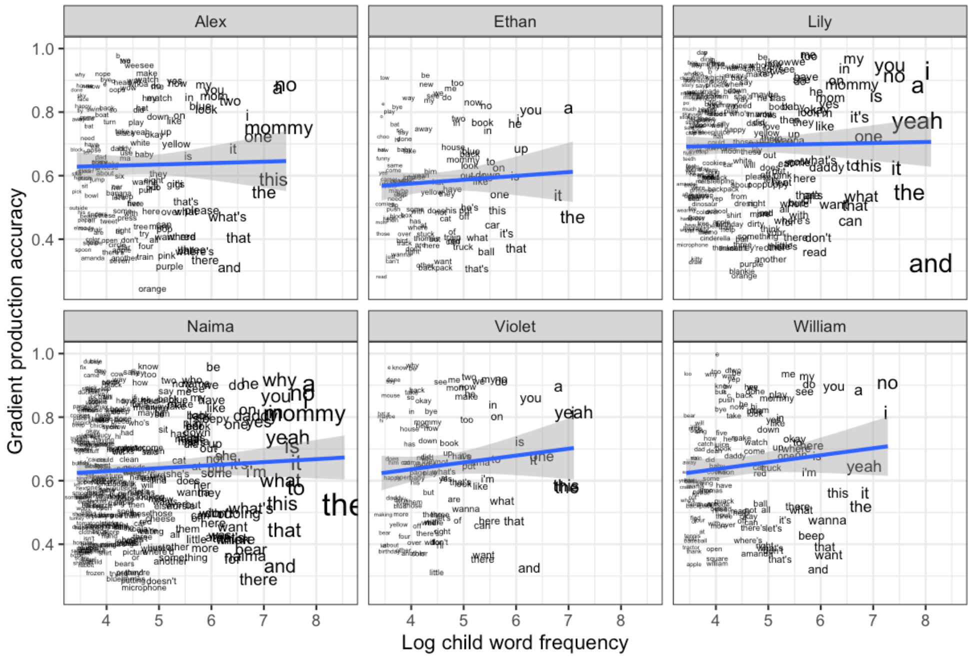
longer.

Minimal pairs. In addition to word length, the number of minimal pairs have an impact of production accuracy on the word level. The minimal pair counts used in this section are specific to each child rather than pooled across all the children. Figure 4.7a shows gradient word production accuracy[1] and child minimal pair counts for all the words in each of the children's production. There is generally an upward trend for all of them. The words clustered around the upper right corner (i.e., more accurately produced words) appear to be shorter. It is possible that these trends are the result of word length differences than minimal pair differences. To further evaluate the effect of minimal pairs on word production, 2-phoneme words for each of the children were plotted them themselves in Figure 4.7b. When limited to 2-phoneme words, the same trend is observed: Words with more minimal pairs are more accurately produced.

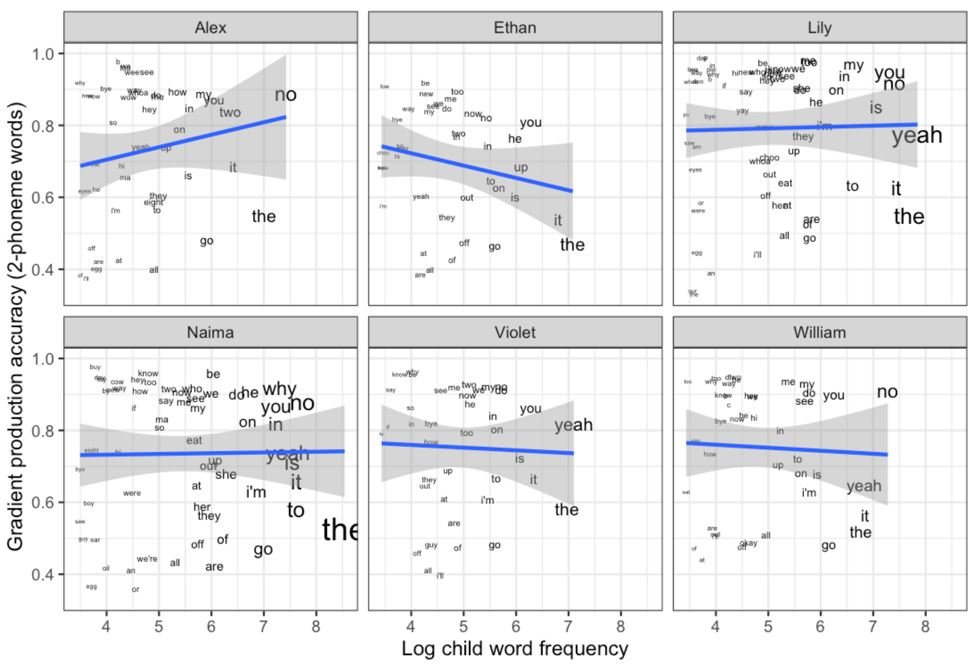
Word frequency. On the other hand, the same trend does not occur for word frequency. Figure 4.8a and Figure 4.8b plots word accuracy against child specific word-frequency counts for all the words and for 2-phoneme words respectively. Unlike the clear trends observed for minimal pairs, there are no patterns between word frequency and word production accuracy.

Phonotactic probability. Interestingly, there is to be a negative trend for phonotactic probability between word production accuracy and phonotactic probability. The trend appeared to be heavily affected by outliers for some of the children, but nevertheless it is a consistent pattern for all six children.

[1] Categorical word accuracy shows very similar trends. See Appendix A for a brief discussion.

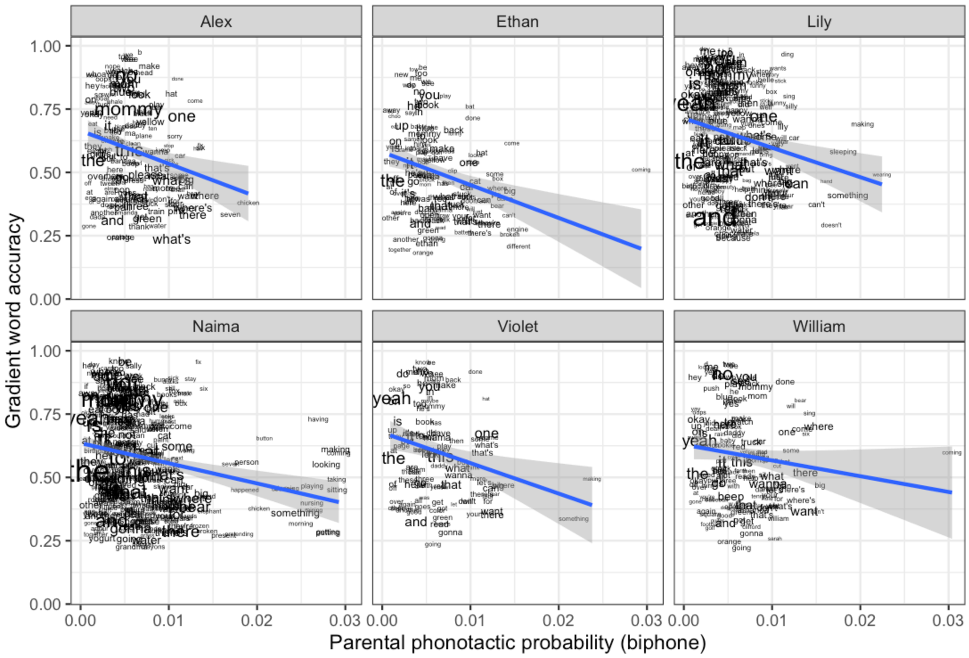


(a) Child word frequency and production accuracy for all words. There is no trend that more frequent words are more accurately produced.

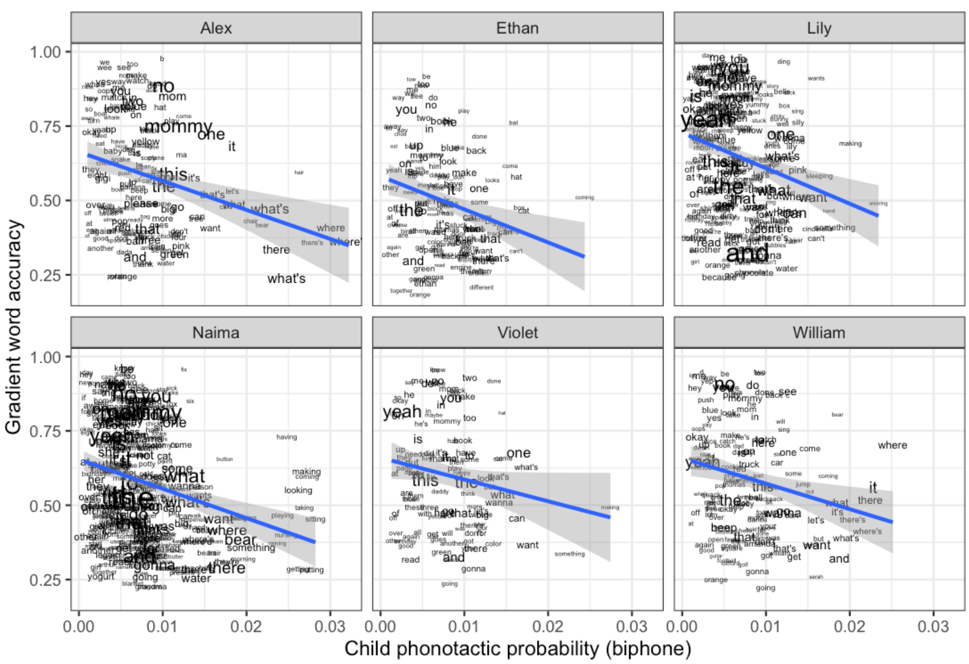


(b) Child word frequency and word production accuracy for the six children for 2-phoneme words.

Figure 4.8: Word frequency and word production accuracy.



(a) Parental phonotactic probabilities and child production accuracy for all words.



(b) Child phonotactic probability and production accuracy for all words.

Figure 4.9: Parental and child phonotactic probability and word production accuracy.

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
Alex				
(Intercept)	0.7268	0.0615	11.81	0.0000***
word length	-0.0448	0.0168	-2.67	0.0084**
child minimal pairs	0.0106	0.0023	4.51	0.0000***
child phonotactic probability	-8.6570	2.3533	-3.68	0.0003***
child word frequency	0.0000	0.0000	1.10	0.2708
$F(4,165) = 22.98, p < 0.0001,$ Adjusted $R^2 = 0.3422$				
Ethan				
(Intercept)	0.6592	0.0586	11.25	0.0000***
word length	-0.0587	0.0160	-3.68	0.0003***
child minimal pairs	0.0103	0.0020	5.29	0.0000***
child phonotactic probability	-6.3902	2.9075	-2.20	0.0295*
child word frequency	0.0000	0.0001	0.12	0.9015
$F(4,150) = 35.4, p < 0.0001,$ Adjusted $R^2 = 0.4719$				
Lily				
(Intercept)	0.7275	0.0598	12.17	0.0000***
word length	-0.0304	0.0164	-1.86	0.0640
child minimal pairs	0.0096	0.0019	5.04	0.0000***
child phonotactic probability	-6.4253	3.2096	-2.00	0.0463*
child word frequency	-0.0000	0.0000	-1.46	0.1442
$F(4,281) = 24.6, p < 0.0001,$ Adjusted $R^2 = 0.2488$				
Naima				
(Intercept)	0.7144	0.0448	15.94	0.0000***
word length	-0.0365	0.0108	-3.38	0.0008***
child minimal pairs	0.0066	0.0015	4.34	0.0000***
child phonotactic probability	-5.9205	1.9780	-2.99	0.0029**
child word frequency	-0.0000	0.0000	-0.46	0.6461
$F(4,381) = 29.12, p < 0.0001,$ Adjusted $R^2 = 0.2261$				
William				
(Intercept)	0.7435	0.0549	13.55	0.0000***
word length	-0.0497	0.0141	-3.53	0.0005***
child minimal pairs	0.0110	0.0018	6.28	0.0000***
child phonotactic probability	-7.8649	2.3223	-3.39	0.0009***
child word frequency	-0.0000	0.0001	-0.38	0.7022
$F(4,157) = 36.24, p < 0.0001,$ Adjusted $R^2 = 0.4669$				
Violet				
(Intercept)	0.7206	0.0894	8.06	0.0000****
word length	-0.0603	0.0279	-2.16	0.0330*
child minimal pairs	0.0102	0.0025	4.05	0.0001***
child phonotactic probability	-3.1716	3.8438	-0.83	0.4110
child word frequency	0.0001	0.0001	0.87	0.3883
$F(4,113) = 18.26, p < 0.0001,$ Adjusted $R^2 = 0.3711$				

Table 4.5: Linear regression results for the six children for word production accuracy.

Statistical modeling. For all children except Lily, the multiple regression models show that word length, the number of minimal pairs, and phonotactic probability are predictive of production accuracy, while word frequency is not. For Lily, word length is not a significant predictor, but minimal pairs and phonotactic probability are significant like the the models for the other five children. The results are summarized in Table 4.5.

4.4.2 Phoneme level production

Production accuracy was also quantified on the phoneme level for each child in the Providence Corpus. In this section, I look at minimal pair counts and phoneme frequency as predictors as phoneme production accuracy.

4.4.2.1 Methods

Phoneme level production accuracy was measured for each phoneme for each of the six children in the Providence Corpus. The quantification of phoneme production accuracy was more difficult than word level accuracy. When producing many words, children frequently omitted parts of the words and simplified consonant clusters. These words should not be discounted when measuring phonemic production accuracy. In cases with missing phonemes, the produced phonemes needed to be best matched with the target forms. This was done by converting the phonetic and phonemic transcription to a templatic representation in terms of sound type and syllable structure. For example, the word "pop" has the target form /pɑp/, and this would be converted to *CVC*. If the child produces [bɑ] for *pop*, this production would be converted to *CV*. The converted CV matches with the first two letters of CVC, and therefore [b] is compared to /p/, and [ɑ] is compared to /ɑ/, while the final /p/ is ignored. Individual minimal pair counts on the phoneme level and phoneme type frequency for each child are taken from the Section 4.3. The effects of minimal pairs and phoneme frequency on phoneme production accuracy are investigated through visualization and confirmed via linear regressions. Because of the high correlation between type frequency and token frequency (Pearsons's $r = 0.872$, $p < 0.0001$), only type frequency is used in these linear regression models.

4.4.2.2 Results

Minimal pairs. Figure 4.10 visualizes the relationship between phoneme production accuracy and minimal pair counts for each child. For all six children, there is a clear trend between the number of minimal pairs a phoneme has and how accurately it is pronounced. Certain phonemes, like /r/ and /ð/, appear to be especially difficult even though they have many minimal pairs.

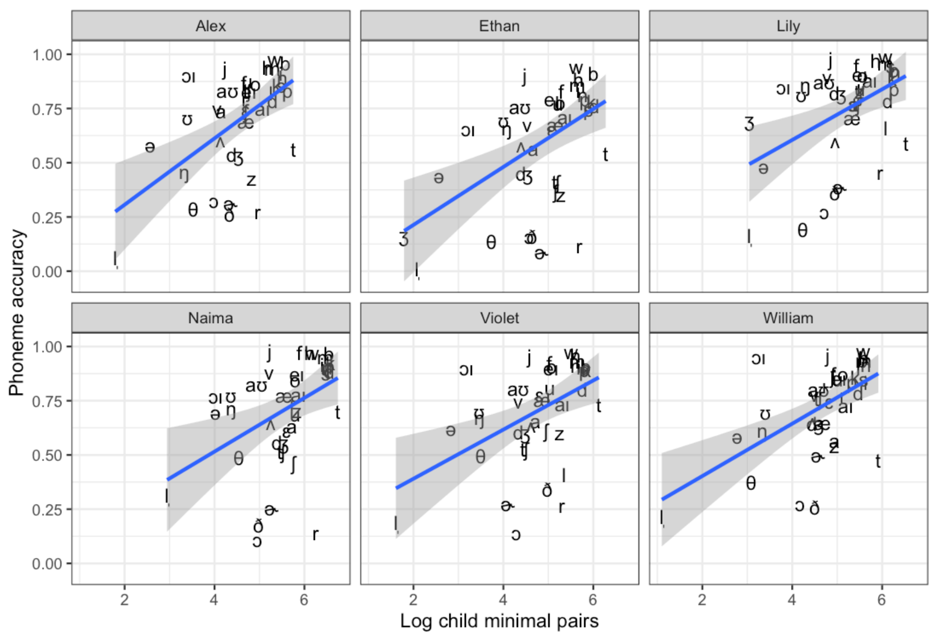
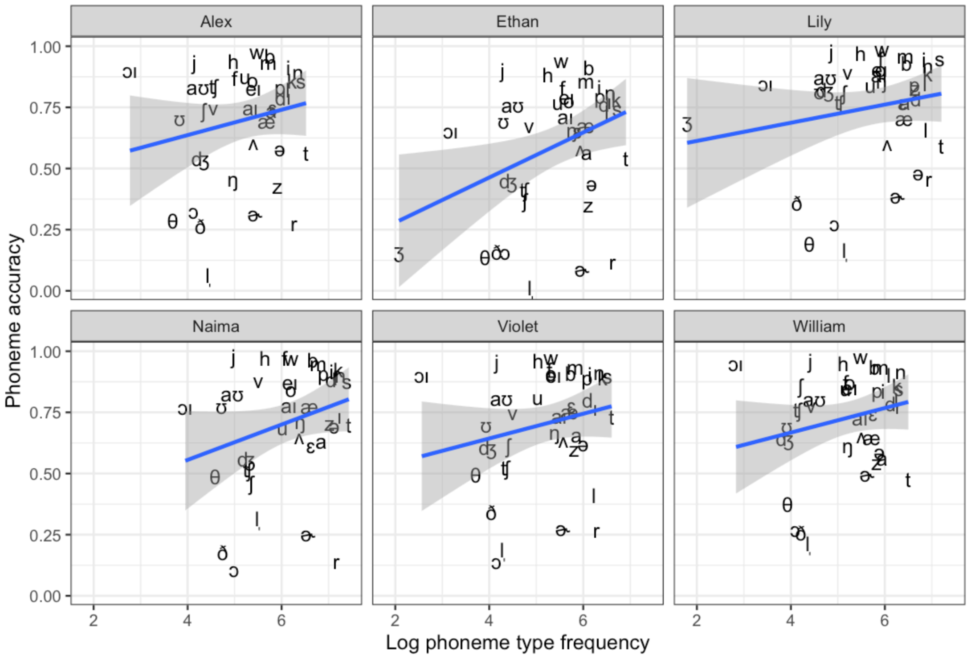


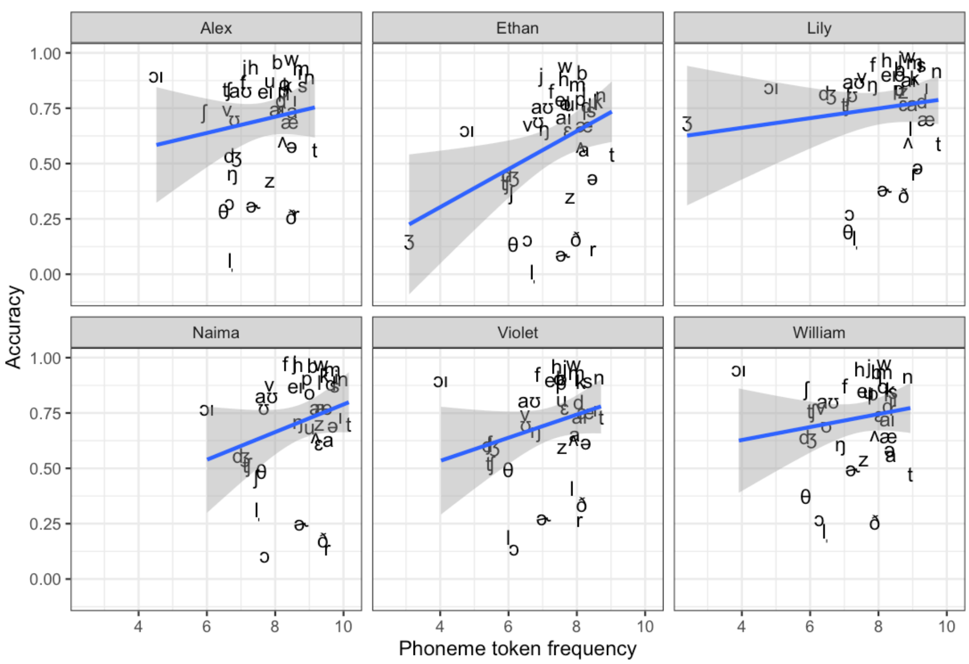
Figure 4.10: Child phoneme production accuracy and the number of minimal pairs.

Phoneme type frequency. There is a slight tendency for phoneme type frequency also. More frequent phonemes appear to be more accurately produced for some children, but the relationship is a lot weaker than minimal pairs.

Phoneme token frequency. Similar to type frequency, there is a slight tendency for phoneme token frequency. More frequent phonemes appear to be more accurately produced for some children, but the relationship is a lot weaker than minimal pairs.



(a) Child phoneme type frequency and child phoneme production accuracy.



(b) Child phoneme token frequency and child phoneme production accuracy.

Figure 4.11: Frequency and production accuracy

	Estimate	Std. Error	t value	Pr(> |t|)
Alex				
(Intercept)	0.5475	0.0636	8.61	0.0000***
child minimal pairs	0.0021	0.0006	3.62	0.0009***
child type frequency	-0.0005	0.0003	-1.75	0.0890
$F(2, 36) = 7.306, p = 0.00217$, Adjusted $R^2 = 0.2492$				
Ethan				
(Intercept)	0.4375	0.0698	6.27	0.0000***
child minimal pairs	0.0013	0.0005	2.70	0.0103*
child type frequency	-0.0002	0.0002	-0.97	0.3368
$F(2, 37) = 5.214, p = 0.01012$, Adjusted $R^2 = 0.177$				
Lily				
(Intercept)	0.6410	0.0602	10.65	0.0000***
child minimal pairs	0.0008	0.0003	2.93	0.0058**
child type frequency	-0.0002	0.0001	-1.61	0.1164
$F(2, 37) = 4.692, p = 0.01528$, Adjusted $R^2 = 0.1592$				
Naima				
(Intercept)	0.5554	0.0678	8.19	0.0000***
child minimal pairs	0.0006	0.0002	2.78	0.0086**
child type frequency	-0.0001	0.0001	-0.88	0.3824
$F(2, 36) = 5.116, p = 0.01108$, Adjusted $R^2 = 0.178$				
Violet				
(Intercept)	0.5729	0.0611	9.38	0.0000***
child minimal pairs	0.0014	0.0004	3.15	0.0033**
child type frequency	-0.0003	0.0002	-1.37	0.1778
$F(2, 36) = 5.858, p = 0.006272$, Adjusted $R^2 = 0.2036$				
William				
(Intercept)	0.5835	0.0576	10.14	0.0000***
child minimal pairs	0.0015	0.0004	3.37	0.0018**
child type frequency	-0.0003	0.0002	-1.26	0.2161
$F(2, 36) = 6.679, p = 0.003411$, Adjusted $R^2 = 0.2301$				

Table 4.6: Linear regression results for the six children for phoneme production accuracy. The difference in degree of freedom is the result of the phoneme /ʒ/, which is missing in some children's production.

Statistical modeling. In order to test whether these trends are significant, linear regressions were run for each of the children, and the results are summarized in Table 4.6. For all six children, the regression results show that minimal pair counts significantly predict phoneme production accuracy for all six children, while phoneme type frequency does not. To ensure that the lack of frequency effects is not the result of collinearity between minimal pairs and frequency counts, simple linear regressions are estimated for both type frequency (accuracy∼type frequency) and token frequency (accuracy∼token frequency) for each child.

Neither token or type frequency is a significant predictor in all twelve simple regression models.

4.5 Discussion

This chapter has two main goals. The first is to quantify the amount of minimal pair cues in the interaction between parents and young children, and the second goal is to investigate the lexical and sub-lexical factors in phonological acquisition. The results from both parts have important implications for the study of phonological acquisition.

4.5.1 Minimal pair cues in parental input and child speech

To accomplish the first goal, a detailed examination of minimal pairs in parental and child speech was carried out. To address the potential concern that very young children might not be able to make use of functional words or morphologically complex words, minimal pair counts were collected with six levels of word category exclusion. The results show that when certain word categories are excluded, the relative numbers of minimal pairs for each phoneme remain similar. There is very high correlations between minimal pair counts with varying degrees of word category exclusion as well as between parent and child counts. These results show that although the exclusion of certain word categories might make sense from the point of view of the child's linguistic ability, word exclusion in minimal pair counting should not make a huge impact on the results of any statistical analysis due to the high correlations between the different counts.

Moreover, the results from the minimal pair counts show that parental speech contain a surprising amount of words that differ by one segment even within a single hour. Even though some words that form minimal pairs might be rare in child-directed speech, many minimal pairs are highly frequent. Within the first few years of life, children are constantly exposed to minimal pair cues in natural speech, and the abundance of minimally-contrastive words allows the child to refine their phonological knowledge. Additionally, a pairwise phoneme minimal pair analysis shows that contrastive words exist for most pairs of phonemes as well, with exception of the phonotactically limited /ʒ/. There is clearly copious information for the learner to acquire phonological contrasts based on lexical contrast as quantified by minimal pairs.

4.5.2 Lexical contrast and minimal pairs

As reviewed in Section 4.1.2.4, developmental and historical studies of language use concepts related to minimal pairs commonly associated with phonological analysis. Studies in first language acquisition often look at the effects of phonological neighborhood density. The computation of phonological neighborhood density is similar to minimal pairs except that it typically includes words that differ through the addition or deletion of a phoneme. Although phonological analysis often only uses words of the same length in minimal pair analysis, the definition of minimal pairs does not exclude contrast with a "null" phonological unit. Functional load, most often used in diachronic studies, is in fact mostly measured through the number of minimal pairs a phoneme distinguishes.

While there is very little difference in quantification of phonological neighborhood density, functional load, and minimal pairs, these ideas are conceptualized and used differently in the literature. To advance our understanding of language acquisition and sound change, it is important to recognize that these different terms in fact measure the same thing: the amount of lexical contrast a phonological unit carries in the lexicon. The continued separation of these ideas is unnecessary and will only impede future efforts to better understand the interaction between lexical and phonological development and change.

4.5.3 Minimal pairs and phonological learning

To address the second goal of this chapter, word production and phoneme accuracy were used as approximate measures for phonological acquisition. On both the word level and the phoneme level, minimal pair counts are significant predictors of production accuracy as shown by linear regression results. This holds true for all six children in the corpus, based on their individual accuracy and minimal pair data. The consistency of the results suggest that minimal pairs are in fact an important cue for phonological acquisition.

In learning native sound categories, acoustic distributions and acoustic salience can play an important role in perceptual tuning. However, once linguistic cues are available in the form of contrastive lexical items, the learner can rely on these cues to acquire phonological distinctions. A lexical contrast model of learning does not require the learner to hear many words frequently to acquire a contrast; this model only requires the learner to have enough experience to understand that two words have distinct meanings. Although the learner may acquire a contrast on the most frequent pair "go" and "know" faster and earlier, the learner can just as well learn from "ball" and "call" as soon as they acquire these words and understand that these words have distinct meanings.

Studies in other domains of language acquisition suggest that it is often not the quantity of input that matters, but rather the quality. For instance, in word segmentation, while infants can make use of statistical cues (Saffran et al., 1996), they used speech cues such as stress rather than statistical cues when both cues are available (Johnson and Jusczyk, 2001; Peña et al., 2002; Thiessen and Saffran, 2003; Yang, 2004; Shukla et al., 2011). Similarly, for word learning, while word frequency and the amount of input clearly have an effect on the vocabulary size of the learner (Hart and Risley, 2003), the clarity of referential cues can also affect learning of new words (Cartmill et al., 2013; Trueswell et al., 2016).

The results from the minimal pair study can be related to other studies that use metrics similar to minimal pairs. Because of the similarities in phonological neighborhood and minimal pair measures on the word level, it is not surprising that phonological neighborhood density and minimal pairs make similar predictions on the word level (e.g., Carlson et al., 2014). Likewise, there is parallel between minimal pair findings here and functional load in studies of sound change. This is not surprising since sound change occurs as a result of language acquisition.

4.5.4 Phonotactic probability

The effect of phonotactic probability is investigated on the word level. For all six children, there is a negative trend – words with higher phonotactic probabilities are less accurately produced. This trend is significant for five of the children. There are several possible interpretations for this result, especially as the downward trend appears to be driven by relatively few words for all six children. It is possible that the downward trend is merely an artifact of frequency. First, since phonotactic probability employs phoneme sequence frequency as part of the calculation, it is possible that some of the highly frequent sequences contain relatively more difficult sounds, like [ð], and [ɹ] in words like "there" and "where". Second, another contributing factor is that the gerund ending *-ing* [ɪŋ] is highly frequent morphological suffix that can inflate the phonotactic probability of words like "making", "taking", and "sitting". These results indicate that perhaps stem-level phonotactic probability should be used to better evaluate the overall effect of phonotactic probability on phonological acquisition. Also, rather than using the suggested algorithm by Vitevitch and Luce (2004), phonotactic probability based on type rather than log token frequencies may yield more insightful results. If these results are not artifacts of token frequency, the developmental interpretation would be that children are more likely to pay attention to unfamiliar sound sequences, resulting in more accurate learning of words. Further study is necessary to determine the interaction between frequency and phonotactic probability.

4.5.5 Frequency

On both the word level and the phoneme level, the lack of frequency effects is consistent. More frequent words and phonemes are not more accurately produced. Although there is a slight trend of frequency for phonemes, the trend is not statistically significant. These results clearly show that hearing a word or a sound more frequently does not lead to better acquisition results for the given word or sound, and they directly contradict previous findings in Edwards and Beckman (2008) and Beckman and Edwards (2010). The combined lack of frequency effects and significant minimal pair effects indicate that it is how a sound functions in a lexical system that determines its acquisition trajectory.

4.5.6 Relation to the computational model

This dissertation investigates the role of lexical contrast in phonological acquisition. Because of the significant variation and overlap in the acoustic signal of distinct phonological categories, language learners must make use of additional information in the acquisition of phonological categories. The learning model outlined in Chapter 3 proposes that lexical contrast is an important cue in the acquisition of phonological categories. This learning mechanism is supported by the corpus study of child production accuracy in this chapter. Remarkably, minimal pair counts are predictive of production accuracy on both the word level and the phoneme level, and this pattern is very consistent for all six children in the Providence corpus. It is clear that phonological acquisition is more than the acquisition of phonetic patterns and that it is crucial that the learner is able to identify meaningful contrasts in the phonetic patterns from lexical cues.

4.6 Conclusion

In this chapter, I use developmental evidence to show that minimal pair cues are abundant in parental speech and that minimal pair cues, along with word length and phonotactic probability, are predictive of child production accuracy. Similar effects are not found for frequency. These results indicate that linguistically relevant cues, such as lexical contrast, play an important role in phonological acquisition.