A phonological, lexical, and phonetic analysis of the new words that young children imitate

Tania S. Zamuner & Andrea Thiessen
Department of Linguistics, University of Ottawa

Address for Correspondence:
Dept. of Linguistics
University of Ottawa
Hamelin Hall, 70 Laurier Ave. East
Ottawa ON, Canada, K1N 6N5
tzamuner@uottawa.ca

Acknowledgements:
This research was supported by a SSHRC grant awarded to TSZ. We thank Cynthia Core, Yvan Rose, Sophia Stevenson, Ana Sofia Niembro, Katherine Lam, and anonymous reviewers for helpful input and feedback.
Abstract

As children learn language, they spontaneously imitate the speech of those around them. This research investigates the new words that five children imitated between the ages of 1;0 and 2;11. Children were more likely to imitate new words as they aged and as their productive language developed. After controlling for age, children also were more likely to imitate new words that were shorter and with high neighborhood densities, and that contained sounds the children had previously produced accurately. Together, the findings demonstrate that both the patterns of the target words and children’s productive abilities are predictors of children’s imitated speech. This supports models of language development where there are influences stemming not only from phonological and lexical representations, but also phonetic representations.

Keywords: language development; speech production; phonology; lexicon; phonetics; imitation
A phonological, lexical, and phonetic analysis of the new words that young children imitate

1. Introduction

Imitation and language development has a long history (Kymissis and Poulson 1990). Imitation has been analyzed at different linguistic levels, such as phonology, word learning, syntax, and pragmatics (Fraser et al. 1963, Macken 1975, Clark 1977, Goldinger 1998, Clark 2007). Imitation has been argued to have a number of diverse functions (Speidel and Nelson 1989), such as providing a pragmatic function (Clark 2007), and as a means of internalizing language (Saville-Troike 1988), or with different functions across development (Lewis, 1936/1999). Imitation is such a pervasive behaviour in early development, it has been argued to be an innate ability (Meltzoff 2005), involving dedicated mirror neuron systems (Rizzolatti and Arbib 1998, Vihman, 2002).

Notwithstanding the wealth of research on imitation, the topic of imitation has received little attention in the past 50 years. This reflects the shift away from behaviourist theories (Skinner 1957, Olmsted 1971) to generativist theories (Chomsky 1957, 1959), although Chomsky himself did acknowledge that imitation does play a function in language acquisition (Speidel and Nelson 1989:14). Moreover, in the body of the research on imitative speech in language development that does exist, little work on imitation examined the relative influence of phonological, lexical, and phonetic (output) factors on children’s imitative speech. This research provides a corpus analysis of new words presented to children aged 0;11-2;11 to determine the predictors for children’s imitations of the new words presented by their parents.

Young learners have substantial receptive knowledge about the phonological and grammatical structures of their language even before they begin producing speech
(Werker and Curtin 2005, Curtin and Zamuner 2014). However, in typical language development, learners move beyond their initial comprehension skills to both comprehend and express language. We approach imitation as it relates to speech production and language acquisition. When children imitate, they produce speech. We investigate imitative speech productions for insights into children’s developing language system, specifically, the relationship between language development and phonological, lexical, and articulatory factors. In the next section, we review the role that these factors have been found to play in previous research, focusing on children under 3 years of age.

2. Phonological, lexical, and phonetic factors in language development

Research on phonological, lexical, and phonetic (output) factors in language development spans the period in development before the onset of speech production to a stage where children have acquired productive language (approximately 3 to 4 years of age). One approach has been to examine the types of words that infants and young children can recognize or learn in a controlled experimental setting. There is also research using non-word repetition and picture-word elicitation tasks, to examine topics such as vocabulary development, production accuracy, and variability (e.g., Hoff et al. 2008, Sosa and Stoel-Gammon 2012, Macrae 2013, Zamuner and Kharlamov 2016). Other studies take a longitudinal approach, examining the words that are added to children’s vocabulary over time, based on spontaneous speech samples or lexical normative data (e.g., de Bree et al. 2014; Maekawa and Storkel, 2006). We focus on these types of longitudinal research.

There has been a recent focus on phonological and lexical factors in language development, which is highlighted in a number of review articles (e.g., Saffran and Graf Estes 2006, Stoel-Gammon 2011, Curtin and Zamuner 2014). One commonly
investigated phonological factor is phonotactic probability, which refers to the likelihood that sounds will occur in a given environment. Infants are better at learning words with common phonotactic patterns (Gonzalez-Gomez et al. 2013, though see Storkel 2009). Commonly investigated lexical factors are word length in phonemes and neighborhood density (ND), which refers the amount of phoneme overlap a word has with other words in the lexicon. Children are more likely to learn shorter words and words that reside in dense phonological neighbourhoods (Stoel-Gammon 1998; Storkel 2004, 2006; Coady and Aslin 2004; Maekawa and Storkel 2006).

Turning to phonetic factors, it has been argued that there is a tight link between articulation and development (Rvachew 1994, Vihman 2009). Researchers are beginning to investigate whether infant speech perception involves articulatory and sensori-motor representations, before the onset of meaningful speech (DePaolis et al. 2011, Yeung and Werker 2013, Bruderer et al. 2015). This is relevant for the emergence of speech production because the findings suggest that these representations are an early and integral part of language processing. There is continuity across the sounds founds in infants’ babbling repertoires and their first words (Vihman et al. 1985) and there is evidence in early child language for some children displaying lexical selection and avoidance (Ferguson and Farwell 1975, Leonard et al. 1979, Schwartz and Leonard 1982, Vihman et al. 2016). Schwartz and Leonard (1982) found that young children were more likely to imitate new words that contain sounds IN rather than OUT of the child’s own production repertoire. However, the stimuli were confounded with other phonological and lexical factors. For example, the OUT words /uʃuʃ, ozoʃ/ contain infrequent English phonemes, have few phonological neighbours, have a less frequent English syllable
structure, have a less typical English stress pattern, among others. Thus, it is difficult to ascertain whether imitation patterns reported in Schwartz and Leonard’s experiment stem from the child’s own articulatory abilities or from other factors. Lexical avoidance and selection can be captured by an ‘articulatory filter’ (Vihman 1993, Vihman et al. 2014, Vihman 2017), which hypothesizes that sounds that the learner has produced are more salient in the learner’s input (also see Stoel-Gammon, 2011 for her discussion of an auditory–articulatory loop). The result is that words comprised of produced sounds are facilitated by the learner’s previous production experience, which in turn supports the memory and recognition and learning of word forms with those segments (Keren Portnoy et al. 2010).

A subset of studies has looked at a combination of phonological, lexical, and phonetic factors in children’s vocabulary development. Maekawa and Storkel (2006) performed backward regression analyses on longitudinal production data from three children in the CHILDES database between the ages of 1;4-3;1. Word length was a predictive factor for all children (children produced shorter words before longer words); however, phonotactic probability, ND, and word frequency varied in their predictability across the children. They also found that for one child, a significant predictor for vocabulary development was the normative production accuracy for the words’ final consonants (but not the words’ initial consonants). That is, early-acquired words were more likely to contain final consonants that 3 year olds typically produce accurately (based on norms established by Smit et al. 1990). This is similar to the approach taken in the current study; however, we examine individual children’s own production accuracy rather than based on normative data. Carlson et al. (2014) conducted a study on
longitudinal production data from 64 children between the ages of 1;2-4;2, and analyzed which factors (such as age, word frequency, word length, ND, phonotactic probability, but no phonetic factors) were significant predictors of when a new word entered the children’s productive vocabularies. They found that the words that emerged in children’s productive vocabulary were those that sounded similar to other words (high ND), but which did not have a lot of interference from other words in the lexicon. (Interference was measured by the variables clustering coefficient and coreness – which is argued to reflect lexical access during speech production).

Theories vary in the relative weight that is given to phonological, lexical, and phonetic (output) factors in language development, with phonetics typically having a more peripheral role. Exemplar theories and usage-based models propose that speech output representations are integrated into phonological and lexical representations (Pierrehumbert 2003, Sosa and Bybee 2008, Munson et al. 2012, Sosa and Stoel-Gammon 2012). In these models, the development of phonological, lexical and phonetic representations is multidirectional as all form part of the long-term representations. Acoustic-motor representations are the focus of the newly developed A-map model, which highlights the importance of the child’s own articulations in language development (McAllister Byun et al. 2016). While the A-map model was developed to specifically account for child-specific phonological patterns, the model can be adopted to make predictions for language development. Central to the A-map model is “the pressure to match adult productions of a given word… and the pressure to attempt a pronunciation that can be realized reliably” (McAllister Byun et al. 2016:2). Given these pressures, one would predict that children would be more likely to produce and/or learn words that
contain sounds that the child has previously produced accurately because these items would both match the adult target and be reliably produced by the child.

The broad aim of this research was to examine what children imitate for insights on the developing language system. The more specific aim was to examine the relative contribution of phonetic factors on children’s speech, while taking into account phonological and lexical factors. We conducted an analysis of imitation on longitudinal data from the Providence Corpus from CHILDES (MacWhinney 2000, Demuth et al. 2006). Following Clark (2007), we identified syntactic frames in which parents made direct offers of new words. For example, ‘It’s called a kid’ (Lily, 1;5.21). Each direct-new-offer was coded for whether it was imitated and for a number of predictor variables where were defined as phonological, lexical, and phonetic factors (phonetic was defined as children’s sound inventory and children’s production experience with the sounds in the direct-new-offers). Binary hierarchical logistical regressions were used to determine which variables were significant predictors for whether children would imitate a direct-new-offer, after taking into account the control variable of chronological age.

3. Corpus Study

3.1. Methodology

Analyses were done on corpora of five mother-child dyads (Alex, William, Lily, Naima, and Violet), from monolingual English-speaking children from the Providence Corpus (MacWhinney 2000, Demuth et al. 2006). These data comprise spontaneous speech collected between the ages of 0;11-4:0, and were typically collected bimonthly. Separate analyses were done for each child, and analyses were restricted to nouns, as done by Maekawa and Storkel (2006), Zamuner (2009), among others. Each session was analyzed
to identify new words, which were defined as words not produced during the previous recording sessions by either the child or other people present. The set of new words from each session was then re-checked to ensure that the words did not appear in any previous transcript in the singular or plural form. Each new word was also checked to ensure that the child did not first produce the word before the direct offer within the concurrent session.

The context for each new word was then coded, by checking the CHAT transcript for whether the new word was presented as direct-offers. Direct offers were defined by syntactic frames identified in Clark and Wong (2002), for example, “What is /what’s . . . ?; This is . . . ; That is /that’s . . . ; and This is /that’s called . . . ” (Clark and Wong 2002:185). Direct offers also included new words that were presented in isolation, e.g., when a parent labeled an item in a book or in the immediate environment. The coding for the final set of direct-new-offers was additionally verified by checking the audio-video recordings. A subset of the words that were identified as direct-new-offers from the CHAT transcript were removed from the analyses after watching the audio-video recordings for a variety of reasons: e.g., multiple people speaking at the same time; exchange was between two adults; mother talking to herself. In limited cases, words were excluded because the immediate context indicated that a word was not new to the child, such as discussing an event that occurred the day before. The final set of direct-new-offers were considered “unlikely to be already known to the child” based on the fact that the words did not previously appear in the corpus (Clark 2007:164). Direct offers of new words were restricted to those that occurred in sentence-final position, to ensure that the new words were in relatively the same prominence. Past research found that direct offers
of new words are more likely to occur in sentence-final position (Clark and Wong 2002). Each direct-new-offer was coded for whether the child imitated the direct-new-offer (107 imitated, 454 not imitated), either immediately (n=104) within the next 5 conversational turns (n=3). Six additional responses by the children were coded as unintelligible. Direct-new-offers that were coded as not imitated included a variety of responses, such as acknowledgments, continuation or change of the topic. See work by Clark (2007) who examines the pragmatic function of these responses.

One difficulty in using spontaneous speech is that it can contain infrequent words, which can skew the results. To control for this, the direct-new-offers were searched for in the Child Mental Lexicon (Storkel and Hoover 2010), which is an on-line corpus of child spoken language (4,832 types, 1,028,417 tokens). The assumption was that words found in the Child Mental Lexicon are more representative of child language. For example, some direct-new-offers not found in the Child Mental Lexicon were dowel (Naima 1;6.4), reflection (Alex 1;11.27), and pimientos (Lily 2;9.18). While the analyses focus on the subset of words found in the Child Mental Lexicon, the analyses were also repeated with the full set of direct-new-offers and similar results were found. Table 1 presents the number of direct-new-offers for each of the five children, broken down by chronological age, and whether the words were found in the Child Mental Lexicon.
The imitation rates by children in the Providence Corpus are lower (19%) than those reported by Clark (2007), who found that children imitate direct offers of new words an average of 54% of the time. Clark (2007) also reported that parents presented direct offers of new words in phrases 81% of the time, and as isolated words 19% of the time. In our data, direct-new-offers in phrases accounted for 54% of the new words, and those presented as isolated words accounted for 46% of the new words. Differences in imitation rates and the context of direct-new-offers could stem from situation contexts.
which vary across the recordings, such as playing games, meal times, and reading books. Data collection from the Providence corpus was less dense after 3;0 (Table 1). For this reason, only the data from children between the ages of 0;11 and 2;11 were used for the analyses. A list of the direct-new-offers is presented in Appendix A. The following analyses examine phonological, lexical, and phonetic (output) variables, to determine were variables significantly predictors for whether a child imitated a direct-new-offer.

3.2. Expressive Language Measures

Imitation has been linked to other skills, such as chronological age and vocabulary growth (Snow 1989, Masur and Eichorst 2002). Five measures of each child’s expressive language development were calculated. Cumulative type and token counts for each child’s productive vocabulary were calculated for each session using the freq command in CLAN, transformed to logarithm form (Expressive Vocabulary Type, Expressive Vocabulary Token). Each child’s consonant phoneme inventory was calculated using the Phoneme Accuracy tool in PHON (https://www.phon.ca, Rose et al. 2006, Rose and MacWhinney 2014). This tool provides a list of the child’s phoneme inventory, their frequency, and their accuracy. As with children’s productive vocabulary, each child’s phoneme inventory was calculated cumulatively, for example, at session 1, session 1 and 2 combined, etc. The measure Accurate Consonant Inventory was based on the number of accurately produced phonemes in the child’s inventory, and a phoneme had to occur only a single time to be included. The measure Accurate Consonant Inventory 10 provided a more conservative estimate, which required that a phoneme be produced a minimal of 10 times accurately. We also calculated phoneme inventories independent of accuracy (e.g., a production of ‘bottle’ as ‘mottle’ could count as an instance of the child producing an
However, these measures very quickly reached at ceiling and were not informative.

Lastly, the child’s mean-length-utterance (MLU) was calculated for each session.

Correlations between chronological age (referred to as age from this point forward) and the Expressive Language measures are in Table 2. There were significant correlations between age and all measures. In addition, there were strong correlations (over .85) among expressive vocabulary and the children’s accurate consonant inventories. This is problematic for the logistic regressions used in the following analyses because of high multicollinearity. Therefore, only MLU and Accurate Consonant Inventory 10 were used in the analyses. It is important to keep in mind that because of the strong correlations, Accurate Consonant Inventory 10 cannot be interpreted as only referring to the size of the child’s accurately produced phoneme inventory, because this variable overlaps with other measures of expressive language.

Table 2: Correlations among age and Expressive language measures (N=264), correlations over .85 are in bold

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>−</td>
<td>.78***</td>
<td>.80***</td>
<td>.75***</td>
<td>.81***</td>
<td>.67***</td>
</tr>
<tr>
<td>2. Expressive Vocabulary Type</td>
<td>−</td>
<td>.98***</td>
<td>.96***</td>
<td>.97***</td>
<td>.80***</td>
<td></td>
</tr>
<tr>
<td>3. Expressive Vocabulary Token</td>
<td>−</td>
<td>.95***</td>
<td>.96***</td>
<td>.79***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Accurate Consonant Inventory</td>
<td>−</td>
<td>.96***</td>
<td>.68***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Accurate Consonant Inventory 10</td>
<td>−</td>
<td>.74***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. MLU</td>
<td>−</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 (two-tailed), ** p < .01 (two-tailed), *** p < .001 (two-tailed), using Bonferroni adjusted alpha levels for multiple comparisons

Binary hierarchical logistical regressions were used for the analyses. Control variables are added in Block 1 and predictor variables are added in Block 2 (Reed and Wu 2013).

A significant change in the model between Blocks 1 and 2 indicates that the predictor variables improves the membership of the dependent variable, after taking into account the control variable. The goal was to determine whether after controlling for participants’
age, Accurate Consonant Inventory 10 and/or MLU significantly accounted for whether children would imitate a direct-new-offer. The standard errors of the independent variables were below 2, indicating that there was no evidence of multicollinearity. The full model was statistically significant, $\chi^2(2, N = 164) = 7.75, p < .05$ and explained 13% of the variance (Nagelkerke $R^2$) (Table 3). After controlling for age, only Accurate Consonant Inventory 10 was a significant predictor for distinguishing between the direct-new-offers that children did and did not imitated ($\chi^2(1, N = 164) = 8.32, p < .01$). The value of Exp(B) or the Odds Ratio implies that direct-new-offers were 1.12 times more likely repeated by children with larger accurate phoneme inventories. The same results were found when Accurate Consonant Inventory 10 was replaced by Expressive Vocabulary Type. The analyses were also run using the full set of direct-new-offers with the same results as when based on the subset of words found in the Child Mental Lexicon.

Table 3. Binary hierarchical logistic regression models predicting probability of child imitating direct-new-offer, age and expressive language measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>S.E.</th>
<th>Wald</th>
<th>$p$</th>
<th>Exp(B)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 1: Age</td>
<td>.003</td>
<td>.001</td>
<td>12.92</td>
<td>.000</td>
<td>1.10</td>
<td>1.00 - 1.01</td>
</tr>
<tr>
<td>Block 2: Age</td>
<td>.002</td>
<td>.001</td>
<td>2.85</td>
<td>.091</td>
<td>1.00</td>
<td>1.00 - 1.00</td>
</tr>
<tr>
<td></td>
<td>MLU</td>
<td>.021</td>
<td>.24</td>
<td>.008</td>
<td>.93</td>
<td>.64 - 1.63</td>
</tr>
<tr>
<td></td>
<td>Accurate Consonant Inventory 10</td>
<td>.11</td>
<td>.04</td>
<td>8.32</td>
<td>.004</td>
<td>1.04 - 1.21</td>
</tr>
</tbody>
</table>

Block 1: Nagelkerke $R^2 = .08, \chi^2 = 13.10, df = 1, p < .001; \text{ Block 2: Nagelkerke } R^2 = .13, \chi^2 = 7.75, df = 2, p < .05.$

The relationship between children’s Age in months, Accurate Consonant Inventory 10, and children’s imitation of direct-new-offers is illustrated in Figure 1.
Children were more likely to imitate a direct-new-offer as the size of their consonant inventory grew.

Figure 1: Relationship between direct-new-offers imitated, children’s age in months and Accurate Consonant Inventory 10

3.3. Target Word Measures

The second set of analyses examine whether children were more likely to imitate direct-new-offers based on the words’ phonological and lexical properties, obtained from the Child Mental Lexicon (Storkel and Hoover 2010). This child-based corpus was used because comparative analyses have found differences in phonological and lexical patterns between child- and adult-based corpora (Storkel and Hoover 2010). The four Target Word measures were all based on logarithm forms and were as follows: word length, phonotactic probability, ND, and word frequency. Word Length in phonemes and ND were chosen because they have been established in the literature as significant predictors of vocabulary development (Storkel 2004, Stokes 2006, Carlson et al. 2014). We also included phonotactic probability (PP) because it has been shown to be a factor in the
accuracy of children’s speech production (e.g., Zamuner et al. 2004, Munson et al. 2005), though the factor has been less predictive in analyses of children’s vocabulary growth (Maekawa and Storkel 2006, Carlson et al. 2014). Lastly word frequency was included in the analyses. While lexical frequency has been shown to be a significant factor in phonological acquisition (Ota et al. 2013), the expectation was that this would not be a significant predictor because all the words were assumed to be new to the children, within the context of the recording sessions.

Correlations between age and the Target Word measures are provided in Table 4. Age was not significantly correlated to any of the Target Word measures; however, the high correlation between Word Length and ND is problematic due to high multicollinearity. In fact, a binary hierarchical logistical regression analyses that included both Word Length and ND variables gave a standard error above 3 for Word Length, indicating multicollinearity. Therefore, only the ND measure was used, while keeping in mind that this variable is highly correlated to and cannot be isolated from Word Length.

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>−</td>
<td>.12</td>
<td>.10</td>
<td>−.15</td>
<td>−.18*</td>
</tr>
<tr>
<td>2. Word Length</td>
<td>−</td>
<td>.58***</td>
<td>−.84***</td>
<td>−.28***</td>
<td></td>
</tr>
<tr>
<td>3. PP</td>
<td>−</td>
<td>−.33***</td>
<td>−.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. ND</td>
<td>−</td>
<td></td>
<td></td>
<td>.19*</td>
<td></td>
</tr>
<tr>
<td>5. Word Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 (two-tailed), ** p < .01 (two-tailed), *** p < .001 (two-tailed), using Bonferroni adjusted alpha levels for multiple comparison

A binary hierarchical logistical regression was used to assess whether characteristics of the Target Word predicted whether children would imitate a direct-new-offer, after controlling for participants’ age (Table 5). There was no evidence of multicollinearity. The full model was statistically significant, accounting for 15% of the
variance (Nagelkerke $R^2$). After controlling for age, the variable ND was a significant predictor for the direct-new-offers that children imitated ($\chi^2(1, N = 164) = 7.36, p < .01$). The value of Exp(B) for the variable ND Log implies that direct-new-words were 3 times more likely to be imitated when the word had more phonological neighbours.

Recall though that ND was highly correlated to Word Length and therefore, the effect cannot be uniquely attributed to ND. When the same model was run with Word Length in place of ND, the results were the same.

Table 5. Binary hierarchical logistic regression models predicting probability of child imitating direct-new-offer, age and target word measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>$p$</th>
<th>Exp(B)</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 1:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.003</td>
<td>.001</td>
<td>12.92</td>
<td>.000</td>
<td>1.10</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td>Block 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.004</td>
<td>.001</td>
<td>17.01</td>
<td>.000</td>
<td>1.04</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td>PP</td>
<td>.24</td>
<td>.61</td>
<td>.16</td>
<td>.69</td>
<td>1.28</td>
<td>.39</td>
<td>4.20</td>
</tr>
<tr>
<td>ND</td>
<td>1.05</td>
<td>.39</td>
<td>7.36</td>
<td>.007</td>
<td>2.85</td>
<td>1.34</td>
<td>6.09</td>
</tr>
<tr>
<td>Word Frequency</td>
<td>1.92</td>
<td>1.47</td>
<td>1.70</td>
<td>.19</td>
<td>6.84</td>
<td>.38</td>
<td>122.80</td>
</tr>
</tbody>
</table>

Block 1: Nagelkerke $R^2 = .08, \chi^2 = 13.10, \text{df} = 1, p < .001$; Block 2: Nagelkerke $R^2 = .15, \chi^2 = 10.53, \text{df} = 3, p < .05$.

As expected Word Frequency was not a significant predictor of children’s imitations of direct-new-offers. The same results were found when running correlations and regression analyses using the full set of direct-new-offers, as when using the subset of words found in the Child Mental Lexicon.

The relationship between children’s age in months, the target word’s ND, and whether or not children imitated the direct-new-offer is provided in Figure 2. Children were more likely to imitate a direct-new-offer when it had a higher ND.
3.4. Target Word Production Experience

The third analyses focused on children’s production experience with the consonants in the direct-new-offers, to determine whether children were more likely to imitate direct-new-offers that contained sounds the children had previously produced. Each child’s experience producing the consonants in the direct-new-offers was measured, using the Phoneme Accuracy tool in PHON. The child’s production experience was evaluated independent of the child’s accuracy (Independent Target, Independent Target 10), and relative to how accurately children had produced those sounds (Relational Accurate, Relational Accurate 10). For Independent Target and Relational Accurate, a consonant had to only occur once to be counted, and the more conservative estimates of Independent Target 10 and Relational Accurate 10 required that a consonant occur a minimal of 10 times. To illustrate, take the direct-new-offer peacock (Naima, 1;0;14), with three consonants: /p, k, k/. The /p/ had not yet occurred in Naima’s productive
vocabulary up to this point in the database, and /k/ occurred in 15 of her words, of which /k/ was produced accurately seven times. Therefore, both the Independent Target and Independent Target 10 cells equaled .66 (2 of the 3 consonants appeared in her target words). For Relational Accurate, the cell also equaled .66 (2 of the 3 consonants appeared accurately at least once in her productive vocabulary); however, Relational Accurate 10 equaled 0 because none of the consonants had been produced accurately by Naima over 10 times. It is possible to conduct a more nuanced analysis of segmental accuracy, taking into account whether a child accurately produces segments or makes substitutions or deletions. It is also possible to take positional effects and complexity into account when coding for accuracy. For example, take the word *peacock*, one could code for whether Naima had any words in her lexicon in which /p/ was produced correctly in word-initial position, in singleton onsets. Our initial goal was to analyze the data taking into account prosodic position and complexity; however, the coding became too fine-grained for the dataset, so it was decided to code at a broader level. Table 6 provides the correlations between age in days and the target word production variables. All variables were significantly correlated with each other, though no correlations were over .85.

Table 6: Correlations among age and Target Word Production Experience Measures (N=264)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>−</td>
<td>.33***</td>
<td>.51***</td>
<td>.35***</td>
<td>.58***</td>
</tr>
<tr>
<td>2. Independent Target</td>
<td>−</td>
<td></td>
<td>.70***</td>
<td>.62***</td>
<td>.48***</td>
</tr>
<tr>
<td>3. Independent Target 10</td>
<td>−</td>
<td></td>
<td></td>
<td>.60***</td>
<td>.73***</td>
</tr>
<tr>
<td>4. Relational Accurate</td>
<td>−</td>
<td></td>
<td></td>
<td></td>
<td>.72***</td>
</tr>
<tr>
<td>5. Relational Accurate 10</td>
<td>−</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 (two-tailed), ** p < .01 (two-tailed), *** p < .001 (two-tailed), using Bonferroni adjusted alpha levels for multiple comparison
A binary hierarchical logistical regression assessed whether the child’s production experience with the consonants in the target word predicted whether children would imitate a direct-new-offer, after controlling for participants’ age (Table 7). No outliers were detected and there was no evidence of multicollinearity. The full model was statistically significant and explained 17% of the variance (Nagelkerke $R^2$). After controlling for Age, the variable Relational Accurate 10 was a significant predictor for distinguishing between the direct-new-offers that children imitated ($\chi^2(1, N = 164) = 8.87, p < .01$). The value of Exp(B) for Relational Accurate 10 indicates that direct-new-words were 14 times more likely to be repeated when it consisted of consonants that the child had produced accurately at least 10 times previously. The same analyses were conducted using the full set of direct-new-offers, with the same results.

Table 7. Binary hierarchical logistic regression models predicting probability of child imitating direct-new-offer, age and child production experience

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.003</td>
<td>.001</td>
<td>12.92</td>
<td>.000</td>
<td>1.10</td>
<td>1.00 - 1.01</td>
</tr>
<tr>
<td>Block 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.001</td>
<td>.001</td>
<td>1.53</td>
<td>.22</td>
<td>1.00</td>
<td>.99 - 1.00</td>
</tr>
<tr>
<td>Independent Target 10</td>
<td>-.92</td>
<td>1.07</td>
<td>.73</td>
<td>.39</td>
<td>.40</td>
<td>.05 - 3.25</td>
</tr>
<tr>
<td>Relational Accurate 10</td>
<td>2.62</td>
<td>.88</td>
<td>8.87</td>
<td>.03</td>
<td>13.72</td>
<td>2.49 - 76.85</td>
</tr>
</tbody>
</table>

Block 1: Nagelkerke $R^2 = .08$, $\chi^2 = 13.10$, df = 1, $p < .001$; Block 2: Nagelkerke $R^2 = .17$, $X^2 = 13.71$, df = 2, $p < .001$.

A figure of the relationship between children’s age in months, Relational Accurate 10 (average accuracy of children’s production for the consonants in the direct-new-offer), and whether or not children imitated the direct-new-offer is given in Figure 3. Children were more likely to imitate a direct-new-offer when it contained more consonants that the child had accurately produced at least 10 times before.
Figure 3: Relationship between direct-new-offers imitated, children’s age and Relational Accurate 10

![Graph showing relationship between direct-new-offers imitated, children’s age and Relational Accurate 10.](image)

3.5. Combined Measures

The last analyses combined the significant predictor variables from the previous analyses: Accurate Inventory 10, ND, and Relational Accurate 10). Correlations among the variables is given in Table 8. There were significant correlations between many of the variables, though none were over .85.

Table 8: Correlations among age and all measures, subset of words (N=264)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td></td>
<td>.81***</td>
<td>-.15</td>
<td>.58***</td>
</tr>
<tr>
<td>2. Accurate Inventory 10</td>
<td></td>
<td>-.18*</td>
<td>.78***</td>
<td></td>
</tr>
<tr>
<td>3. ND</td>
<td></td>
<td></td>
<td>-.06</td>
<td></td>
</tr>
<tr>
<td>4. Relational Accurate 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 (two-tailed), ** p < .01 (two-tailed), *** p < .001 (two-tailed), using Bonferroni adjusted alpha levels for multiple comparison

A final binary hierarchical logistical regression analysis assessed a combination of the predictor variables, after controlling for participants’ age (Table 9). The model was run
with outliers removed that had standard residuals of ±2.58; however, this did not improve classification by more than 2% so the original model with all data was used. There was no evidence of multicollinearity. The full model was statistically significant, $\chi^2(3, N = 164) = 22.13, p < .001$, explaining 22% of the variance (Nagelkerke $R^2$). After controlling for age, the variables ND ($(\chi^2(1, N = 164) = 8.47, p < .01)$ and Relational Accurate 10 ($(\chi^2(1, N = 164) = 3.91, p < .05)$ were significant predictors for distinguishing between the direct-new-offers that children imitated. Direct-new-offers were 3 times more likely repeated when the word resided in a denser phonological neighbourhood, and direct-new-offers were 4 times more likely to be imitated when the target word contained phonemes that the child had previously produced accurately. The variable Accurate Inventory 10, which was a measure of children’s overall language (size of the child’s accurate phoneme inventory), was not a significant predictor when the other variables were included in the model.

Table 9. Binary hierarchical logistic regression models predicting probability of child imitating direct-new-offer, age and all measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.003</td>
<td>.001</td>
<td>12.92</td>
<td>.000</td>
<td>1.10</td>
<td>1.00 - 1.01</td>
</tr>
<tr>
<td>Block 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.001</td>
<td>.002</td>
<td>.18</td>
<td>.67</td>
<td>1.00</td>
<td>.998 - 1.00</td>
</tr>
<tr>
<td>Accurate Inventory 10</td>
<td>.07</td>
<td>.06</td>
<td>1.48</td>
<td>.22</td>
<td>1.07</td>
<td>.96 - 1.20</td>
</tr>
<tr>
<td>ND Log</td>
<td>1.10</td>
<td>.38</td>
<td>8.47</td>
<td>.004</td>
<td>3.01</td>
<td>1.43 - 6.31</td>
</tr>
<tr>
<td>Relational Accurate 10</td>
<td>1.46</td>
<td>.74</td>
<td>3.91</td>
<td>.048</td>
<td>4.32</td>
<td>1.01 - 18.37</td>
</tr>
</tbody>
</table>

Block 1: Nagelkerke $R^2 = .08$, $\chi^2 = 13.10$, df = 1, $p < .001$; Block 2: Nagelkerke $R^2 = .22$, $X^2 = 22.13$, df = 3, $p < .001$.

The combined models were run with Inventory 10 replaced by Expressive Vocabulary Type, and with ND replaced by Word Length, with the same results. The same analyses
were also conducted using the full set of direct-new-offers with partially different results. The full model was significant, $\chi^2(3, N = 514) = 29.34, p < .001$, explaining 24% of the variance (Nagelkerke $R^2$). After controlling for age, the variable ND was a significant predictor ($\chi^2(1, N = 164) = 11.35, p < .001$), but Relational Accurate 10 only approached significance ($\chi^2(1, N = 164) = 3.26, p = .071$). Lastly, the variable Accurate Inventory 10 was significant ($\chi^2(1, N = 164) = 4.92, p < .05$).

3.6 General perceptual or articulatory ease

One remaining and important issue is whether it is the child’s own production accuracy that is predictive of which words they imitate, or whether children’s imitation patterns reflect more general perceptual or articulatory ease. To test for this possibility, we coded the initial and final consonants of the direct-new-offers with the percentage of children who accurately produced the consonants at age 3, using the normative data from Smit et al. (1990). We then performed two additional logistic regressions to determine whether, after controlling for age, normative data for initial and final consonant accuracy was a significant predictor for whether children would imitate a direct-new-offer. There were no significant effects for either the initial or final consonant.

4. Discussion

This study analyzed children’s imitation of direct-new-offers from longitudinal production data. First, we found that there was an increase in the number of direct-new-offers that children imitated between 1 and 2 years of age. We then examined a variety of predictor variables: the child’s productive language maturation (Expressive Language Measures), phonological and lexical characteristics of the direct-new-offers (Target Word Measures), and phonetic (output) factors which were measures as the child’s production
experience with the sounds in the direct-new-offers (Target Word Production Experience). When analyzing the measures separately, all had significant predictors for children’s imitations. For the Expressive Language Measures, the size of the child’s consonant inventory was a significant predictor of imitation rates (though this could not be solely attributed to consonant inventories because it was highly correlated to vocabulary size). When examining the phonological and lexical properties of the target words (Target Word Measures), the target word’s ND was a significant predictor of children’s imitations (though again, ND was highly correlated to word length). In addition, the child’s own production experience with the sounds in the direct-new-offers (Target Word Production Experience), was a significant predictor of children’s imitations. When all variables were combined, Target Word and Target Word Production Experience were significant predictors, but not the measure of Expressive Language.

What emerges is that both the properties of a word and the child’s own accuracy in producing the words’ sounds, predicts whether or not the word is imitated. These results are in-line with previous findings which found that the child’s age, word length, and ND are all significant predictors of children productive vocabulary development (Stoel-Gammon 1998, Maekawa and Storkel 2006, Carlson et al. 2014). The findings are were different from Maekawa and Storkel (2006), who found that for one of the three children examined, a significant predictor for the child’s productive vocabulary development was the words’ final (but not initial) consonant accuracy at age 3, based on normative data from Smit et al. (1990). Our primary analyses differed from Maekawa and Storkel’s in that we examined the child’s own production accuracy with all of the consonants in the direct-new-offers, rather than based on normative data of the initial and final consonants.
When we analyzed the imitation using normative data for initial and final consonant accuracy, these were not significant predictors for whether children would imitate a direct-new-offer. Thus, our data suggest that it is the child’s own production experience with the sounds of a word that predicts the child’s imitation patterns. However, the effects of perceptual salience and ease of articulation on children’s production patterns is a question that is worth examining in more detail in the future.

Our findings are also in-line with studies of lexical selection and avoidance at the earliest stages of speech production (Ferguson and Farwell 1975, Leonard et al. 1979, Schwartz and Leonard 1982, Vihman et al. 2016, Vihman 2017). It has been suggested that these effects are strongest when children have a small productive vocabulary, i.e., under 50 words. It is possible that the effects found in our study were driven by the earliest stages, when children had fewer than 50 words in their vocabulary. To test for this possibility, we restricted the data to when the children’s productive vocabulary was over 50 words and redid the analyses. After controlling for age, we found that there were no longer any significant Expressive Language predictors on children’s imitation rates. However, we did find the same significant effects with the Target Word and Target Word Production Experience measures (also see Storkel, 2006 for an analyses of data beyond 50 words). Thus, the findings held that the properties of a word and the child’s own accuracy in producing the words’ sounds, were significant predictors of children’s imitations beyond the 50-word stage.

These findings can be captured by models in which representations have multiple dimensions, and in which phonetic details are stored bundled together with semantic and other lexical information. In exemplar and usage-based theories of speech perception and
speech production (e.g., Goldinger 1998, Pierrehumbert 2003, Sosa and Bybee 2008), the speaker keeps track of instances of a particular input to create an exemplar cloud. Each token leaves a unique memory trace, which includes detailed perceptual and conceptual information. In perception, traces are activated according to their similarity to the input stimulus. The average features of the traces with the highest activation provide the meaning to the incoming token (Goldinger 1998). In speech production, a trace is selected randomly from the exemplar cloud and the target for production is a weight of the surrounding exemplars (Pierrehumbert 2003). The assumption is that the production and perception system are closely tied together. Imitation has been conceptualized as pattern matching, which describes the learner’s ability to match and to detect differences between their vocal productions and the target (Olmsted 1971, Locke 1983, Vihman 1993). The recently developed A-map model (McAllister Byun et al. 2016) captures this notion of pattern matching. It predicts that children would be more likely to learn words that they are able to accurately produce, because these words would have the least amount of pressure for both matching the adult target and for the child to produce these words reliably (also see McAllister and Tessier 2016).

While naturalistic studies provide a large-scale overview of development, they may miss crucial data and may not contain certain words or structures. Future studies could be conducted to control the structures of the new words presented to children, while measuring which words are imitated; moreover, to measure whether the words that are imitated are subsequently learned better. This type of controlled approach is seen in the literature on the production effect, where a subset of words are produced by participants during training (akin to imitation) and another subset of words are not produced. In most
cases, adults’ and children’s production leads to better recall and recognition of words (e.g., MacLeod et al. 2010, Icht and Mama 2015, Zamuner et al. 2016). However, the effect of production is not always systematic. A reverse production effect has also been documented, where there is an advantage in recall and recognition for items that are not produced during training. The reverse production effect has been found both with adults and with children (Kaushanskaya and Yoo 2011, Zamuner et al. 2015). In sum, the effects of production appear to be subject to task-, attentional-, linguistic- and experience-related factors (Zamuner et al. 2017). Controlled studies would also allow us to address issues of individual variability. A shortcoming of this corpus analyses is that the data were not dense enough to analyze each child separately. To know how generalizable the results are, we want to know whether the effects described in this research are similar across children or whether their relative weighting varies by child. The same questions can also be asked about whether the results are generalizable across different languages. Lastly, it is important to point out that the selection of variables used in hierarchical logistic regression analyses can impact the results given that one has to decide which variables to include in the analyses (Reed and Yu, 2013). As this is one of the first studies to apply these analyses to examine phonological, lexical, and phonetic factors in children’s imitative speech, it can serve as a comparison for future research on language development.
References


Saffran, Jenny R., and Katharine Graf Estes. 2006. Mapping sound to meaning:


Appendix A: List of direct-new-offers between 0;11-2;11, found in Child Mental Lexicon

- alligator
- alphabet
- angel
- **antelopes**
- **ants**
- **arrow**
- arrow
- bacon
- **ball**
- balloon
- **barn**
- barn
- bass
- bat
- beak
- **bell**
- bell
- **bells**
- bill
- bird
- blanket
- bottles
- bow
- **boy**
- **bridge**
- bubbles
- bug
- bulldozer
- butterfly
- butterfly
- cabbage
- cake
- candle
- cardinals
- carrot
- caterpillar
- ceiling
- chair
- chair
- chef
- chef
- chick
- chimpanzee
- chipmunks
- **fork**
- ghost
- glove
- gloves
- goat
- goose
- goose
- grapes
- gray
- hair
- handle
- hat
- hay
- heels
- hen
- hen
- holly
- hook
- horns
- **hose**
- **hose**
- hose
- hunters
- icecream
- icecream
- irises
- iron
- iron
- jam
- jar
- keys
- kid
- **king**
- king
- **kite**
- kitten
- ladle
- lamp
- lawnmower
- lemon
- lettuce
- line
- **lips**
- lizard
- lobster
- loop
- magazine
- man
- map
- **Mars**
- mice
- mixer
- monster
- monsters
- moose
- mop
- mountain
- mustache
- mustache
- napkin
- needles
- nest
- net
- nightgown
- note
- octopus
- octopus
- oven
- pail
- pajamas
- pants
- parrot
- peach
- peaches
- peacock
- **pelican**
- penguin
- pepper
- piglets
- pilot
- pipes
- planet
- plant
- platform
- playdoh
- porcupine
- potato
- present
- puck
- rabbit
- rabbit
- rake
- rat
- reindeer
- rhinoceros
- rhinoceros
- ring
- rocket
- roller
- rollingpin
- rollingpin
- rooster
- rooster
- rug
- ruler
- ruler
- saddle
- sand
- scales
- scarf
- screen
- screwdriver
- seal
- seeds
- shadow
- shapes
- sheep
- shirt
- shovel
- sign
- skirt
- slippers
- smoke
- soldier
- spider
- spider
- sponge
- spring
- square
- squirrel
- stamps
- star
<table>
<thead>
<tr>
<th>steak</th>
<th>table</th>
<th>tomato</th>
<th><strong>vanilla</strong></th>
<th>wire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>steam</strong></td>
<td>tail</td>
<td><strong>tongue</strong></td>
<td><strong>vine</strong></td>
<td><strong>witch</strong></td>
</tr>
<tr>
<td>stem</td>
<td>tan</td>
<td>tricycle</td>
<td>violet</td>
<td>wolf</td>
</tr>
<tr>
<td>stocking</td>
<td>taxi</td>
<td>truck</td>
<td>wagon</td>
<td>woman</td>
</tr>
<tr>
<td>stool</td>
<td>teacher</td>
<td>trunk</td>
<td><strong>walrus</strong></td>
<td><strong>worm</strong></td>
</tr>
<tr>
<td>strap</td>
<td>tickets*</td>
<td>tunnel</td>
<td><strong>walrus</strong></td>
<td>wrench</td>
</tr>
<tr>
<td><strong>string</strong></td>
<td>tie</td>
<td>turkey</td>
<td><strong>walrus</strong></td>
<td>zebra</td>
</tr>
<tr>
<td>swing</td>
<td><strong>tiger</strong></td>
<td>turtle</td>
<td>whale</td>
<td>zebra</td>
</tr>
<tr>
<td>swing</td>
<td>tires</td>
<td>umbrella</td>
<td>wing</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates words that were found in the singular. Imitated words in bold.*