Cross-linguistic Patterns in Infant Babbling

Andreea Geambașu, Mariska Scheel, and Clara C. Levelt

1. Introduction

Infants begin to vocalize very soon after birth, and they begin to babble about six months after they are born (Oller, 1980). The babbling stage is distinct from the previous phase of vocalizations in that sounds – or gestures in infants acquiring sign language – are now clearly organized in a syllabic structure. As such, these utterances are the infant's first linguistic productions.

In the works of Stark (1980) and Oller (1980), two stages were identified within the babbling phase. Babies start with reduplicated babbling when they are six to eight months old, and progress into "variegated" (Oller, 1980) or "non-reduplicated" (Stark, 1980) babbling at 10 to 12 months. Work by Koopmans-van Beinum and van der Stelt (1986) outlines a similar line of development, with reduplicated babbling beginning at six months and lasting up until at least 12 months. They do not identify a specific non-reduplicated stage during this period. In addition, Roug, Landberg, and Lundberg (1989) also identified babbling stages similar to those proposed by Oller and Stark, with reduplicated (consonant) babbling beginning at seven months, and variegated babbling beginning at 12 months. The stages identified by these researchers differ only slightly. Where they crucially converge is on the consensus that infants begin their babbling at around six to eight months old, that they begin with reduplicated utterances, and that they transition into producing variegated utterances at around 10 to 12 months. The existence of these two stages has been disseminated in introductory linguistics textbooks for years (e.g., Hoff, 2008). Yet, the data on which these particular stages have been based came from relatively small-scale studies and from primarily English-acquiring infants.

In the past decades, some studies have focused on finding out whether the proposed stages still hold when data from a larger number of infants is taken

[1] Andreea Geambașu, Leiden University Centre for Linguistics, and Leiden Institute for Brain and Cognition, Corresponding author: a.geambasu@hum.leidenuniv.nl. Mariska Scheel, Leiden University. Clara C. Levelt, Leiden University Centre for Linguistics, and Leiden Institute for Brain and Cognition.

We are grateful to all researchers contributing to the PhonBank database and to Greg Hedlund and Yvan Rose for their help in scripting and implementation in Phon. We also thank Bingjing Gu and Shenmin Wang for their help with statistical analyses. This work was supported by the Netherlands Organisation for Scientific Research (NWO) Vrije Competitie grant number 360.70.452 to Clara C. Levelt.

into account, and whether they take the same form across languages. Smith, Brown-Sweeney, and Stoel-Gammon (1989) in one notable example followed 10 English-learning infants from six to 18 months old, at three four-month intervals. They found that while reduplicated utterances occurred more frequently in the first two age groups (six to nine months and 10 to 13 months respectively), both reduplicated and variegated utterances co-occurred across all age three groups. Crucially they did not find a change with respect to the frequency of reduplicated versus variegated babbling at the time when this change was expected according to the above-mentioned literature, namely between the first and second age groups. More recently, a study comparing songbird and infant babbling by Lipkind et al. (2013) analyzed data of nine English-acquiring infants from the Davis corpus (Davis & MacNeilage, 1995) and showed that there was no difference between the ages of 8- through 29-months old in American English-learning infants with respect to the amount of reduplicated babbling found. Thus they found no discrete stage of reduplication followed by one of variegation. Instead, what they found was that infants and birds both tended to reduplicate newly-acquired sounds more than other sounds.

While Lipkind and colleagues present a crucial first step in a larger-scale approach by making use of the data available in the CHILDES database (MacWhinney, 2000), a wealth of previously-collected data that can give us an even better understanding of babbling has remained unexplored up until now. For the present study, we collected and systematically analyzed available corpus data from different languages in order to finally determine unequivocally what the universal developmental babbling patterns are, and to assess language-specific influences on babbling. With this work, we thus present the first large-scale, cross-linguistic survey of the babbling patterns of infants.

Our hypothesis was that if there are distinct developmental stages with respect to babbling patterns (Oller, 1980; Stark 1980; see de Boysson-Bardies, 2001, Hoff, 2008, Vihman, 1996, etc. for reviews), we should see predominantly reduplicated babbling in the period of seven to ten months and a switch to predominantly variegated babbling in the period of ten to 12 months. An alternative hypothesis, based on the findings of Koopmans-van Beinum and van der Stelt (1986), and Roug et al. (1989), was that we should see predominantly reduplicated babbling throughout the entire period of seven through 12 months old. This would indicate that variegated babbling only truly begins later, with the ontogeny of the lexicon. The aim of this work is thus to pinpoint if and where a transition from one stage to another may occur. In addition, by conducting a cross-linguistic analysis, our aim is to determine which aspects of babbling development are universal, and to identify language-specific variation.

2. Language and Corpora Selection

The present study was carried out on datasets of child utterances from different languages, made available through the PhonBank corpus (Rose & MacWhinney, 2014). The datasets that were suited for our purpose were collected and transcribed by seven individual principle investigators for the purpose of their own varied research goals. The transcriptions have since been made available by these researchers to be used and analyzed by others with the aid of the Phon software (Rose et al., 2006; Rose & MacWhiney, 2014).

From all the available datasets (n=33, from 12 different languages) represented in PhonBank, we selected only those datasets that included phonetic transcriptions in the International Phonetic Alphabet (IPA) and that included child utterances recorded between our target age range of seven through 12 months. The lower limit of seven months was chosen, as this was the first age group for which phonetically transcribed data in at least one language was available, and, from a theoretical point of view, it is the lower end of the age range at which infants begin to babble. The upper limit was chosen because it is the age at which we expect variegated babbling to either begin to appear or even to overtake reduplicated babbling.

Table 1. Languages studied, the names of the corpora from which the data was taken, the country of recording, and the number of infants per age group represented in each language.

			Age in Months					
Language	**Corpus**	**Country of Recording**	**7**	**8**	**9**	**10**	**11**	**12**
Arabic	Kern	Tunisia	0	3	3	4	4	3
Dutch	Zink	Belgium	0	4	4	4	4	4
English	Providence Stanford	USA	4	6	10	17	18	19
French	Kern Lyon Stanford	France	1	2	5	8	9	12
Japanese	Ota Stanford	USA	0	0	1	1	0	3
Portuguese	Freitas	Portugal	0	0	0	1	5	6
Romanian	Kern	Romania	1	4	4	4	4	3
Swedish	Stanford	Sweden	0	0	5	2	2	2

Datasets from eight different languages – Arabic (Kern, Davis & Zink, 2009), Dutch (Kern et al., 2009), English (Providence corpus, Demuth, Culbertson, & Alter, 2006; Stanford corpus, de Boysson-Bardies & Vihman, 1991), French (Kern corpus, Kern et al., 2009; Lyon corpus, Demuth & Tremblay, 2008; Stanford corpus, de Boysson-Bardies & Vihman, 1991), Japanese (Ota corpus, Ota, 2003; Stanford corpus, de Boysson-Bardies & Vihman, 1991), Portuguese (Freitas, 1997), Romanian (Kern et al., 2009), and Swedish (de Boysson-Bardies & Vihman, 1991) – fit the criteria described above. Table 1 lists the languages, the corpora from which those languages were extracted, and the number of infants represented at each age group between seven and 12 months (per month, per corpus). Note that some languages are represented by more than one corpus, and that some languages are not represented at all ages.

The data included recordings from monolingual households. In most cases, the infants were growing up in their country of origin. The Japanese infants were recorded in the United States (San Francisco in the case of the Stanford corpus, Washington DC in the case of the Ota corpus), but were raised in monolingual Japanese households.

3. Corpus Data Collection

In order to identify and analyze reduplicated and variegated patterns, the data had to be processed in such a way that each syllable was identified, labeled, and compared to the other syllables in the utterance. In this way syllables could be identified as the same (reduplicate) or different (variegated). Our interest was not in the kind of syllables that were being reduplicated or used in variegated sequences, but in the extent to which reduplication and variegation was present. Thus after syllables in each utterance were identified, they were labeled in an abstract way (assigned a unique letter within each utterance). Then, each type of pattern was tallied for analysis. This process was done with the help of two scripts that will be described in more detail below.

The first step that was necessary to identify the patterns in infant utterances was syllabification. This was done with Phon's built-in syllabifier. Syllabification was applied to all records of each language. For all languages except for two, a language-specific syllabifier was available that syllabifies according to the rules of the language. The exceptions were Romanian and Japanese. For these languages, the English syllabifier was used, and complex utterances (eg., those with consonant clusters and geminates, longer utterances, etc.) were manually inspected. No odd or incorrect syllabification was found.

For the next steps, namely abstraction from the phonetic content of the syllable and comparison of two subsequent syllables, two scripts were written by Greg Hedlund and Yvan Rose and added to the Online Scripts Library in Phon (Phon 2.1, Rose et al., 2006). The first script, called SyllablePatternListing (Hedlund & Rose, 2015, https://www.phon.ca/phontrac/wiki/scriptlibrary), identifies the first syllable of an utterance, assigns it a letter (A), moves on to the second syllable, and performs an identity comparison with the previous syllable. If the first and second syllables are identical, the second syllable is also labeled A, yielding a reduplicated AA pattern. If the first and second syllables are not identical, the second syllable is labeled B, yielding a variegated AB pattern. For utterances of longer length, the identity matching is performed on each syllable as described above, and the comparison made with each of the previous syllables. Table 2 shows the possible patterns identified by this script. A second script named SyllablePatternSummary, tallied each of the patterns identified per infant.

Table 2. All possible combinations of syllables in two-, three-, and four-syllable utterances, categorized roughly as reduplicated or variegated by their predominant pattern. Bold pairs of syllables are reduplicated. Underlined pairs of syllables are variegated.

	Utterance length		
Utterance Type	Two-syllable	Three-syllable	Four-syllable
Complete Reduplication	**A A**	**A A A**	**A A A A**
Combination - 3 reduplications			
Initial		**A A** B	**A A A** B
Final		A **B B**	A **B B B**
Internal			**A A B B**
Combination - 2 reduplications			
Initial			**A A** B C
			A A B A
Final			A B **C C**
			A B **A A**
Internal			A **B B** C
			A **B B** A
Complete Variegated	A B	A B A	A B A B
		A B C	A B A C
			A B C B
			A B C A
			A B C D

4. Results

Our data collected from Phon consisted of 42,348 total utterances of one-, two-, three- and four-syllable utterances from a total of 64 infants. Utterances longer than four syllables were uncommon, and, from a practical point of view, beyond the scope of the current exploratory work. Of the one- through four-syllable utterances, 47% were one-syllable utterances (1SU), which cannot inform us about syllable patterns. We thus focused our analysis primarily on two- (2SU, 38%), three- (3SU, 11%), and four-syllable utterances (4SU, 4%). Figure 1 illustrates how frequently infants in each language produced each type of utterance, as well as the means across languages. Note that the longer the utterance, the less frequently it is represented across almost all languages, the exception being Portuguese, where 2SU (57.6%) are more frequent than 1SU (25.4%). In Dutch, English, French, and Romanian on the other hand, 1SU made up 60% and above of collected utterances.

In our primary analysis we will focus on the 2SU and 3SU. Because of their complexity (see Table 2), 4SU will be discussed separately (see section 5).

Across languages and across our age groups, the distribution of 2SU was 18.7% reduplicated and 81.3% variegated. When all age groups were pooled together, Romanian had a higher instance of reduplicated 2SU than other languages (30.3%), while English had a notably low rate of reduplication (6.2%). Figure 2 shows the distribution of reduplicated (AA) and variegated (AB) utterances in each language.

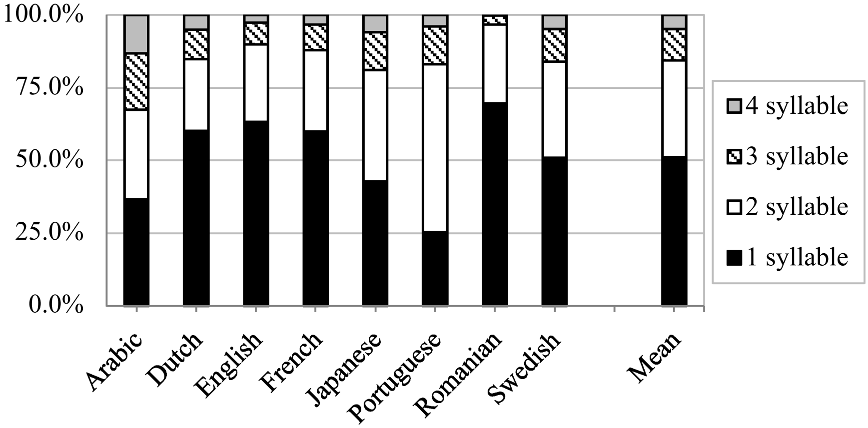

Figure 1. Percentage of syllables represented in each language, and the mean of all languages.

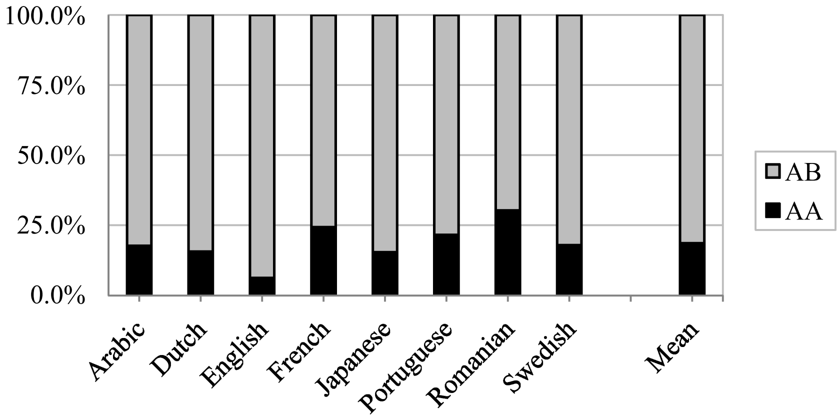

Figure 2. Distribution of two-syllable utterances across all ages following a reduplicated pattern (AA) and a variegated pattern (AB), in each language individually, and a mean of all languages.

As shown in Table 2, 3SU can be broken down into five different patterns: completely reduplicated (AAA), completely variegated (ABC), or containing some element of both reduplication and variegation (ABB, AAB, and ABA).

Figure 3 presents these patterns separately per language, pooling together all age groups. It is immediately evident that the overwhelming majority of utterances fell into the completely variegated pattern (ABC). Across all languages, this pattern represented a mean of 65.5% of all three-syllable utterances. In English, 83.3% of 3SU were completely variegated, while only 3.2% were completely reduplicated. All languages showed at least 50% completely reduplicated utterances in the distribution of the patterns (Romanian was at 50%, while all other languages were at approximately 65%).

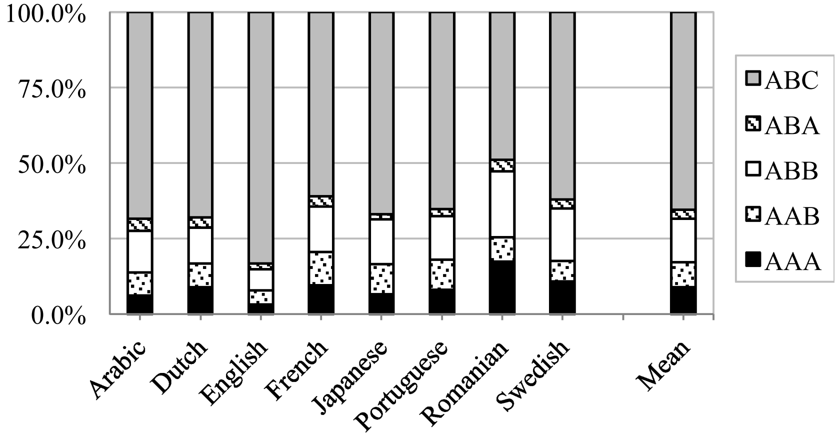

Figure 3. Distribution of three-syllable utterances following a totally reduplicated pattern (AAA), a totally variegated pattern (ABC) and patterns containing some intermediary level of reduplication (ABB, AAB, ABA).

The initial picture of the distribution of patterns across all age groups of 2SU and 3SU clearly shows that across languages, within the seven through 12-month-old range, infants actually tend to produce more variegation than reduplication. We now turn to answering our main question: can we differentiate an early stage with mostly reduplicated babbling and a late stage with mostly variegated babbling in our large, cross-linguistic dataset?

4.1. Analysis

4.1.1. Two-Syllable Utterances

Although our total dataset included 65 infants across eight languages, and all of these infants are presented in the figures for complete visual inspection (see Figure 4), for the purpose of a proper statistical analysis, some pruning of the data had to be done. Of the 65 infants, 14 individuals contributed only one month worth of data. We removed those infants who were not followed longitudinally in their respective studies, leaving a total of 51 infants. Of these 51, we find that only six infants (12%) in three languages have data in the seven-month range. This month is thus also removed from analysis.

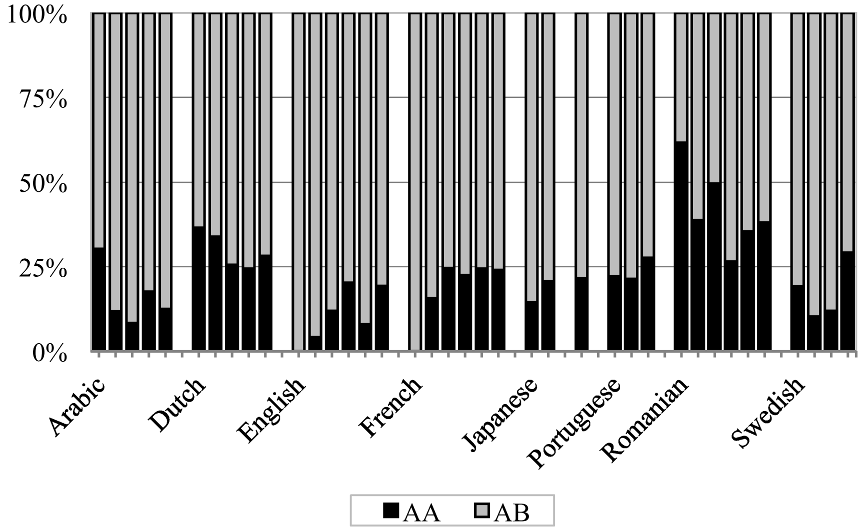

Figure 4. Development of two-syllable utterance patterns from seven to 12 months old across all languages. Y-axis label indicates the beginning of the time course for the language.

As the original recordings were collected by several different researchers for different research purposes, they also varied with respect to their frequency of collection. Some of the corpora contained multiple recording sessions per month while others included only one session. In order to make the data uniform with the respect to the number of data points per child, the tallies of multiple sessions within one month were averaged. In this way, each child now had one datapoint per month, indicating how many utterances of each type he or she produced on average in that month.

This dataset was analyzed with a generalized estimating equation [GEE; geeglm in R, R Development Core Team (2015).]. GEEs, as opposed to linear mixed models, do not attempt to model the individual variability in the data, and can thus tell us more about the mean performances. The analysis was run per pattern, with fixed effects of month and language. The factor of month was centered at 10 months to avoid correlation between estimates of the intercept and the slope. In addition, English was chosen as the reference language as this language had the most data.

On pattern AA, no interaction between month and language was found (X^2=3.2, df=6, p=0.78), and no effect of month was found (X^2= 0.07, df=1 p=0.79). An effect of language was found (X^2=29, df=6, p<0.001). As no interaction was found, this factor was removed, and a second model was run.

Here too we found no effect of month (X^2=8.06, df=6, p=0.09), and an effect of language (X^2=29, df=6, p<0.001). A Wald test reveals that Dutch and Romanian differed significantly from English (Wald=5.46, p=0.02 for Dutch; Wald=14, p<0.001 for Romanian). Both languages produced significantly more AA patterns than English (see Figure 2).

The same analysis was run on the AB pattern and while no interaction between month and language was found (X^2=5.24, df=6, p=0.5), a significant effect of month (X^2=4.25, df=1, p=0.04) and a significant effect of language (X^2=17.26, df=6, p=0.008) were found. A second model with the interaction effect removed yielded identical effects for month and language respectively. Although significant effects of month and language were found, in the initial model, further inspection of the details with a Wald test revealed no effect of month (Wald=3.39, p=0.07), and no individual language differed from English.

4.1.2. Three-Syllable Utterances

All 3SU were analyzed in the same way as 2SU. Their development can be inspected visually in Figure 5.

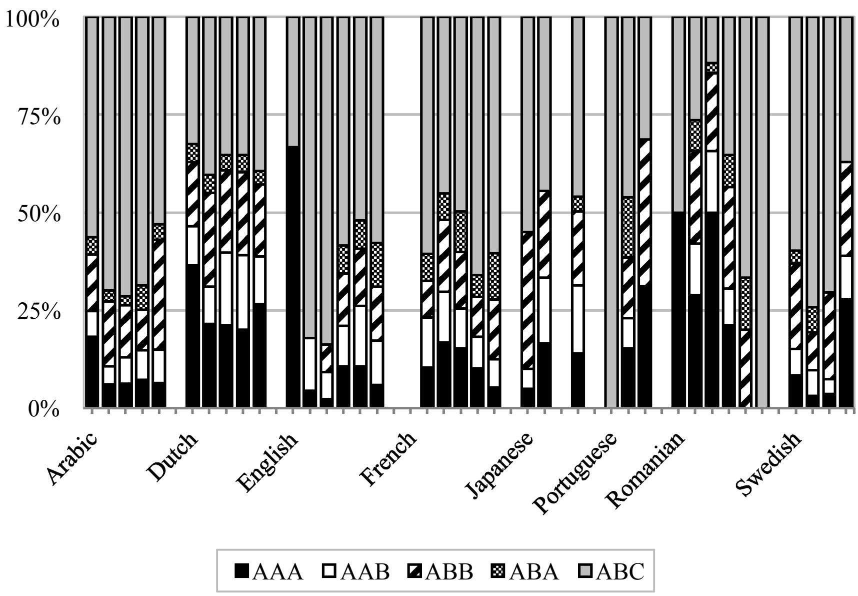

Figure 5. Development of three-syllable utterance patterns from seven to 12 months old across all languages. Y-axis label indicates the beginning of the time course for the language.

For pattern AAA, no interaction of month and language was found (X^2=9.68, df=6, p=0.14) and no effect of month was found (X^2=0.79, df=1, p=0.37. An effect of language was found (X^2=12.57, df=6, p=0.05). A second model that excluded the interaction effect was run, confirming the presence of the language effect (X^2=12.71, *df*=6, p=0.05). Wald tests reveal that only Dutch differed significantly from English with respect to this pattern (Wald=8.94, p=0.003), with higher rates of AAA overall.

In pattern AAB, there was a significant interaction between month and language (X^2=14.5, df=6, p=0.025) and a significant effect of language (X^2=115.8, df = 6, p<0.001). However, no effect of month was found (X^2=0.1, df=1, p=0.772). Dutch (Wald=11.11, p<0.001) and Romanian (Wald=4.94, p=0.02) were significantly different from English for this pattern. In addition, this model also found an interaction between language and month in French (Wald=5.85, p=0.02) and Romanian (Wald=5.24, p=0.02), both showing less of this pattern over time. Particularly in Romanian, this pattern largely disappears after 10 months.

For pattern ABA, there was an interaction between month and language (X^2=15.8, df=6, p=0.02), and an effect of language (X^2=21.5, df=6, p=0.002), but no effect of month (X^2=2.1, df=1, p=0.15). Only Portuguese differed significantly from English over time (Wald=8.48, p=0.004), with Portuguese babies producing this pattern at 11 months but not at 10 or 12 months.

For pattern ABB, there was no interaction between month and language (X^2=11.3, df=6, p=0.08), or of month (X^2=0.8, df=1, p=0.39). However an effect of language was found (X^2=114.3, df=6, p<0.001). As no interaction was found, the interaction effect was removed from the model, and the same language effect was found. Arabic (Wald=79.37, p<0.001), Dutch (Wald=26.55, p<0.001), French (Wald=12.94, p<0.001), and Swedish infants (Wald=4.63, p=0.03) were found to have significantly higher production rates of this pattern than infants acquiring English.

Finally, for pattern ABC, there was no interaction between month and language (X^2=3.2, df=6, p=0.78), and no month effect (X^2=0.7, df=1, p=0.39). However a language effect was found (X^2= 136.4, df=6, p < 0.001). With the interaction effect removed, the effect of the month and language did not change. Arabic (Wald=14.94, p<0.001) and Romanian (Wald=17.56, p<0.001) differed with respect to this pattern from the baseline English, with Arabic showing significantly more ABC productions than English, and Romanian significantly fewer.

5. Four-Syllable Utterances

In comparison with shorter utterances, 4SU are both more numerous and more complex with respect to the categorization of patterns. Therefore, they are not analyzed in the same way as the 2SU and 3SU. Table 2 shows how the utterances might be categorized as including both reduplication and variegation. It is not straightforward whether to characterize certain utterances as either reduplicated or variegated as they fall under both categories. For the scope of the present work it thus suffices to present an overview of the results.

Figure 6 shows the distribution of all possible 4SU over the eight languages. There are 15 possible patterns. Visual inspection of the figure shows that the pattern with the highest representation across most languages is the completely variegated ABCD. The exceptions are Portuguese and Romanian, which have more instances of the ABBB pattern. In Portuguese, the AAAA pattern is not found at all, while the ABAA pattern occurs 17% of the time. This pattern is found rarely or not at all in all other languages. Dutch also stands out as having the highest frequency of the AAAA pattern, at just under 25% vs. 7.14% in English and 6.33% in French.

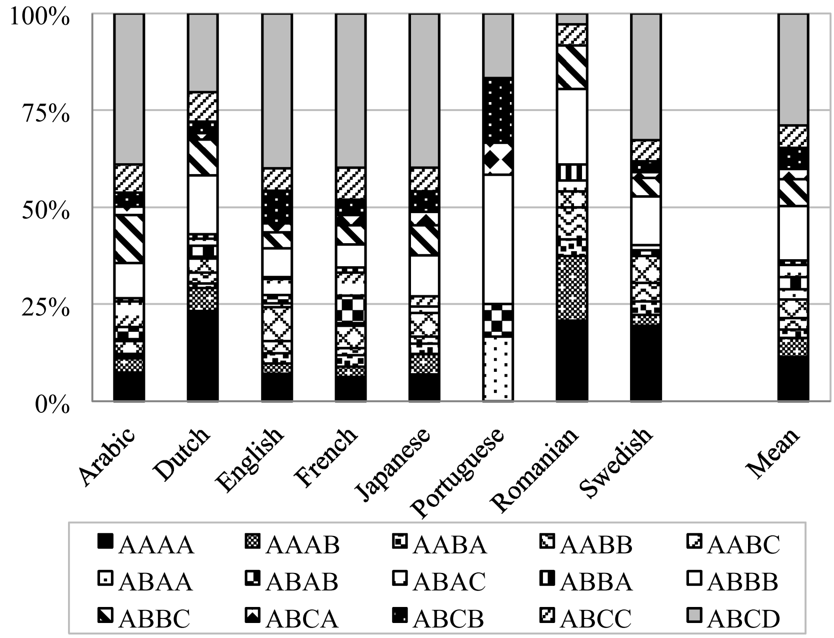

Figure 6. Distribution of four-syllable utterances following a totally reduplicated pattern (AAAA), a totally variegated pattern (ABCD) and patterns containing some intermediary level of reduplication.

In order to get a better idea of the development of 4SU, we grouped the 15 possible patterns into larger categories, depending on whether reduplication was present and where this reduplication occurred in the string. Figure 7 shows the grouping of the patterns into six categories, and their frequencies across the represented age groups. When considering any pattern containing reduplication as separate from a full variegation pattern, we can clearly see that full variegation is well represented in most languages. In Arabic, English, French, and Swedish, full variegation thus accounts for more than 50% of the utterances in most age groups (although English, Japanese, and Swedish have fewer variegated utterances over time, and Swedish contains no full variegation at all in month 12). Dutch on the other hand contains approximately 25% variegated utterances, with most of the 4SU productions containing a reduplicated syllable. Notably Arabic, Dutch, and French do not show a change over time, with their frequency bars remaining constant across age groups.

Thus with the 4SU too, while we did identify some language specific differences, no evidence was found for two distinct stages, nor for a development towards increasingly more variegation.

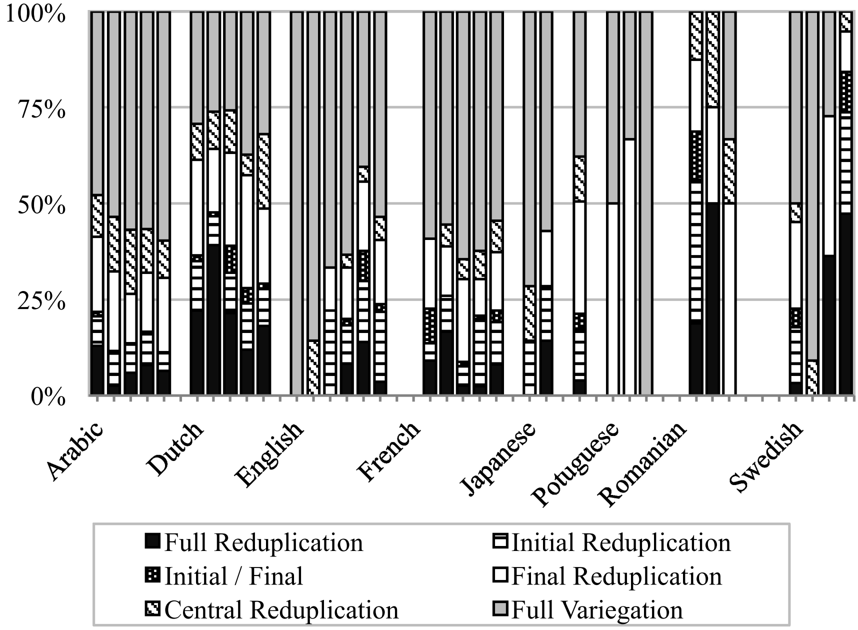

Figure 7. Development of four-syllable utterance patterns, grouped by utterance type. Each bar indicates data from one month, from seven to 12 months, across all languages. Y-axis language label indicates location of first age group for that language.

6. Discussion

The present study presents the first attempt at a large-scale and cross-linguistic analysis of infant babbling utterances. Taken together, our results from the 2SU, 3SU, and 4SU show that there is no evidence for clear stages of development across the babbling period with respect to reduplication versus variegation. Month was never an explanatory factor in our models for 2SU and 3SU, while for 4SU, no clear age effects could be discerned either. All age effects interacted with language to show that in certain languages, certain specific patterns (such as AAB) are produced more or less over time. However, in no language did we find the predicted line of development in which reduplicated babbling is most prominent initially and is either gradually or suddenly overtaken by variegated babbling. Instead we found that variegated babbling is actually the most prominent pattern from the earliest stages of babbling. This was the case in most languages in 2SU, 3SU, and 4SU.

Our data showed language effects on each individual pattern analyzed, showing that there are some important differences between languages with respect to babbling. For example, Dutch differed significantly from English, a closely related language, with respect to the production of completely reduplicated utterances like AA and AAA. Dutch infants had significantly more such utterances than English-acquiring infants. Such between-language differences raise many specific questions related to the connection between each language and the frequency of various patterns. This presents an interesting point of focus for future work. The linguistic profile (the phonology and the morphology) of the native language may have an important role in shaping the types of babbling patterns infants produce and the distribution of these patterns.

Finally, the work presented here may also be continued and expanded by identifying exactly the types of syllables making up reduplicated and variegated utterances. For example, is the variegation or reduplication taking the form of a consonant change, a vowel change, or both? Does this change over time, and does it differ between languages? Both reduplication and variegation may take many forms. In the present work, the patterns were identified on the basis of exact identity, yet this is another area where different language backgrounds may yield different results.

Our large-scale analysis has thus shown that while the types of patterns infants produce differ depending on their language background, the idea of discrete stages may be abandoned in favor of more fruitful research into explaining particularities of babbling patterns in different language backgrounds.

References

de Boysson-Bardies, Bénédicte, Hallé, Pierre, Sagart, Laurent, & Durand, Catherine. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, *16*(01), 1-17.

de Boysson-Bardies, Bénédicte (2001). How language comes to children: From birth to two years. (M.B. DeBevoise, Trans.) Cambridge, MA: MIT Press. (Original work published 1996).

de Boysson-Bardies, Bénédicte & Vihman, Marilyn M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, *67*, 297-319.

Davis, Barbara L. & MacNeilage, Peter F. (1995). The articulatory basis of babbling. *Journal of Speech Language and Hearing Research,* 38, 1199–1211.

Demuth, Katherine, Culbertson, Jennifer, & Alter, Jennifer. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech*, *49*, 137–174.

Demuth, Katherine & Tremblay, Annie. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, *35*, 99–127.

Freitas, Maria João. (1997). Aquisição da Estrutura Silábica do Português Europeu. Ma, Universidade de Lisboa.

Hoff, Erika. (2008). *Language Development*. Belmont, CA: Wadsworth/Thomson Learning.

Kern, Sophie, Davis, Barbara L., & Zink, Inge. (2009). From babbling to first words in four languages: Common trends, cross language and individual differences. In F. d'Errico & J. M. Hombert (Eds.), *Becoming eloquent: Advances in the Emergence of language, human cognition and modern culture* (Vol. *6*(289), pp.205–232). John Benjamins' Publishing Company.

Koopmans-van Beinum, Florien J., & van der Stelt, Jeannette M. (1986). Early stages in the development of speech movements. In B. Lindblom & R. Zetterström (Eds.), *Precursors of early speech*. Basingstoke, Hampshire: Macmillan Press.

Lipkind, Dina, Marcus, Gary F., Bemis, Douglas K., Sasahara, Kazutoshi, Jacoby, Nori, Takahasi, Miki, Suzuki, Kenta, Feher, Olga, Ravbar, Primoz, Okanoya, Kazuo. & Tchernichovski, Ofer. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*, *498*, 104–108.

MacWhinney, Brian. (2000). The CHILDES Project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Ota, Mitsuhiko. (2003). The development of prosodic structure in early words: Continuity, divergence and change. Amsterdam/Philadelphia: John Benjamins.

R Development Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Rose, Yvan, MacWhinney, Brian, Byrne, R., Hedlund, Greg, Maddocks, K., O'Brien, P., & Wareham, Todd. (2006). Introducing Phon: A Software Solution for the Study of Phonological Acquisition. In D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 489–500). Somerville, MA: Cascadilla Press.

Rose, Yvan, & MacWhinney, Brian. (2014). The PhonBank Project: Data and Software-Assisted Methods for the Study of Phonology and Phonological Development. In J. Durand, U. Gut & G. Kristoffersen (Eds.), *The Oxford Handbook of Corpus Phonology* (pp. 308–401). Oxford: Oxford University Press.

Roug, L., Landberg, I., & Lundberg, L.-J. (1989). Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. *Journal of Child Language*, *16*(01), 19-40.

Smith, Bruce L., Brown-Sweeney, Sharon, & Stoel-Gammon, Carol. (1989). A quantitative analysis of reduplicated and variegated babbling. *First Language, 9*, 175–19.

Stark, Rachel E. (1980). Stages of speech development in the first year of life. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 1, pp. 73–90). New York: Academic Press.

Vihman, Marilyn M. (1996). *Phonological development: The origins of language in the child*. Blackwell Publishing.

Proceedings of the 40th annual Boston University Conference on Language Development

edited by Jennifer Scott
and Deb Waughtal

Cascadilla Press Somerville, MA 2016

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-066-1 (2 volume set, paperback)
ISBN 978-1-57473-166-8 (2 volume set, library binding)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com