UC Berkeley

UC Berkeley Previously Published Works

Title

Child phonological responses to variegation in adult words: A cross-linguistic study.

Permalink

Authors

Vihman, Marilyn May
Ota, Mitsuhiko
Keren-Portnoy, Tamar
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Running head: Child responses to variegation

Child phonological responses to variegation in adult words: A cross-linguistic study

Marilyn May VIHMAN[1, 2], Mitsuhiko OTA[3], Tamar KEREN-PORTNOY[1], Shanshan LOU[1] and

Rui Qi CHOO[1]

[1] University of York

[2] University of California Berkeley

[3] University of Edinburgh

Address for correspondence:

Marilyn Vihman

Language and Linguistic Science

University of York

Heslington YO10 5DD

marilyn.vihman@york.ac.uk

Key words: phonological development, word production, variegation, reduplication, consonant harmony

Abstract

Variegation – the presence of more than one supraglottal consonant per word – is a key challenge for children as they increase their expressive vocabulary toward the end of the single-word period. Here we consider the prosodic structures of target words and child forms in English, Finnish, French, Japanese, and Mandarin to determine whether children learning these languages respond similarly to the challenge or instead differ in ways related to the phonological structure of the adult language. Based on proportional occurrence of each structure, we find that the word forms of children learning Mandarin and Japanese show more variegation than do those of children learning the European languages, although their target words do not; proportions of reduplication, consonant harmony and single-consonant words also differ by language. We conclude that experience with the structure of the language – and thus representation, as well as immature articulatory skills – shapes children's responses to variegation.

Variegation – or the presence of two different supraglottal consonants in one lexical unit – is a central challenge for children in the early word learning period. The difficulty of producing words containing differing consonants (other than glottals or glides, which are present from an infant's earliest vocalizations) is evidenced in the fact that children attempt few such words when they first begin to talk and produce such forms even less (Menn & Vihman, 2011). As the child's expressive vocabulary grows, the bias in favour of producing only a single supraglottal consonant per word form continues to be in evidence.

On the one hand, the children continue to target many words that include only a single consonant. For example, in the sessions analysed for this study, from the end of the single word period, seven of the 15 English-learning children produced *baby* (as [beɪbi; beːbi; pebi; beɪːbiː, beɪʔbiç; beɪbeɪːi; bəbi; biːbiːç]). On the other hand, to meet the challenge of expanding the types of words they can produce, children typically accommodate this bias by either partial or full reduplication, consonant or syllable omission or replacement of consonants by glottals or glides. To illustrate, drawing on our English data, examples of partial reduplication (consonant harmony) from our sample include *all gone* [guːgɒː], *bagel* [bʌbu], *circle* [geːgə], *doggy* [gɒgiː], *ladder* [dɛːdəʰ], *lady* [jɛːji] and *piggie* [pɛpi]. Full reduplication is less common in English but can also be observed, in both target words (*bye-bye, choochoo, mama, walk-walk, woofwoof*) and child forms adapting adult words to this prosodic structure: *cookie* [kɛkɛ], *glasses* [kæ̃kæ̃], *picture* [p'œp'œ], *snowman* [mɪmɪː], *balloon* [bɛbɛ], *bee* [bʷiːbʷiːː]. Note that some adaptations to reduplicated form involve repeating a monosyllable (e.g., *bee*), while others involve changing one or both vowels (in words whose target form already has consonant harmony, such as *cookie*), and yet others (for target forms without harmony) involve replacing

one of the syllables by repeating the other (*glasses*) or producing two identical syllables that resemble one of the target syllables but accurately match neither (the remaining examples). Finally, we also see consonant omission (with metathesis: *Nicky* [ɪnːi]), glottal substitution (*sofa* [ʔəfːæ]), and gliding (*lady* [lɛji]), as responses to variegation in target words, which suggests the difficulty of accessing and producing such forms.

We will briefly review previous research on the processes that result in child production of word forms that lack variegation in response to variegated targets. The omission of target segments or whole syllables and the inclusion of non-target-based glottals or glides is widely recognized and has been illustrated in studies of individual children (see. for example, diary data from children acquiring Estonian, French, German and Hindi as well as English in Vihman & Croft, 2007) or of developmental phonology more generally (e.g., Grunwell, 1982, provides a useful overview of the typical chronology of phonological processes in English, including final consonant omission, which is commonly seen up to age two years.)

Reduplication in child language has so far received relatively little attention, especially from a cross-linguistic perspective; early studies based on English include Fee and Ingram (1982), Ferguson (1983), and Schwarz, Leonard, Wilcox and Folger (1980). In contrast, consonant harmony (CH), which describes a process in which the consonantal features of one syllable 'spread' to another within the same word, has been widely reported and discussed (e.g., Levelt, 2011; Lleó, 1990; Smith, 1973; Stoel-Gammon & Stemberger, 1994). Most studies of consonant harmony in child forms have featured only languages spoken in Europe. However, one early cross-linguistic account analysed the extent of CH in three children learning Cantonese in addition to one to three children learning each of five European languages (Vihman, 1978); the

4

data covered a wide range of ages and periods of data collection. In that study the Cantonese differed from the European data in its low level of both variegated target forms and child use of CH. These findings, though limited by the small number of children per group as well as the uneven sampling, suggest that, despite its common occurrence, CH might not be a 'universal tendency' of child phonology (Smith, 1973). If the occurrence of such common processes is not universal, it is worth asking how the phonological structure of different languages may guide children to different solutions to the variegation challenge.

Evidence of the nature and extent of deployment of these solutions in different languages may also provide insight into two opposing hypotheses that have been proposed as to the type of challenge variegation poses. Some have heavily emphasized the role of motoric advances or the maturation of articulatory skills. For example, Davis, MacNeilage and Matyear (2002) present evidence and argument to support their Frame and Content model, which sees early child word forms as based on the motoric principles that underlie canonical babble. Similarly, McAllister Byun, Inkelas and Rose (2016) have argued for articulatory limitations as the primary factor shaping child outputs in the first years of word use.

In contrast, Aitchison and Chiat (1981) tested their intuition that (long-term) memory (or representation) is likely a key source of adult-child discrepancies in early word forms by carrying out an experiment in word production with 4- to 9-year-olds. They found that, in naming novel picture-book images as they recurred, these older children produced errors that resembled those of the early-word period, such as [tu:du:] for *kudu* and [kəku:n] for *racoon*, both with consonant harmony; articulatory limitations are unlikely to have been involved at these ages.[1]

---

1 The first author, upon first encountering the name of an old Japanese road between Kyoto and Tokyo (the Nakasendo Way), initially recalled the word as /nakendo/ (with loss of some medial

More recently, Hodges, Munro, Baker, McGregor, Docking and Arciuli (2016) studied the effects on longer-term learning (in two-to-three-year-olds) of accuracy in initial imitation of nonwords, testing immediately, for short-term memory effects, and 5 minutes later and then again 1 to 7 days later, for longer-term effects. The results showed that accuracy of imitation had an immediate effect only (i.e., on short term memory) and that 'expressive phonology' (based on standard phonological skill test scores) was unrelated to performance, whereas expressive vocabulary – or long-term memory for a range of different word forms – 'predicted performance at all time intervals' (p. 457). This again provides evidence that memory or representation plays a role in constraining word learning in development.

Finally, cases of metathesis in early child word forms, though rare, suggest representational, not articulatory challenges, since the child in such cases is able to produce the target sounds (with or without voicing change), but appears to lack a robust memory for their place in the word. We find just 17 such child forms out of the 991 different child disyllabic variants we coded (2%). Examples include Finnish *jalka* 'foot, leg' [gala], Japanese *kiɕa* 'train' [ciga], Mandarin *ʈʂuo1tsi0* 'table' [tɤ1ʈʂai4]. Despite their infrequency of occurrence these examples suggest that long-term memory for word forms (and emergent pattern preferences, or templates: Vihman, 2019) must constitute at least one aspect of child responses to the challenge of targets with more than one supraglottal consonant.

Here we consider phonological data from five languages – English, Finnish, French, Japanese and Mandarin – that contrast sharply in their prosodic structures and accentual patterning. We will be interested in establishing, first, the extent to which children learning different languages

_____

segments) and then, repeatedly over a period of days, as /nakanendo/, with consonant harmony.

are called on to deal with variegation in the target words that they attempt and, second, the ways in which they deal with the challenge it poses. Comparing such cross-linguistic data should also provide insight into the related question of the likely sources of child failures to match their targets in early word production. That is, if *immaturity of articulatory skill* is the main obstacle to accurate child production, we can expect the proportion of matches to variegated forms to be roughly equivalent across all five language groups, once we control for phonotactic structures. On the other hand, if the *memory load, or representational challenge*, involved in retaining complex word forms well enough to reproduce them also plays a role, then this may be expected to lead to different child responses to variegation in different languages. More specifically, we might expect different effects on representation based on target language differences in rhythmic patterning and phonotactics.

Accordingly, we address the issue of child responses to variegation in adult word targets with a specific focus on cross-linguistic similarities and differences. Based on analyses of both target words and child forms, we ask the following research questions:

(1) To what extent does variegation occur in the target words children attempt in five different languages?

(2) Does the structure of the input language affect children's responses to the challenge of variegation? And if so, what structural characteristics of the adult language might support variegated child word form production in response to target variegation?

(3) Can the cross-linguistic data shed light on the question of the sources of child failures to match their target form? In other words, can we weigh the balance of relative

articulatory as compared with representational advance in achieving accurate production? More specifically, does the data provide evidence as to whether ambient language differences – in stress patterning, the presence of medial geminates, or the diversity of syllable structures, for example – affect children's long-term memory for (or representation of) target forms?

Note on phonological structure

To facilitate consideration of possible ambient language effects on child responses to variegation we briefly characterize relevant aspects of the phonological structure of each language included in our analyses. We focus on consonant inventory size, syllable structure and accent; we assume that these are the features most likely to affect child representation and production. We do not discuss vowel inventories, which do not directly affect our analyses.

English words have stress on the first syllable of 75% of disyllabic (trochaic) words but on the second syllable of 75% of disyllabic (iambic) phrases (Delattre, 1965). English has 24 consonants and permits, even in the monomorphemic forms that underlie most early child words, up to three consonants in both onset and coda positions, creating a highly diverse set of syllable types.

French lacks lexical stress but marks phrase-final syllables with lengthening (Delattre, 1965). The consonant inventory (N = 21) and phonotactic structure are similar to that of English in terms of diversity of syllable types, but word-final consonants are far less frequent in French (66% occurrence in American English content words in child-directed speech [CDS], 25% in French: Vihman, Kay, Boysson-Bardies, Durand & Sundberg, 1994).

Finnish has demarcative stress on the first syllable, but its rhythm is described as relatively even, with secondary stress falling on alternate syllables (Suomi & Ylitalo, 2004). The consonant inventory is small (N = 11) but consonants may occur as intersyllabic geminates anywhere in the word; vowel length is also contrastive and unrestricted as to word position. The language also has front-back vowel harmony. Codas are restricted to coronals; no clusters occur word-initially or -finally.

Japanese is a pitch accent language, with a high or falling pitch on any one syllable of a word or on none (Ota, 2013). Like Finnish, Japanese has contrastive length in both vowels and consonants. However, geminate consonants are less well contrasted phonetically than in Finnish (Aoyama, 2000; Kunnari, Nakai & Vihman, 2001); consonant duration is also less consistently produced in input speech, as both consonants and vowels may be lengthened for expressive purposes (Kunnari et al., 2001). There are 15 consonants; aside from geminate clusters, which form a coda with the preceding syllable, the only coda permitted is a nasal, which agrees in place of articulation with the next syllable onset, if any, but is otherwise uvular, often with vowel nasalization as well.

Mandarin lacks lexical stress or word-level accentual patterning but has four lexical tones, with each syllable bearing one tone (conventionally marked as numbers in phonemic transcription: 1 high level, 2 rising, 3 fall-rise, 4 falling); in addition, a 'neutral tone' occurs in some cases, such as on grammatical particles and, more importantly for our purposes, on the second syllable of reduplicated words in CDS. The rimes of these syllables are reduced, with the tone manifested as mid-falling after Tones 1, 2 and 4 and low-rising after Tone 3 (Cheng, 1973; Lin, 2007). The only codas allowed are alveolar or velar nasals. There are two syllable types, C_0V_{1-3} or C_0VV_0N

(that is, an optional onset consonant followed by up to two nuclear vowels, optionally followed by either a third vowel or a nasal consonant); the language is said to have an inventory of only some 400 unique syllables, if tone is disregarded (Deng & Dang, 2007) – while English has some 9000 (Huff, 2017). Mandarin has 25 consonants. It does not allow syllable-initial or -final clusters, but in disyllabic words first-syllable codas create word-medial clusters with second-syllable onsets.

The languages in our sample differ in numbers of contrasting consonants and syllables, dominant accent patterns, and syllabic and metric structure. As suggested above, if articulatory immaturity is the sole or primary source of child failure to accurately match target words, we may reasonably expect the children in each group to show a similar range of responses to the challenge of variegation in producing words with comparable phonotactic structures. That is, regardless of the language they are learning, children should employ similar strategies to cope with variegation in words with an equivalent structure, such as, for instance, C_1VC_2V, where C_1 and C_2 are consonants with different supraglottal places of articulation. If, instead, we find systematic cross-linguistic differences in the distributions of child responses, then this could be taken to reflect rhythmic and other phonological differences in the children's linguistic exposure and experience, which would in turn suggest a role for access to long-term memory or representation in early word form production.

Methods

Our data derive from children acquiring languages representing five distinct families, two of them Indo-European (English [Germanic] and French [Romance]), one Finno-Ugric (Finnish), one a likely isolate (Japanese) and one Sinitic (Mandarin Chinese). The six samples consist of

previously reported data sets for US and UK English, Finnish, French and Japanese and a new

data set for Mandarin Chinese. The American English data were collected in California and New

Jersey, the UK English data in Bangor, Wales, and York, England, the Finnish data in Oulu,

Finland, the French data in Paris and the Japanese data in the United States, one corpus from

California, the other from Washington D.C. The Mandarin data were recorded in Yorkshire,

England. Although English is the dominant language in the communities where the Japanese and

Mandarin data were collected, all the children recorded mainly hear the heritage language from

both parents and none produced more than four English words, if any, in the data sampled for

this study.

We draw on data from the end of the single-word period, as this is a time when the child's

phonetic skills are still slowly developing, while at the same time the rate of vocabulary learning

shows a steady, rather rapid increase (Ganger & Brent, 2004). Most of the typical phonological

processes can be observed in this period, including those that we have picked out as possible

responses to the challenge of variegation (i.e., of production of whole words, not individual

segments or clusters): reduplication, consonant harmony, inclusion of non-target glottals and

glides and consonant or syllable omission (Grunwell, 1982). Accordingly, this is an optimal

developmental point for assessing the effects, if any, of different linguistic structures on

children's deployment of these processes in their attempts to produce variegated target words.

In order to ensure that the children's lexical level was similar enough cross-linguistically to

permit a meaningful comparison, we selected samples from longitudinal data based on extent of

spontaneous word production, a reliable index of child lexical advance in this period (Vihman,

2019). We identified the first half-hour session in which each child produced 25 or more lexical

items (words or phrases) spontaneously (see Table 1). This corresponds to the end of the single-word period for most children, with the child's first word combinations typically occurring within a month of the '25-word-point' session. (The longitudinal recording sessions for three of the Japanese children – Hiromi, Kenta and Takeru – were over an hour long. We thus chose sessions with 25 words in the first 30 minutes but included all the words produced in the session in our analyses.) Imitations are also included in the analyses but do not contribute to our estimates of vocabulary size. The data we analyse consist of both the target words the children attempt and the child forms themselves.

Table 1. Language groups and child names and ages.

Language group; child pseudonym	Child age in months	Total words produced	Spontaneous words produced
US English[2]			
Alice	16	34	33
Deborah	17	37	25
Emily	16	36	30
Molly	16	45	31
Sean	16	29	27
Timmy	17	34	27
UK English			
Ella	15	34	34
Lewis	17	31	29
Patrick	19	27	27
Rachel	19	61	31
Tobias	19	34	28
Flora	19	46	35
Ivy	18	36	36
Jude	15	42	36
Tania	19	37	35
Finnish			
Atte	20	25	25

2 Note that the American children are drawn from a study of 20 children that covered ages 9 to 16 or 17 months only; only the six children included here attained the relevant lexical level within that period. This may have resulted in a small bias toward more rapidly developing children in this sample.

Eelis	22	33	33
Eliisa	15	31	30
Mira	17	29	29
Sini	15	53	37
French			
Basile	20	46	33
Carole	15	27	27
Charles	18	27	25
Julien	20	34	34
Laurent	17	27	27
Marie	17	28	27
Noël	18	29	29
Vincent	18	42	27
Japanese			
Haruo	19	40	30
Hiromi	21	46	46*
Kazuko	15	37	27
Kenji	19	26	24
Kenta	27	72	67*
Takeru	21	49	46*
Taro	23	32	28
Mandarin			
Didi	17	35 (+ 1 Eng.)	35
Keke	18	34 (+ 2 Eng.)	34
Shi	17	39	39
Xinyu	17	37 (+ 2 Eng.)	37
Yiyi	17	43 (+ 4 Eng.)	35

* Twenty-five spontaneously produced words identified in the first 30 minutes of hour-long recording; total spontaneous words recorded indicated here.

Note: Data sources are, for US English, Vihman & McCune (1994); for UK English, Keren-Portnoy, Vihman, DePaolis, Whitaker & Williams (2010) and DePaolis, Keren-Portnoy & Vihman (2016); for Finnish, Kunnari (2000) and Savinainen-Makkonen (2001); for French, Boysson-Bardies, Vihman, Roug-Hellichius, Landberg, Durand & Arao (1992) and Wauquier & Yamaguchi (2013); for Japanese, Boysson-Bardies et al. (1992) and Ota (2003); and for Mandarin, Lou (2021).

Target words and child forms by prosodic structure

We began by making a preliminary analysis of the target words produced in each recorded session, counting the targets of all child word forms produced as single units, not combinations (i.e., including some forms targeting fixed expressions such as *what's this*). Table 2 indicates the distribution by length in syllables. The table shows that one- and two-syllable words account for over 90% of the words the children attempt in every group except Japanese, where they account for 78%. Table 3 also shows that the occurrence of monosyllabic and disyllabic targets differs cross-linguistically. Monosyllables outweigh disyllables in both the English groups; disyllables dominate in the other languages, but most sharply in Finnish and Japanese. Because monosyllables are variegated only when they include codas or clusters and these occur only rarely in Finnish or Japanese, children learning those languages encounter variegation mainly in longer words – and this mostly means disyllables. Accordingly, we restrict our analyses largely to disyllabic adult words, which are targeted in sufficient numbers in all our language groups to permit a controlled comparison. In addition, we include target monosyllabic words and words of more than two syllables whenever they result in child disyllabic forms; this makes it possible to gain a more complete idea of the children's responses to variegation, as 74% are variegated (109 of 148 monosyllabic or longer-than-disyllabic targets). Disyllabic targets make up 83% of the words included.

Table 2. Target words by length in syllables.

	Monosyllables	*Disyllables*	*Longer words*	*Total words*
Finnish (5)	0.11	0.79	0.10	171
Japanese (7)	0.14	0.64	0.22	290
Mandarin (5)	0.36	0.61	0.04	185
French (8)	0.32	0.61	0.07	261
UK English (9)	0.55	0.40	0.07	355
US English (6)	0.59	0.36	0.06	219
Mean	*0.34*	*0.57*	*0.09*	

| SD | | 0.20 | 0.16 | 0.07 | | | |

Note: The number of children in each group is indicated in parentheses after each language name, in this table only. Ordered by proportion of disyllables; group means based on individual child means for each word length.

For each child learner we then established the *variant forms* of each word that fall into the four basic prosodic structures that we distinguish for the purposes of this analysis: variegation (or words with more than one supraglottal consonant type: VRG), full reduplication (RED), partial reduplication (or consonant harmony: CH), and OTHER, which includes words with no more than a single supraglottal consonant (e.g., VCV or glottal or glide in any position plus at most a single supraglottal consonant). In addition, we include, for adult target words only, monosyllabic (MONO) and longer-word (LONGER) targets for disyllabic child forms (e.g., French *tiens* 'here, take it' /tjɛ̃/, produced as disyllabic [tete]; Japanese *inaijo* 'I'm not here', produced as [naɪjɔ:], or *omeme* 'eyes', produced as [m:emɛʔ]). Any target form that contained two supraglottal consonants was classified as VRG, including reduplicated syllables with a coda (e.g., *night-night*), words with a medial cluster with two or more supraglottals (e.g., *Elma*), variegated monosyllables (e.g., *bus*) and variegated longer targets (e.g., *spaghetti*).

Selecting examples from one child from each language group, Table 3 shows all the types of prosodic structures targeted for that child's disyllabic forms and all the types produced. For each child we show one form from every category represented in that child's sample.

Table 3. Target words and child forms by prosodic structure.[3]

child name (child age)		Targets				

---

[3]We provide transcription in IPA as well as orthography for all but English target words and in IPA only for Japanese and Mandarin targets. Child forms are in broad IPA.

			VRG	RED	CH	OTHER
US English Deborah (1;5)	VRG	*pickle*		bɑbɑ		
	RED	*mama*			mʌmɑ	
	CH	*baby*			be:bi	
	OTHER	*hello*				ʌwo
UK English Flora (1;7)	VRG	*balloon*	bɪlɔn			
	CH	*baby*	mbibiç	bibi		
	OTHER	*hello*				æla:
Finnish Atte (1;8)	VRG	*kaksi* 'two' /kaksi/			tati	
	RED	*pupu* 'bunny' /pupu/		pupu		
	OTHER	*ukko* 'old man' /uk:o/	uk:on			uk:o
French Julien (1;8)	VRG	*ballon* 'ball' /balɔ̃/	belɔ̃			
	RED	*dodo* 'sleep (BT)' /dodo/		tutu		
	CH	*maman* 'mother' /mamɑ̃/			mɛmɑ̃	
	MONO	*tiens* 'here, take it' /tjɛ̃/		tete		
	OTHER	*allo* 'hello' (telephone) /alo/				alo
Japanese Kazuko (1;3)	VRG	/dak:o/ 'hug'	tɔk:o:		gʌk:ɔ:	
	RED	/mimi/ 'ear'		mɪmɪ		
	OTHER	/aɕi/ 'foot'				ʔɑdʒi
Mandarin Keke (1;6)	VRG	/xɤ2tsi0/ 'box'		kʰɤ3kɤ1		
	RED	/ɕiɛ4ɕiɛ0/ 'thanks'	kʰɤ1tɕʰiɛ4	ɕiɛ1ɕiɛ4		

	CH	/ɕiuo1ɕi0/ 'rest'			ɕio1ɕi4	
	OTHER	/jy4tou2/ 'taro'				ju2tou2

Note. Mandarin tones are indicated using a number for each syllable in both target and child forms: 1 high level, 2 rising, 3 falling-rising, 4 falling. RED Reduplication, CH Consonant harmony, VRG or V Variegation, Mono Monosyllabic target. Variegated targets and child forms are shaded.

Note that for Finnish and Japanese, which have medial geminates, we disregard differences in either vowel or consonant duration in categorizing word forms as reduplicated, so that target words like Finnish *pappa* /pap:a/ 'grandpa' or Japanese *nen:e* 'sleep' (both CVCCV) are treated as reduplicated despite the fact that, strictly speaking, the first syllables /pap/ and /nen/ are not repeated. Similarly, for Mandarin we disregard tone differences in categorising word forms, as they do not contribute to the variegation challenge per se. Furthermore, accuracy in disyllabic tone pattern production has been found to be independent of accuracy in segmental pattern production at this developmental point (Choo, 2022). Recall that the second syllable of Mandarin reduplicated words tends to have neutral tone in CDS (71% of child targets here, based on the means for each child), but some repeat the tone (20%) and a few change tone (9%). Thus 'reduplication' refers to the segmental sequence only.

In our analysis of child forms we count no more than one variant per word in any one structure (following Vihman, 2019), but include as many structures per word as the child forms warrant (see, in Table 3, Flora's forms for English *baby,* Kazuko's forms for Japanese /dak:o/ 'hug' or Keke's forms for Mandarin /ɕiɛ4ɕiɛ0/ 'thanks'). Note that differences in voice onset time (voicing or aspiration: VOT), which are not reliably produced in this developmental period (Macken, 1980), are not taken to constitute variegation (e.g., Deborah's form [k̬ɪ̝gʷɛ̝] for

English *spaghetti* is categorized as CH with velar harmony despite the voicing change); VOT

contrasts are also disregarded in categorising target words. Also, clusters, which may occur in

any position in English or French but only medially in Finnish, Japanese or Mandarin,

complicate the picture to some extent. All word types with variegation – whether across vocalic

intervals or in clusters – are combined in the variegation tallies for both targets and child forms,

so that, for example, we treat as variegated both Finnish *kaksi* 'two', with its two distinct

supraglottal onset consonants /k/ and /s/, and *ankka* 'duck', the sole medial consonantal slot of

which is filled by a cluster of two supraglottal consonants. Finally, note that although we treat

consonants with full or partial supraglottal closure differently from glides or glottals for the

purpose of identifying variegated forms, reduplicated and harmony forms may consist of glides

or glottals only (e.g., for targets, reduplicated French *ouioui* /wiwi/ 'yes, yes', Finnish *hauhau*

'woofwoof'; for child forms, harmonized US English *hello* [hɔʔhi:::], Japanese /dʑadʑa/ 'bath,

pouring water' [ȷe̞ȷaʰə]). These are then treated as RED or CH, not as OTHER.

Results: Variegation in targets and child word forms

We begin with an overview of the relative frequency of occurrence of variegation in the target

words children produced as disyllables in each language group. Table 4 presents the median

counts of each prosodic type targeted by the children. In Figure 1, these are converted to

proportions with respect to the total disyllables produced by each child and presented as means

for language groups in order to allow comparisons within and between languages. Here we see,

in direct response to RQ1, that variegation is by far the most common prosodic structure

attempted overall, accounting for 52% of the target words on average. Nevertheless, we can also

see that the language groups differ, with Japanese (at 42%) and Mandarin (at 29%) having the

smallest proportion of variegated targets. Mandarin is the only language in which reduplication

(at 52%) exceeds variegation in target words (see Fig. 1).

Table 4. Median counts (and range of counts) of prosodic structure types in the targets of child disyllabic forms.

	VRG	*RED*	*CH*	*OTHER*	*MONO or LONGER*	*Total disyllables*
US English (N = 6)	12.0 (9 - 14)	1.0 (1 - 3)	2.0 (0 - 3)	0.0 (0 - 2)	1.0 (0 - 3)	17.0 (13 -21)
UK English (N = 9)	12.0 (4 - 16)	2.0 (0 - 6)	2.0 (0 - 7)	2.0 (0 - 6)	3.0 (1 - 6)	21.0 (12 - 33)
French (N = 8)	13.5 (9 - 19)	5.0 (0 - 8)	2.0 (0 - 4)	1.0 (0 - 9)	0.0 (0 - 2)	24.5 (17 - 33)
Finnish (N = 5)	17.0 (10 - 39)	4.0 (2 - 5)	4.0 (0 - 7)	4.0 (1 - 8)	0.0 (0 - 1)	27.0 (23 - 54)
Japanese (N = 7)	12.0 (5 - 29)	10.0 (3 - 12)	2.0 (0 - 4)	5.0 (2 - 8)	1.0 (0 - 3)	25.0 (20 - 52)
Mandarin (N = 5)	7.0 (4 - 11)	13.0 (6 - 18)	0.0 (0 - 2)	5.0 (1 - 6)	0.0 (0 - 1)	25.0 (15 - 32)

Note. VRG variegation (shaded), RED reduplicated, CH consonant harmony. OTHER no more than one supraglottal consonant. MONO or LONGER Target with one syllable or more than two syllables that are produced by the child as disyllable.

Figure 1.

Mean proportions of prosodic structures in target words
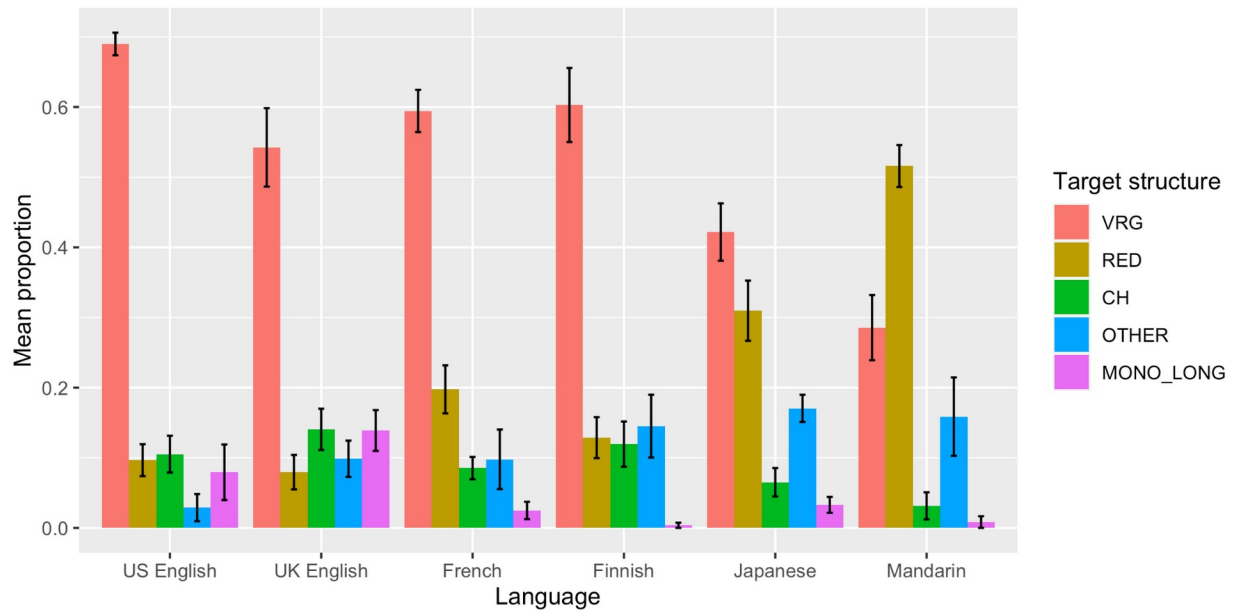
Note. Error bars show +/-1 standard error of mean. VRG variegation, RED reduplicated, CH consonant harmony. OTHER no more than one supraglottal consonant. MONO or LONGER Target with one syllable or more than two syllables that are produced by the child as disyllable.

The distribution of target word counts by language and prosodic structure was submitted to a mixed-effects Poisson regression analysis. As fixed effects, Language and Structure were deviation-coded against the grand mean for each category, with French and OTHER chosen as the levels coded as -1. The procedure was run using the lmer4 package on R. The initial model also contained by-participant random intercepts and slopes for Language and Structure (Count ~ Structure * Language + (1 + Structure * Language | Child), but it failed to converge. Model reduction was performed by reducing the random factor structure, first by removing the interaction between the slopes and then the individual slopes, first for Language, then for Structure, until convergence without singularity was obtained. The final model included an interaction term for Language and Structure and random intercepts for individual children (Count

The full results are presented in Appendix A.[4] Against the

expected frequencies based on the overall distributions along Language and Structure, the

observed frequencies for VRG targets were significantly higher in US English and Finnish while

they were lower in UK English and Japanese. Observed frequencies of RED targets were

significantly higher than expected in Japanese and Mandarin but lower in the two varieties of

English. Frequencies of CH targets were higher in Finnish but lower in Japanese. Finally, both

US and UK English had more MONO/LONGER targets than expected.

Next, as a first step toward responding to RQ2, we consider the distribution of the disyllabic

child word forms produced for variegated targets in each language group. Table 5 presents the

median counts of each prosodic type, and Figure 2 shows the distributions as proportions with

respect to the total disyllables produced by each child for variegated targets. The results reveal

that children learning Mandarin or Japanese produce a higher proportion of disyllabic variegated

forms for variegated targets than do the children learning any of the European languages. This is

striking, given the relatively small proportion of variegated targets they attempt (Figure 1).

Table 5. Median counts (and range of counts) of prosodic structure types produced for variegated targets.

Language group	VRG	RED	CH	OTHER	Total disyllables
US English	3.5 (1 - 6)	1.0 (0 - 4)	2.5 (0 - 4)	4.0 (2 - 9)	12.0 (9 -14)
UK English	3.0 (0 - 11)	0.0 (0 - 2)	2.0 (0 - 6)	4.0 (3 - 8)	12.0 (4 - 16)
French	3.5 (1 - 8)	1.5 (0 - 4)	2.5 (1 - 4)	6.0 (0 - 10)	13.5 (9 - 19)
Finnish	1.0 (0 - 4)	2.0 (0 - 7)	6.0 (0 - 11)	9.0 (1 - 24)	17.0 (10 - 39)
Japanese	6.0 (2 - 9)	1.0 (0 - 4)	2.0 (0 - 10)	3.0 (0 - 7)	12.0 (5 - 29)
Mandarin	3.0 (3 - 7)	2.0 (1 - 3)	0.0 (0 - 1)	0.0 (0 - 2)	7.0 (4 - 11)

---

4 In this and all subsequent reports of the Poisson regression analysis the main fixed effects for Language indicate whether there was a difference in the overall tokens produced for each language. These effects can reflect differences in the baseline sample size between languages; they are not directly relevant to the current analysis.

Note. VRG variegation (shaded), RED reduplication, CH consonant harmony, OTHER no more than one supraglottal consonant. Means for each group based on individual means in each structure.

Figure 2.

Mean proportions of prosodic structures in child productions for variegated targets
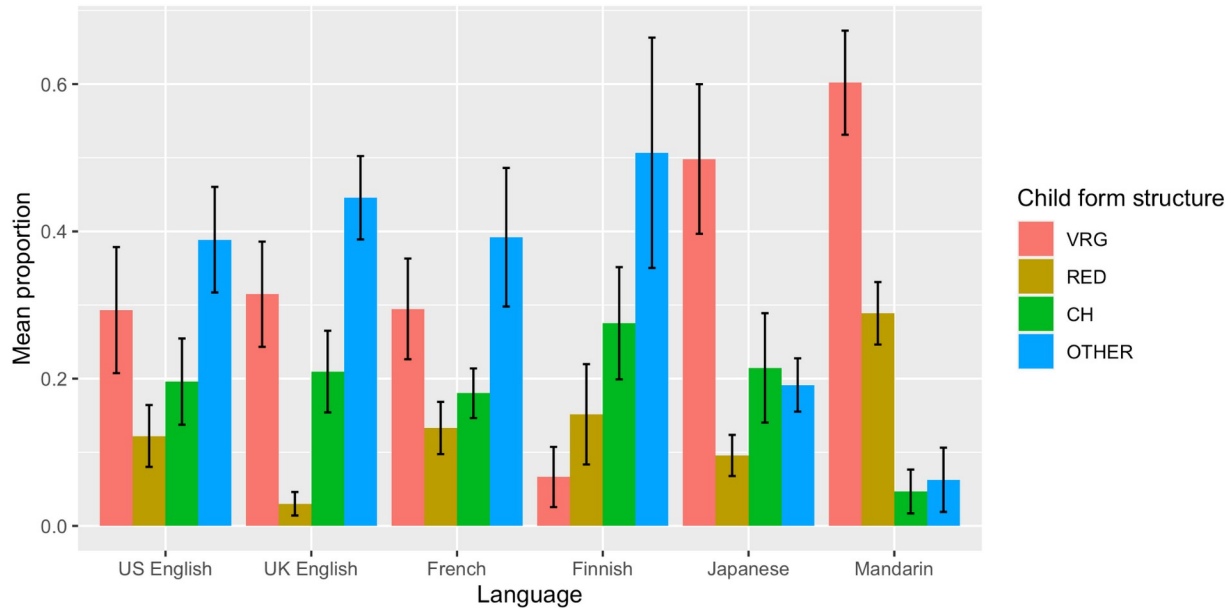
Note. Error bars show +/-1 standard error of mean.

As with the target words, we conducted a mixed-effects Poisson regression to compare the distributions of the child forms belonging to each prosodic category. Details of the category coding and model selection ==procedure were the same as those used for the target analysis. The final model consisted of both Language and Structure as fixed effects and by-participant random intercepts.== The full results are presented in Appendix B. Against the expected frequencies based on the overall distributions along Language and Structure, the observed frequencies for VRG child forms were significantly higher in Japanese and Mandarin while they were lower in Finnish. Observed frequencies of RED child forms were significantly higher than expected in Mandarin but lower in UK English. Frequencies of CH child forms were higher in Finnish but

lower in Mandarin. On the whole, these outcomes are not inconsistent with what we saw in the target analysis, but one crucial difference is that Finnish has a relatively low proportion of child VRG forms despite having a relatively high proportion of VRG targets and, conversely, Japanese and Mandarin have a relatively higher proportion of child VRG forms despite having relatively low proportions of VRG targets (although the latter was not statistically significant for Mandarin).

Child responses to variegated targets

Having established the extent and distribution of the children's responses to variegated targets in each language group, we now examine these results more closely. Our target word analysis showed that both Japanese and Mandarin words provide a high proportion of reduplication and a small proportion of harmony alongside a comparatively low proportion of variegation. Our child data show that the relative use of reduplication and consonant harmony tends to be complementary: All groups make use of both, but harmony is rare in Mandarin while it is about equally common in the other groups; the balance between the two processes is roughly related to occurrence in the targets, with the high proportion of reduplication in Mandarin child forms, in particular, seeming to reflect its occurrence in targets. Combining the proportions for reduplication and harmony, we can see that, for all groups, some form of repetition is a resource for responding to variegation in target words. The third option, OTHER, is less clearly related to the proportions seen in the targets. In the next section we look more closely into the cases where children have recourse to that option in producing variegated forms.

Child uses of OTHER

In the European language groups OTHER accounts for a higher proportion of child forms for variegation (Fig. 2) than of targets (Fig. 1), but that is not the case in Japanese or Mandarin. We carried out an analysis to determine whether children in the different language groups also tend to differ as to their preferred single-consonant output forms. Table 6 illustrates children's adaptations of variegated target words to OTHER forms, with at least one example for each subcategory, where possible; empty cells indicate that no instances occurred in our data.

Table 6. Subcategories of OTHER child forms.

	[GLIDE]VCV	*CV[GLIDE]V*	*VCV*	*[GLOT]VCV*	*CV.V*	*CV[GLOT]V*
US English	*water* [wɑ:mɪ]	*shiny* [taji], *gran'ma* [mɛʊwʌ]	*bus* [əb:aɪ]	*balloon* [ʔæbæ]	*bunny* [baɪ́:ɪ]	*car* [kɑ:ʔɑ]
UK English	*zebra* [wɛbəʰ]	*snake* [ʃeɪ:jɛʰ]	*open* [ɑʔpə:], *caterpillar* [abaʔ]			*bath* [baʰʔɪ]
Finnish		*piiloon* '[go] into hiding' [i[j]o:]	*kissa* 'cat' [iç:a], *loppu* 'finish, all done' [op:u]		*korva* 'ear' [ko:a]	
French	*Dalila* [jeila].		*bravo* [avo], *debout* 'standing' [əbu]	*c'est bon* 'it's good' [habõ]		*canard* 'duck' [aʔa]
Japanese	*dʑɯ:sɯ* 'juice' [jʊt:'ʊ]	*denwa* 'telephone' [tʃɪ::wɑ]	*zo:sãn* 'elephant' [odoɳ]	*dak:o* 'hold me' [ʔak:a]	*gohãn* 'meal' [dɔã]	*tɕi:zɯ* 'cheese' [ɟiʔɳ]
Mandarin	*fei1tɕi1* 'airplane' [wɤ2ji1]	*xau3ma0* 'okay?' [xa3wa4]				

As some subcategories occurred only rarely, for the purpose of quantification we combined (i) forms with either onset or medial glide, (ii) forms either lacking an onset or having a glottal

onset and (iii) forms lacking a medial consonant or having a glottal medial consonant. Table 7

shows the distribution of OTHER forms by subcategory for each language.

Table 7. OTHER child forms for variegated targets.

	CV[GLIDE]V, [GLIDE]VCV	VCV, [GLOT]VCV	CV.V, CV[GLOT]V	Total	N/child
US English	**0.44**	**0.44**	0.12	25	4.17
UK English	**0.51**	0.37	0.11	35	3.89
Finnish	0.04	**0.76**	0.20	49	9.80
French	0.09	**0.89**	0.02	44	5.50
Japanese	**0.39**	**0.33**	0.28	18	2.57
Mandarin	**1.00**	0.00	0.00	6	1.20
Mean	*0.41*	*0.46*	*0.12*		
SD	*0.35*	*0.32*	*0.11*		

Note. The most-used structures in each group are in **bold face**. Means for each group based on individual means in each structure.

Forms lacking a word-initial supraglottal consonant account for a particularly high proportion of

the Finnish and French forms. Heavy use of onset consonant omission, affecting even early-

learned consonants, has been reported elsewhere for both Finnish (e.g., *nalle* 'teddy' [al:e], *pallo*

'ball' [al:o]) and French (e.g., *debout* 'standing' [əbu], *garçon* 'boy' [aʑa]) and ascribed to the

perceptual effect of medial geminates in Finnish (Savinainen-Makkonen, 2007) and of phrase-

final accent in French (Vihman & Croft, 2007). Both these effects of perceptual salience

(lengthened medial consonant closure, final syllable lengthening) have been shown

experimentally to detract from attention to (and thus representation of) the word-initial consonant

(cf. Vihman & Majorano, 2017, on geminates in Italian; Hallé & Boysson-Bardies, 1996,

Vihman, Nakai, DePaolis & Hallé, 2004, on final syllable accent in French, and Segal et al.,

2020, on iambic accent in Hebrew).

In English, on the other hand, strong first-syllable stress means that initial consonant omission is rare (Vihman et al., 2004); the children's VCV forms seldom derive from disyllables with a supraglottal initial consonant but include vowel-initial target words, monosyllables and longer words with truncated first syllables (Table 6). Instead of initial consonant omission, glide substitution is the most common pattern; this relates in part to the common process of gliding liquids (e.g, *balloon* [bɛː[j]ʊ] or, from another child, [ɪwuːun]), but glides in the child forms do not always relate to a target liquid (Table 6). Japanese learners make roughly equivalent use of each of the subcategories. Finally, the very few OTHER child forms for variegated targets in Mandarin all involve glide substitution.

The structure of the input language clearly affects children's responses to variegation in every category, leading to differences in relative use of full and partial consonant harmony and in reduction of the variegated form to a single-consonant output (OTHER), despite the fact that the challenge and the 'in-principle solutions' involved must be similar cross-linguistically. Our analysis also revealed that the challenge involved in variegation is more successfully handled by some language groups than others, as evidenced by the relatively high proportion of Mandarin- and Japanese-learning children's variegated forms for variegated targets. We turn our attention now to the second part of RQ2, the question as to how structural differences between the languages might lead to child differences in the ability to reproduce variegated targets in general.

The effect of structural predictability on child responses to variegation

One source of cross-linguistic differences in child ability to reproduce variegated targets might be the relative complexity of the shapes of those target words. To test that possibility, we reanalysed the European-language data, leaving out all words that include word-final consonants

other than nasals or supraglottal consonant clusters other than medial nasal + consonant, which Japanese and Mandarin structure allow. In other words, we reduced the variegated target words that the children attempted in the European languages to just those that conform to the more constrained structures of Japanese and Mandarin, to test whether the children were more able to produce variegated forms for those simpler variegated targets: That would suggest that the greater success of the children learning Japanese and Mandarin at using variegated forms for variegated targets could be ascribed to the fact that their target words are simpler in shape and therefore easier to match. Table 8 presents the median counts of each prosodic type targeted by the children. In Figure 3, these are converted to proportions with respect to the total disyllables produced by each child and presented as means for language groups to allow comparisons within and between languages. This analysis does not affect Japanese and Mandarin, so the data from those languages are simply repeated from Table 4 and Figure 1.

Table 8. Median counts (and range of counts) of prosodic structure types in the targets of child disyllabic forms, excluding word-final non-nasals and all but medial nasal clusters.

	VRG	*RED*	*CH*	*OTHER*	*MONO or LONGER*	*Total disyllables*
US English (N = 6)	3.5 (0 - 5)	1.0 (1 - 2)	2.0 (0 - 3)	0.0 (0 - 2)	1.0 (0 - 3)	8.5 (4 -10)
UK English (N = 9)	4.0 (1 - 8)	0.0 (0 - 6)	2.0 (0 - 7)	2.0 (0 - 6)	2.0 (0 - 5)	12.0 (5 - 22)
French (N = 8)	6.0 (2 - 8)	5.5 (0 - 8)	2.0 (0 - 4)	1.0 (0 - 9)	0.0 (0 - 2)	15.0 (11 - 20)
Finnish (N = 5)	15.0 (7 - 30)	4.0 (2 - 5)	4.0 (0 - 7)	4.0 (1 - 8)	0.0 (0 - 1)	25.0 (20 - 45)
Japanese (N = 7)	12.0 (5 - 29)	10.0 (3 - 12)	2.0 (0 - 4)	5.0 (2 - 8)	1.0 (0 - 3)	25.0 (20 - 52)
Mandarin (N = 5)	7.0 (4 - 11)	13.0 (6 - 18)	0.0 (0 - 2)	5.0 (1 - 6)	0.0 (0 - 1)	25.0 (15 - 32)

Note. VRG variegation (shaded), RED reduplicated, CH consonant harmony. OTHER no more than one supraglottal consonant. MONO or LONGER Target with one syllable or more than two syllables that are produced by the child as disyllable. Data for Japanese and Mandarin are the same as those from Table 4.

29

Figure 3.

Mean proportions of prosodic structures for targets, excluding word-final non-nasals and all but medial nasal clusters
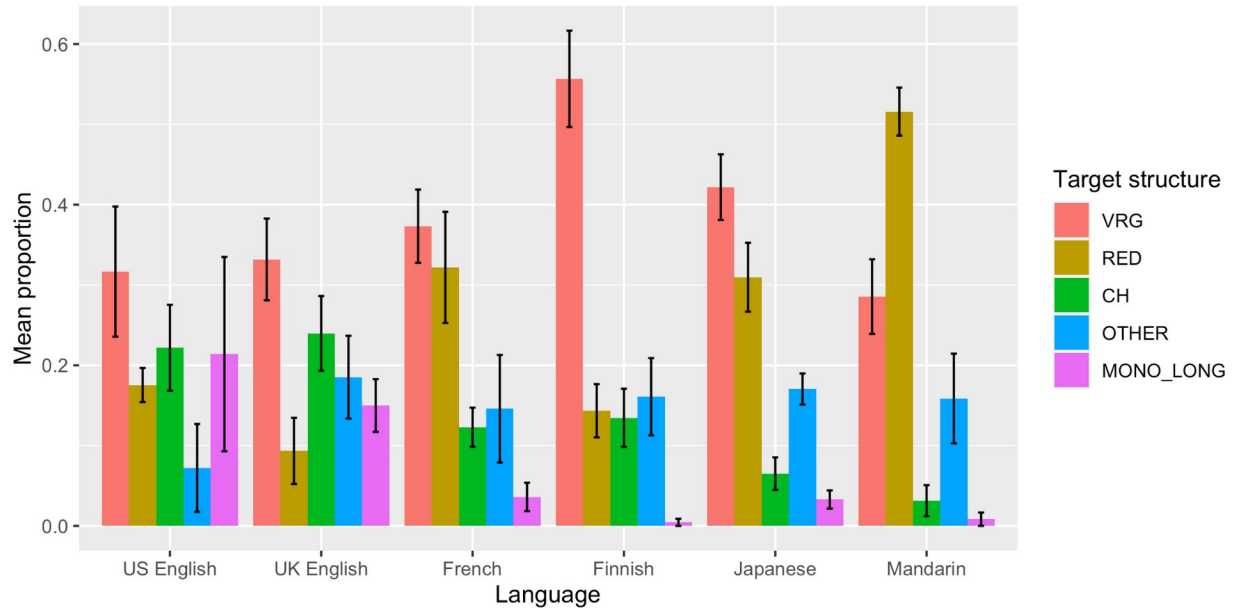
Note. Error bars show +/-1 standard error of mean.

We ran a mixed-effects Poisson regression to compare the distributions of the child forms belonging to each prosodic category, following the same model selection procedure used for the analysis of the data in Table 4 (see Appendix C for full results). The final model consisted of fixed effects for Language and Structure and by-participant random intercepts. Under this analysis, VRG targets were still significantly more frequent than expected in Finnish but lower in UK English. RED targets were significantly more frequent than expected in Japanese and Mandarin, but lower in UK English. Conversely, CH targets were more frequent in UK English but lower in Japanese and Mandarin. MONO and LONGER targets were more frequent in the two English dialects. This analysis of the stripped-down target words shows that the different European languages are affected differently in terms of VRG targets: Finnish, with its restriction

of codas to coronals and of clusters to word-medial position only, remains little changed, while the number of variegated target words in English and French is reduced.

We now turn to the distribution of the disyllabic child word forms produced for variegated targets in each language group, but without those with word-final non-nasals and without all but medial nasal clusters. Table 9 presents the median counts of each prosodic type using this stripped-down data set and Figure 4 shows the distributions as proportions with respect to the total disyllables produced by each child. The results of the statistical analysis using Poisson regression show few differences from those applied to the entire data (full results are presented in Appendix D; the model selection procedure and the final model structure were the same as in the other analyses above). Most importantly, VRG forms were significantly more frequently produced by Japanese and Mandarin children and less by Finnish children (see Appendix D).

Table 9. Median counts (and range of counts) of child disyllabic forms produced for variegated targets, excluding targets with word-final non-nasals and all but medial nasal clusters.

Language group	VRG	RED	CH	OTHER	Total disyllables
US English	1.0 (0 - 2)	1.0 (0 - 1)	0.0 (0 - 3)	1.0 (1 - 3)	4.0 (1 -5)
UK English	1.0 (0 - 6)	0.0 (0 - 1)	1.0 (0 - 2)	1.0 (0 - 4)	3.0 (1 - 8)
French	1.5 (0 - 5)	0.5 (0 - 1)	1.0 (0 - 2)	2.0 (0 - 6)	6.0 (2 - 8)
Finnish	1.0 (0 - 3)	2.0 (0 - 6)	5.0 (0 - 8)	5.0 (1 - 19)	13.0 (5 - 30)
Japanese	6.0 (2 - 9)	1.0 (0 - 4)	2.0 (0 - 10)	3.0 (0 - 7)	12.0 (5 - 29)
Mandarin	3.0 (3 - 7)	2.0 (1 - 3)	0.0 (0 - 1)	0.0 (0 - 2)	7.0 (4 - 11)

Note. VRG variegation (shaded), RED reduplication, CH consonant harmony, OTHER no more than one supraglottal consonant. Data for Japanese and Mandarin are the same as those from Table 5.

Figure 4.

Mean proportions of prosodic structures in child productions for variegated targets, excluding targets with word-final non-nasals and all but medial nasal clusters.
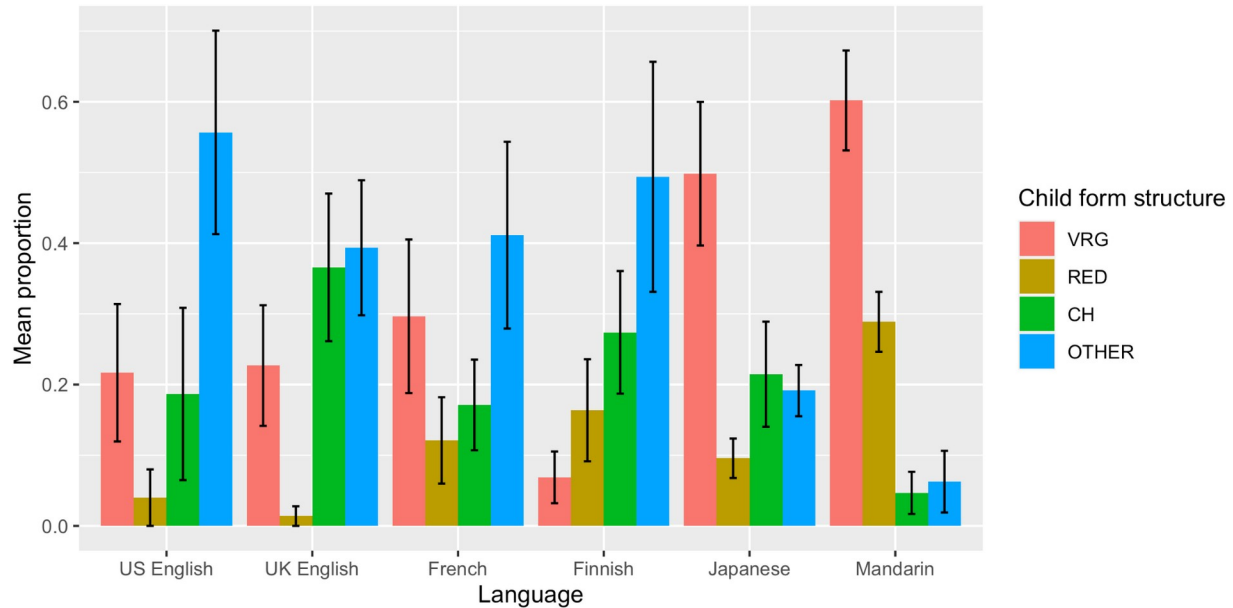
Note. Error bars show +/-1 standard error of mean.

In other words, the reanalysis, which filters the European target words through the syllable-structure constraints of Japanese and Mandarin, has little effect on the proportion of child variegated forms produced in response to those targets. Finnish targets are very little changed by this reduction and child production of variegation for variegated targets remains very low in Finnish in comparison to English and French, which changed more substantially where targets are concerned (Table 8); English and French child variegated forms for variegated targets remain at approximately the same level as in the full analysis (Table 5). Overall, the proportion of variegated forms that we find for Japanese and Mandarin child forms now appears even more extreme, given the lower mean for the European languages.

From this we conclude that the greater proportion of variegated forms that Japanese- and Mandarin-learning children produce for variegated targets cannot be accounted for – or not entirely – by the simplicity, or the restricted set of structures, represented by the particular adult word forms that they are trying to produce: Reducing the targets of children learning European languages to comparably simpler structures does not result in a greater proportion of variegation in the children's word forms. In short, although the overall simplicity of syllable structure in these East Asian languages appears to affect child production of variegated targets, it cannot tell the whole story; the children's wider experience of the language must be playing a role. The challenge for children learning languages such as Finnish, English and French, then, must not be so much the specific difficulty of individual variegated words, such that a selection of simpler variegated forms would pose a lesser challenge. The difficulty seems rather to be in the overall complexity of the phonological structure of those languages. That is, variegation appears to present a problem for the children not only in terms of handling the occurrence of more than one supraglottal consonant per word, but also in terms of retaining the many different structures and positional combinations of such consonants that those languages allow.

We therefore propose that the crucial difference between the East Asian languages and the European languages analysed here may be the number of different syllables required to map out the words of the language. The smaller the syllable inventory, the easier it should be for a learner to retain, represent and reproduce words by combining the possible syllables, allowing them to produce variegated outputs more successfully. To test this idea, we provide two analyses of syllable inventory size in the languages of interest.

The first column in Table 10 shows the syllable inventory size of (UK) English, Finnish, French, Japanese and Mandarin.[5] Note that durational differences in both vowels and consonants are respected but tonal differences were disregarded for Mandarin syllables such that syllables with different tones were considered the same if they had the same segmental structure.[6] To check that this observation extends to a typical child lexicon, we applied the same principle to our data by merging the disyllabic target words of the various children for each language group and calculating the number of unique syllables that make them up. Here again we disregarded tone in classifying Mandarin syllables. The relevant figures, shown in the third column of Table 10, have been scaled to per-10-word values to allow direct cross-linguistic comparison. The results show that children learning Mandarin or Japanese need only about 10 syllables to represent/produce 10 disyllabic words, whereas children learning English, French and also Finnish – despite its relatively simple syllable structure – need more.

Table 10. Syllable inventory size in adult and child language.

	# Unique syllables in adult corpora	# Unique syllables per 10 disyllabic words (child corpora)
UK English	6,949	15.73
US English	n/a	14.72
Finnish	3,844	12.73
French	2,949	12.23
Japanese	643	10.74
Mandarin	416	9.88

Note. All data for adult languages, except Mandarin, are the number of unique syllables found in the most frequent 20,000 words in adult corpora (Oh, 2015). The count for adult Mandarin is the number of segmentally-defined syllable types in the Xinhua dictionary (Xia, 2000). Estimates for child data are based on the number of syllable types per 10 disyllabic words found in the corpora.

_____

5 US English is not included, as comparable data were not available in Oh (2015). However, there is little reason to think that the number would differ greatly from UK English.
6 Even when tonally contrastive syllables are treated as different syllables, the number of syllables in Mandarin is lower than those of the European languages (1,274, according to Oh's (2015) analysis)

A remaining puzzle is the far lower rate of child variegated production in Finnish as compared to English and French. The Finnish syllable inventory is larger than that of Japanese or Mandarin, falling between those of English and French. Finnish phonological structure is less restrictive than that of the two Asian languages – it allows all five of its coronal consonants (/t, s, n, l, r/) to occur as codas, both word-medially and finally – but more restrictive than English or French. As in Japanese, the nasal in Finnish nasal + stop clusters assimilates in place to the adjacent stop; the velar nasal can also occur as a geminate, under morphophonological alternation with /ŋk/. Given the size of the syllable inventory involved, it is not surprising that the Finnish children use fewer variegated productions for variegated targets than the Japanese and Mandarin learners. But what can explain the low rate of variegation in child Finnish as compared with English and French?

The key observation here is that Finnish children respond more often to variegation with the category OTHER than children learning any of the other languages, primarily due to omission of the word-initial consonant – and this is the case even though Finnish consistently stresses the first syllable of content words. Finnish children clearly pay reduced attention to the word-initial consonant (Savinainen-Makkonen, 2000); we assume that this is related to the presence of medial geminates (Vihman & Majorano, 2017). In fact, of the 25 target forms with geminates, all but one (*pallo* 'ball' > [a.o]) are produced as VCV forms; complementarily, of the 25 Finnish child forms produced with initial consonant omission for consonant-initial variegated targets, just four of the target words have medial singletons.

In French the accentual lengthening of the second syllable of disyllabic words similarly draws attention away from the onset consonant (Hallé & Boysson-Bardies, 1996; Vihman et al., 2004);

the considerable difference in child variegation in French as compared to Finnish appears to be due to the higher use of reduplication and harmony in Finnish.

Japanese also has geminates, but as noted earlier, these are less strongly contrasted than is the case in Finnish and less reliably produced in input speech; Japanese children's production of duration also takes longer to become adult-like (Aoyama, 2000). Furthermore, only half as many of the words that the children target have geminates in Japanese (24%, averaged across the 7 children) as in Finnish (51%, for the 5 children). Both the weaker acoustic difference and the lower input frequency help to account for the lesser impact of geminates on child production in Japanese. This leads us to conclude that the basic difference between Finnish and Japanese is likely representational: Finnish children appear to best retain the medial consonant while Japanese children retain something of the variegated sequence.

General Discussion

We set out to establish the extent to which children learning languages of differing phonological structure must face the challenge of producing variegated adult words and to explore how differently they respond to that challenge. Our findings are necessarily limited by the size of our language groups and the relatively small number of words available for analysis for each child. Nevertheless, we found clear answers to our first two research questions. First, we found that variegated adult targets made up over half the words the children attempted to say overall, although there were differences by language. Secondly, we found clear differences by language group in child responses to target-word variegation. The Mandarin and Japanese child forms for variegated targets were considerably more often variegated than those of the other language groups.

36

We found evidence that these differences are related to structural differences in the adult languages. There were sharp differences by language in the children's uses of the processes that reduce the numbers of different supraglottal consonants in word forms. The children learning Mandarin tended to produce full reduplication, reflecting the high incidence of reduplication in the input, while the other groups made greater use of partial reduplication, or consonant harmony; when these are taken together, however, we see that they amount to much the same process, which reflects children's preference for forms with consonant repetition rather than variegation. In addition, initial consonant omission was strongly represented in French and Finnish, a likely reflection of the adult-language rhythmic differences mentioned earlier; (glottal)VCV forms were less often produced in English and Japanese and not at all in Mandarin.

We return now to the more fundamental question of the relative importance, for the shaping of children's word forms, of the maturity of their articulatory skills in comparison with the adequacy of their long-term representations of the words they attempt to produce. Note that vocal practice, first in prelinguistic babbling and then in early word production, contributes to both aspects: Vocal practice necessarily improves articulatory skills, but it also lends salience (and thus memorability) to aspects of input speech that are like what the child is producing, or that are, in other words, familiar from their own often-repeated output (Majorano, Vihman & DePaolis, 2014; Vihman, DePaolis & Keren-Portnoy, 2014; Waterson, 1971). Beyond that, vocal practice lays down the foundations of phonological memory, or the ability to retain novel word forms (Keren-Portnoy et al., 2010; Vihman, 2022); this helps to account for the fact that lexical advance itself supports further word learning (Fernald, Swingley & Pinto, 2001; Torkildsen, Hansen, Svangstu, Smith, Simonsen, Moen & Lindgren, 2009). In short, both

articulatory and representational factors undoubtedly play a role in early word learning; the debate over what supports accurate production likely cannot be resolved in favour of one or the other alone.

We argued that if variegation was difficult for articulatory reasons alone, then we should see essentially the same degree of difficulty – especially in relation to variegated targets – in children learning any language. And indeed we do see that all of the children respond to variegated targets with less variegation than is found in the targets and produce more of the simpler prosodic structures (reduplication, consonant harmony or forms involving consonant omission or reduction to glottals or glides). However, we found that children learning different languages differ in the extent to which they resort to solutions that do not require variegation. Interestingly, these differences do not accord with the degree to which the languages present children with the need to produce variegation. Although we might expect that experiencing proportionately more variegation in the words they attempt provides children with more opportunities for vocal practice with variegation, our results go in the opposite direction: Children learning either Mandarin or Japanese, whose targets are less often variegated, still succeed more often in producing variegated forms for those targets than do children in the other groups. Thus, our results suggest that it is highly unlikely that immature articulatory abilities are the single most important source of difficulty with producing words in an adult-like manner around the end of the single-word period.

Although we have found that the structure of the input language does affect child responses to variegation, we have no conclusive answer as to just how those structural effects translate into production differences. Target language shaping is partially rooted in *perceptual salience*:

Recall our earlier comments on the effects of both geminates and accentual patterning on infant word-form recognition. However, as noted above, the ambient language structure also affects infants' *production experience*, based on the options the language provides, which leads to ongoing accommodation to the structural requirements of the language. That is, sensorimotor experience (production practice) helps to develop children's articulatory skills while at the same time tuning up their sensitivity to the phonological patterning of their language.

The range of syllable choices in a language and differences in prosodic structure both seem to contribute to the differences in child responses to variegation (see Post & Payne, 2018). Languages like English, French and Finnish, with their greater range of syllables, are more challenging for some aspects of early word learning than Japanese or Mandarin. Both hearing and producing forms that are phonotactically simple and relatively predictable (disregarding effects of lexical tone, pitch change or segmental duration) makes the remembering or planning of the production of such forms an easier task than can be the case when each word may involve any one of a considerable range of syllable shapes and of possible consonants for each syllabic slot. That multiplicity of options must increase the challenge of remembering and/or planning word production.

Furthermore, it has been demonstrated that *overall familiarity* (from production practice) with subcomponents (syllables or segments) of a novel form leads to greater accuracy in production (Cychosz, Erskine, Munson & Edwards, 2021; Dollaghan et al., 1995; Keren-Portnoy et al., 2010). This suggests that having a less diverse inventory of possible syllable types to learn and produce might well support more ambitious and more accurate word learning, as the child's repertoire of distinct syllabic motor routines could more readily be recruited to first retain and

then reproduce novel patterns. Here Japanese and Mandarin would present some advantages – although learning, for each lexical item, contrastive vowel and consonant duration and pitch accent (in Japanese) or tone patterns (in Mandarin) adds a layer of difficulty that we have disregarded here. In general, the lesson we draw from our comparison of 'stripped down' English, Finnish and French with Japanese and Mandarin is that it is not just the structure of individual word targets but the entire system or set of possible structures a child has experienced that shapes child responses to variegation.

References

Aitchison, J. & Chiat, S. (1981). Natural phonology or natural memory? The interaction

between phonological processes and recall mechanisms. *Language and Speech, 24,* 311-326.

Aoyama, K. (2000). A psycholinguistic perspective on Finnish and Japanese prosody:

perception, production and child acquisition of consonantal quantity distinctions. Doctoral

dissertation, University of Hawaii.

Boysson-Bardies, B. de, Vihman, M. M., Roug-Hellichius, L., Durand, C., Landberg, I. & Arao

F. (1992). Material evidence of infant selection from the target language: A cross-linguistic

phonetic study. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological*

development: models, research, implications. Parkton, MD: York Press.

Cheng, C.C**.** (1973). *A synchronic phonology of Mandarin Chinese*. (v. 4). Amsterdam: Walter

de Gruyter.

Choo, R. Q. (2022). The acquisition of segments and tones in children learning Mandarin: An

observational and experimental study. Unpublished PhD thesis. University of York.

Cychosz, M., Erskine, M., Munson, B. & Edwards, J. (2021). A lexical advantage in four-year-

old children's word repetition. *Journal of Child Language, 48,* 31-54.

Davis, B. L., MacNeilage, P. F. & Matyear, C. L. (2002). Acquisition of serial complexity in

speech production: A comparison of phonetic and phonological approaches to first word

production. *Phonetica, 59*, 75-107.

Delattre, P. C. (1965). *Comparing the phonetic features of English, French, German and*

Spanish. Heidelberg: Julius Groos Verlag.

Deng, L. & Dang, J. (2007). Speech analysis: The Production-Perception perspective. In Hai-

    Zhou Li & Chin-Hui Lee (eds.), *Advances in Chinese spoken language processing* (pp. 3–

    32). New Jersey: World Scientific.

DePaolis, R. A., Keren-Portnoy, T. & Vihman, M. M. (2016). Making sense of infant

    familiarity and novelty responses to words at lexical onset. *Frontiers in Psychology, 7,* 715,

    1-11.

Dollaghan, C. A., Biber, M. E. & Campbell, T. F. (1995). Lexical influences on nonword

    repetition. *Applied Psycholinguistics, 16*, 211-222.

Fee, J. & Ingram, D. (1982). Reduplication as a strategy of phonological development. *Journal*

    of Child Language, 9, 41–54.

Ferguson, **C.** A. (1983). Reduplication in child phonology. *Journal of Child Language, 10,* 239-

    243.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can

    recognize spoken words using partial phonetic information. *Child Development, 72,* 1003-

    1015.

Ganger, J. & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental*

    Pychology, 40, 621-632.

Grunwell, P. (1982). *Clinical Phonology*. London: Croom Helm.

Hallé, P. & Boysson-Bardies, B. de. (1996). The format of representation of recognized words

    in infants' early receptive lexicon. *Infant Behavior and Development, 19*, 463-481.

Hodges, R., Munro, N., Baker, E, McGregor, K., Docking, K. & Arciuli, J. (2016). The role of elicited verbal imitation in toddlers' word learning. *Journal of Child Language, 43*, 457-471. doi:10.1017/S0305000915000240

Huff, C. L. (2017). A study of Mandarin homophony. Unpublished PhD thesis. Oxford University.

Keren-Portnoy, T., Vihman, M. M., DePaolis, R., Whitaker, C. & Williams, N. A. (2010). The role of vocal practice in constructing phonological working memory. *Journal of Speech, Language, and Hearing Research*, *53*, 1280-1293.

Kunnari, S. (2000). *Characteristics of early lexical and phonological development in children acquiring Finnish*. Acta Universitatis Ouluensis B34, University of Oulu.

Kunnari, S., Nakai, S. & Vihman, M. M. (2001). Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication, 5,* 13-24.

Levelt, C. C. (**2011**). Consonant harmony in child language. In M. van Oostendorp, C. J. Ewen, E. V. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology*, v. 3, pp. 1691–1716. Oxford, UK: Wiley-Blackwell.

Lin, Y.H. (2007). *The sounds of Chinese with audio CD* (v. 1). Cambridge: Cambridge University Press.

Lleó, C. (1990). Harmony and reduplication: On the extended availability of two strategies in phonological acquisition. *Journal of Child Language, 17*, 267-278.

Lou, S. (2021) Early phonological development in Mandarin: An analysis of prosodic structures and tones from babbling through the single word period. Unpublished PhD thesis, University of York.

Majorano, M., Vihman, M. M. & DePaolis, R. A. (2014). The relationship between infants'
production experience and their processing of speech. *Language Learning and
Development, 10*, 179-204.

McAllister Byun, T., Inkelas, S. & Rose, Y. (2016). The A-map model: Articulatory reliability
in child-specific phonology. *Language, 92,* 141-178.

Menn, L. & Vihman, M. M. (2011). Features in child phonology: inherent, emergent, or
artefacts of analysis? In N. Clements & R. Ridouane (eds.), *Where do phonological features
come from?,* pp. 261-301. Amsterdam: John Benjamins.

Oh, Y. M. (2015). Linguistic complexity and information: Quantitative approaches.
Unpublished doctoral thesis, University of Lyon.

Ota, M. (2003). *The development of prosodic structure in early words*. Amsterdam: John
Benjamins.

Ota, M. (2013). Lexical frequency effects on phonological development: the case of word
production in Japanese. In M. M. Vihman & T. Keren-Portnoy (eds.), *The emergence of
phonology: whole-word approaches, cross-linguistic evidence,* pp. 415-438. Cambridge:
CUP.

Post, B. & Payne, E. (2018). Speech rhythm in development: what is the child acquiring? In P.
Prieto & N. Esteve-Gilbert, eds., *The development of prosody in first language acquisition,*
pp. 125-143. Amsterdam: John Benjamins.

Savinainen-Makkonen, T. (2000). Word initial consonant omissions – a developmental process
in children learning Finnish. *First Language, 20*, 161-185.

Savinainen-Makkonen, T. (2001). Suomalainen lapsi fonologiaa omaksumassa. [Finnish
Children Acquiring Phonology.] Publications of the Department of Phonetics, University of
Helsinki, 42.

Savinainen-Makkonen, T. (2007). Geminate template: a model for first Finnish words. *First
Language, 27,* 347-359. Reprinted in Vihman & Keren-Portnoy, 2013.

Schwartz, R., Leonard, L., Wilcox, M. & Folger, M. (1980). Again and again: reduplication in
child phonology. *Journal of Child Language, 7,* 75–87.

Segal O., Keren-Portnoy T., Vihman M. (2020). Robust effects of stress on early lexical
representation. *Infancy, 25*, 500–521.

Smith, N. V. (1973). *The acquisition of phonology: a case study*. Cambridge: The University
Press

Stoel-Gammon, C. & Stemberger, J. (1994). Consonant harmony and phonological
underspecification in child speech. In M. Yavas (ed.), *First and second language phonology,*
pp. 63-80. San Diego, CA: Singular.

Suomi, K. & Ylitalo, R. (2004). On durational correlates of word stress in Finnish. *Journal of
Phonetics, 32,* 35-63.

Torkildsen, J. V. K., Hansen, H. F., Svangstu, J. M., Smith, L., Simonsen, H. G., Moen, I. &
Lindgren, M. (2009). Brain dynamics of word familiarization in 20-month-olds: Effects of
productive vocabulary size. *Brain and Language, 108*, 73-88.

Vihman, M. M. (1978). Consonant harmony: its scope and function in child language. In J.
Greenberg, C.A. Ferguson & E.A. Moravcsik (eds.), *Universals of human language*, *2, 281-
334*. Stanford, CA: Stanford University Press.

Vihman, M. M. (2019). *Phonological templates in development*. Oxford: Oxford University
 Press.

Vihman, M. M. (2022). The developmental origins of phonological memory. *Psychological
 Review.*

Vihman, M. M. & Croft, W. (2007). Phonological development: Toward a 'radical' templatic
 phonology. *Linguistics, 45,* 683-725. Reprinted in Vihman & Keren-Portnoy, 2013, pp. 17-
 57.

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2014). The role of production in
 infant word learning. *Language Learning, 64: Suppl. 2*, 121-140.

Vihman, M. M., Kay, E., Boysson-Bardies, B. de, Durand, C. & Sundberg, U. (1994).
 External sources of individual differences? A cross-linguistic analysis of the phonetics of
 mothers' speech to one-year-old children. *Developmental Psychology, 30,* 652-663.

Vihman, M. M. & Keren-Portnoy, T. (eds.) (2013). *The emergence of phonology: whole-
 word approaches, cross-linguistic evidence.* Cambridge: Cambridge University Press.

Vihman, M. M. & Majorano, M. (2017). The role of geminates in infants' early words and word-
 form recognition. *Journal of Child Language, 44,* 158-184.

Vihman, M. M. & McCune L. (1994). When is a word a word*? Journal of Child Language,
 21,* 517-542.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in
 early lexical representation. *Journal of Memory and Language, 50,* 336-353.

Xia, D. (ed.) (2000). *Xinhua zidian.* Beijing: Commercial Press.

Waterson, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics, 7,* 179-211.

    Reprinted in Vihman & Keren-Portnoy, 2013, pp. 61-92.

Wauquier, S. & Yamaguchi, N. (2013). Templates in French. In M. M. Vihman & T. Keren-

    Portnoy (eds.), *The emergence of phonology: whole-word approaches, cross-linguistic*

    evidence (pp. 317-342). Cambridge: CUP.

Appendix A: Model estimates for the frequency of target words

	Estimate	Std. Error	z	p	
(Intercept)	1.018	0.074	13.82	< 0.001	***
VRG	1.479	0.076	19.46	< 0.001	***
RED	0.424	0.094	4.49	< 0.001	***
CH	-0.343	0.118	-2.91	0.004	**
MONO_LONG	-1.415	0.214	-6.60	< 0.001	***
USEnglish	-0.395	0.165	-2.40	0.016	*
UKEnglish	0.192	0.115	1.67	0.095	
Finnish	0.041	0.203	0.20	0.840	
Japanese	0.376	0.127	2.96	0.003	**
Mandarin	-0.242	0.215	-1.13	0.260	
VRG:USEnglish	0.356	0.170	2.09	0.037	*
RED:USEnglish	-0.549	0.258	-2.13	0.033	*
CH:USEnglish	0.312	0.260	1.20	0.230	
MONO_LONG:USEnglish	1.066	0.337	3.17	0.002	**
VRG:UKEnglish	-0.270	0.121	-2.23	0.026	*
RED:UKEnglish	-1.016	0.193	-5.26	< 0.001	***
CH:UKEnglish	0.249	0.182	1.37	0.171	
MONO_LONG:UKEnglish	1.321	0.255	5.17	< 0.001	***

Appendix A (cont.)

VRG:Finnish	0.421	0.202	2.09	0.037	*
RED:Finnish	-0.228	0.249	-0.92	0.359	
CH:Finnish	0.592	0.256	2.31	0.021	*
MONO_LONG:Finnish	-1.280	0.689	-1.86	0.063	
VRG:Japanese	-0.300	0.131	-2.29	0.022	*
RED:Japanese	0.370	0.150	2.48	0.013	*
CH:Japanese	-0.457	0.227	-2.01	0.044	*
MONO_LONG:Japanese	-0.004	0.332	-0.01	0.990	
VRG:Mandarin	-0.325	0.230	-1.41	0.158	
RED:Mandarin	1.349	0.226	5.98	< 0.001	***
CH:Mandarin	-0.672	0.386	-1.74	0.082	
MONO_LONG:Mandarin	-0.987	0.693	-1.42	0.154	

Notes. Significance codes: '***' 0.001 '**' 0.01 '*' 0.05

Appendix B: Model estimates for the frequency of child forms

	Estimate	Std. Error	z	p	
(Intercept)	0.892	0.075	11.88	< 0.001	***
VRG	0.307	0.092	3.35	< 0.001	***
RED	-0.529	0.116	-4.56	< 0.001	***
CH	-0.135	0.119	-1.13	0.258	
USEnglish	0.088	0.153	0.57	0.566	
UKEnglish	-0.172	0.156	-1.10	0.270	
Finnish	0.411	0.158	2.60	0.009	**
Japanese	0.161	0.143	1.13	0.259	
Mandarin	-0.756	0.233	-3.24	0.001	**
VRG:USEnglish	-0.102	0.190	-0.54	0.589	
RED:USEnglish	-0.065	0.248	-0.26	0.792	
CH:USEnglish	-0.017	0.221	-0.08	0.937	
VRG:UKEnglish	0.305	0.180	1.70	0.089	
RED:UKEnglish	-1.028	0.332	-3.09	0.002	**
CH:UKEnglish	0.283	0.208	1.36	0.175	
VRG:Finnish	-1.322	0.259	-5.12	< 0.001	***
RED:Finnish	0.207	0.219	0.94	0.347	
CH:Finnish	0.603	0.196	3.07	0.002	**
VRG:Japanese	0.340	0.160	2.12	0.034	*
RED:Japanese	-0.210	0.236	-0.89	0.372	
CH:Japanese	0.271	0.192	1.41	0.158	
VRG:Mandarin	0.916	0.256	3.57	< 0.001	***
RED:Mandarin	1.059	0.295	3.58	< 0.001	***
CH:Mandarin	-0.945	0.470	-2.01	0.044	*

Appendix C: Model estimates for the frequency of target words (structurally-adjusted)

	Estimate	Std. Error	*z*	*p*	
(Intercept)	0.858	0.076	11.23	< 0.001	***
VRG	1.042	0.085	12.27	< 0.001	***
RED	0.485	0.100	4.87	< 0.001	***
CH	-0.197	0.119	-1.66	0.097	**
MONO_LONG	-1.341	0.216	-6.22	< 0.001	***
USEnglish	-0.572	0.174	-3.28	0.001	*
UKEnglish	-0.026	0.126	-0.21	0.838	
Finnish	0.163	0.207	0.79	0.432	
Japanese	0.533	0.132	4.04	< 0.001	**
Mandarin	-0.084	0.218	-0.39	0.700	
VRG:USEnglish	-0.307	0.224	-1.37	0.170	*
RED:USEnglish	-0.503	0.281	-1.79	0.074	*
CH:USEnglish	0.497	0.264	1.88	0.060	
MONO_LONG:USEnglish	1.323	0.340	3.89	0.000	**
VRG:UKEnglish	-0.461	0.156	-2.95	0.003	*
RED:UKEnglish	-1.057	0.223	-4.73	< 0.001	***
CH:UKEnglish	0.473	0.187	2.53	0.012	
MONO_LONG:UKEnglish	1.118	0.277	4.04	< 0.001	***

Appendix C (cont.)

VRG:Finnish	0.697	0.208	3.36	0.001	***
RED:Finnish	-0.249	0.251	-0.99	0.321	
CH:Finnish	0.487	0.257	1.90	0.058	
MONO_LONG:Finnish	-1.313	0.689	-1.91	0.057	
VRG:Japanese	0.136	0.137	1.00	0.319	
RED:Japanese	0.309	0.153	2.02	0.043	*
CH:Japanese	-0.603	0.228	-2.65	0.008	**
MONO_LONG:Japanese	-0.077	0.332	-0.23	0.816	
VRG:Mandarin	0.112	0.233	0.48	0.632	
RED:Mandarin	1.288	0.228	5.66	< 0.001	***
CH:Mandarin	-0.818	0.386	-2.12	0.034	*
MONO_LONG:Mandarin	-1.060	0.693	-1.53	0.126	

Notes. Significance codes: '***' 0.001 '**' 0.01 '*' 0.05

Appendix D: Model estimates for the frequency of child forms (Structurally-adjusted)

	Estimate	Std. Error	z	p	
(Intercept)	0.245	0.109	2.25	0.024	*
VRG	0.371	0.131	2.82	0.005	**
RED	-0.728	0.208	-3.50	0.000	***
CH	-0.051	0.153	-0.33	0.739	
USEnglish	-0.686	0.305	-2.25	0.025	*
UKEnglish	-0.665	0.268	-2.48	0.013	*
Finnish	0.814	0.199	4.09	< 0.001	***
Japanese	0.786	0.176	4.46	< 0.001	***
Mandarin	-0.126	0.258	-0.49	0.625	
VRG:USEnglish	-0.197	0.408	-0.48	0.629	
RED:USEnglish	-0.485	0.664	-0.73	0.464	
CH:USEnglish	0.224	0.416	0.54	0.590	
VRG:UKEnglish	0.362	0.308	1.18	0.239	
RED:UKEnglish	-1.104	0.652	-1.69	0.090	
CH:UKEnglish	0.416	0.335	1.24	0.214	
VRG:Finnish	-1.321	0.292	-4.52	< 0.001	***
RED:Finnish	0.471	0.290	1.62	0.104	
CH:Finnish	0.445	0.226	1.97	0.049	*
VRG:Japanese	0.277	0.185	1.49	0.136	
RED:Japanese	-0.011	0.292	-0.04	0.970	
CH:Japanese	0.188	0.215	0.88	0.382	
VRG:Mandarin	0.852	0.272	3.13	0.002	**
RED:Mandarin	1.258	0.342	3.68	< 0.001	***
CH:Mandarin	-1.028	0.478	-2.15	0.031	*