UNIVERSITY of York

This is a repository copy of *A challenge to whole-word phonology? : A study of Japanese and Mandarin*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/195286/

Version: Published Version

Article:
Vihman, Marilyn orcid.org/0000-0001-8912-4840, Ota, Mitsuhiko, Keren-Portnoy, Tamar orcid.org/0000-0002-7258-2404 et al. (2 more authors) (2023) A challenge to whole-word phonology? : A study of Japanese and Mandarin. Language Learning and Development. ISSN 1547-3341

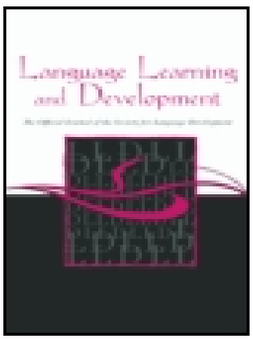
https://doi.org/10.1080/15475441.2022.2149401

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/



Language Learning and Development

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hlld20

A Challenge to Whole-word Phonology? A Study of Japanese and Mandarin

Marilyn May Vihman, Mitsuhiko Ota, Tamar Keren-Portnoy, Rui Qi Choo & Shanshan Lou

To cite this article: Marilyn May Vihman, Mitsuhiko Ota, Tamar Keren-Portnoy, Rui Qi Choo & Shanshan Lou (2023): A Challenge to Whole-word Phonology? A Study of Japanese and Mandarin, Language Learning and Development, DOI: 10.1080/15475441.2022.2149401

To link to this article: https://doi.org/10.1080/15475441.2022.2149401

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 04 Jan 2023.

Submit your article to this journal

Article views: 76

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=hlld20

LANGUAGE LEARNING AND DEVELOPMENT
https://doi.org/10.1080/15475441.2022.2149401

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

A Challenge to Whole-word Phonology? A Study of Japanese and Mandarin

Marilyn May Vihman [a,b], Mitsuhiko Ota [c], Tamar Keren-Portnoy [a], Rui Qi Choo [a], and Shanshan Lou [a]

[a]Language and Linguistic Science, University of York, Heslington, UK; [b]Linguistics, University of California Berkeley, Berkeley, California, USA; [c]School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, Scotland

ABSTRACT

Phonological models of early word learning often assume that child forms can be understood as structural mappings from their adult targets. In contrast, the whole-word phonology model suggests that on beginning word production children represent adult targets as holistic units, reflecting not the exact sound sequence but only the most perceptually salient elements or those that align with their own vocal patterns. Here we ask whether the predictions of the whole-word model are supported by data from children learning Japanese or Mandarin, both languages with phonotactic structures differing from any so far investigated from this perspective. The Japanese child word forms are found to include some characteristics suggestive of whole-word representation, but in Mandarin we find little or no such evidence. Instead, some children are found to make idiosyncratic use of whole syllables, substituting them for target syllables that they match in neither onset nor rime. This result, which neither model anticipates, forces reconsideration of a key tenet of the whole-word model – that early word production is based on word-size holistic representations; instead, at least in some languages, the syllable may serve as the basic representational unit for child learners.

A long-standing approach to accounting for the differences between word forms produced by young children and their adult targets is to assume that the child's forms are directly mapped from the adult forms. This "mapping model" was formalized using ordered rules, following Chomsky and Halle (1968), Menn (1971), Smith (1973), Spencer (1986), and Stampe (1979). It was widely adopted in clinical work in the form of "phonological processes" (cf., Grunwell, 1982; Ingram, 1976; Vihman & Greenlee, 1987) and continues to inform Optimality Theoretic approaches (e.g., Fikkert & de Hoop, 2009; Fikkert & Levelt, 2008; Gnanadesikan, 2004; McAllister Byun et al., 2016; Pater, 2004). In this model the assumption is that the child has a mental representation of the target word that is largely adultlike, in the sense that the target is accurately perceived and is analyzed in terms of adultlike phonological units such as segments and features (although in some analyses, the representation is thought to differ from that of adults due to structural constraints). The discrepancies between this "underlying form" and the "surface form" that the child produces are due to phonologically or phonetically motivated processes applied to the underlying form. For example, a consonant cluster in the target word may be reduced to a singleton because of the markedness of such a phonological structure or the articulatory difficulty of executing successive oral closures. The expectation is that the outcome of a child's attempt to produce a novel word form can readily be predicted based on the

CONTACT Marilyn May Vihman mv509@york.ac.uk Language and Linguistic Science, University of York, Heslington YO10 5DD

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

structure of the adult word and the processes the child has been observed to use. The emphasis is thus on the phonologically-motivated systematicity in the mapping between the targets and the child forms.

An alternative view is that children begin word production by representing as holistic units the adult word forms that they retain in memory (Vihman & Croft, 2007). That approach sees child words as not necessarily reflecting the full segmental and featural sequence of the target but, at least in some cases, only the most perceptually salient elements or those that memorably align with a vocal pattern in the child's own repertoire (Vihman, 1993); other elements may be more variably represented or leave a weaker memory trace. Such holistic representations are shaped not only by the structure of the target words themselves but also by aspects of the overall profile of the word forms of the language that are congruent with the child's lexical, perceptual and vocal production experiences.

The idea that child forms that fail to match their targets may reflect a combination of perceptual salience and representational reinterpretation or reorganization, in addition to level of articulatory skill, was the basis for the "whole-word phonology" approach to phonological development that developed in the 1970s (Ferguson & Farwell, 1975; Macken, 1979; Menn, 1983; Waterson, 1971; for an overview, see, Vihman & Keren-Portnoy, 2013a, 2013b). This approach to understanding early word production leads to the following claims, which differentiate the whole-word approach from the mapping model. First, many phonological properties of the child's word form are represented at the word-level rather than at sub-lexical (principally segmental) levels. Second, the child's representations of the target may not be adultlike, as they are filtered through perceptual and motoric constraints particular to the child. Third, the child form is a product not only of the target form and regular processes but also of other remembered word forms, which feed into the gestalt representations of the targets. Below, we review key observations that support each of these claims.

Evidence that early word forms represent phonological characteristics at the whole word level is found in phenomena that encompass the entire word. The most extreme cases are seen in children who produce word forms containing only consonants or vowels that share some features. For example, Fikkert and Levelt (2008) describe a Dutch-learning child who, at 1;5.11, would produce only words with either coronal consonants and coronal or low vowels (e.g., /di/ [ti] "that one," /zɛs/ [sɛs]) or labial consonants and low/rounded vowels (/pɔp/ [pɔ] "doll," /ap/ [ap] "monkey"). The pattern behind these forms is best described as "one word, one feature," a special type of harmony that encompasses all segmental units in the child's word. Similarly, Menn (1971) details the period of early word production of her son Danny, from about 22 months, during which he produced stop codas when they agreed with the onset in place (*bump* [bʌmp], *cracker* [gæk]), but assimilated coronal stops to labials or velars in either word position (*boat* [bop], *bread* [bʌb], *truck* [gʌk]). More generally, as originally pointed out by Waterson (1971), any type of long-distance assimilation of consonants (e.g., *bucket* [bæbu:]) or repetition of syllables (e.g., *biscuit* [be:be:]) can be understood as a form of phonological dependency between elements that takes the entire word as its domain of representation.

Syllable repetition is also employed to maintain the overall prosodic shape of a target word when some sub-lexical content appears to be only weakly represented. Lleó (1990) illustrates her Catalán-learning daughter's replacement, at 2;10–2;11, of unstressed syllables with material drawn and sometimes recombined from elsewhere in the target word: e.g., *A'melia* [mɛ'mɛlja], *bici'cleta* "bicycle" [ble'bletsa, ble'bleka], *tova'llola* "towel" [βo'βola] (9 examples are shown). As Lleó comments, "a general outline of the word prevails over a linear analysis of its segments" (Lleó, 1990, p. 275). Vihman (1978, pp. 316 f.) provided similar data (12–14 words) from her daughter, at 1;5–1;9, for Estonian, with its fairly consistent first-syllable stress: e.g., *tagasi* "(go, put, take) back" [tasisi], *porgandit* "carrot, partitive singular" [pɔnini], *magustoit* "dessert" [masusu] (with metathesis of /us/). Each of these patterns involves some use of syllable-size repetition, but the source of the repeated syllable is variable, illustrating ways in which children may reorganize the features of a word in what must be taken to reflect holistic representation.

With regard to the claim that early word representations are incomplete due to perceptual or motoric filters, the most relevant supporting observation is found in the tendency for children to omit word-initial consonants when another part of the word contains a structure with greater salience,

whether due to accentual prominence or the presence of geminates. For example, children learning French, which has prosodic prominence expressed as word-final syllable lengthening, often omit even early-learned word-initial consonants (e.g., Béryl, 18 mos.: *marteau* /maʁto/ "hammer" [ato], *nuage* /nɥaʒ/ "cloud" [aça], *parrain* /paʁɛ̃/ "godfather" [apa], with metathesis: Vihman, 2019, p. 148; see also Charles, in Vihman & Kunnari, 2006). Similarly, Italian children often show a < VCCV> pattern for targets with the structure CVCCV (where CC is a geminate), reflecting the effect of medial geminates on attention to the word-initial consonant (e.g., at 21 mos., A.P.[1] *quella* "that one, f." [ɛlːa], both G. A. and G.C.: *latte* "milk" [atːe], L.L.: *gallo* "rooster" [alːo], Vihman, 2019, p. 115; see also, Keren-Portnoy et al., 2010). Similar patterns have been identified in other languages with iambic stress (for Hebrew, see, Keren-Portnoy & Segal, 2016) or geminates (for Finnish, see, Savinainen-Makkonen, 2007; for Hindi; Vihman & Croft, 2007). The omission of word-initial consonants in these child production patterns is remarkable given the well-known bias for word-initial consonants in adult word processing (e.g., Marslen-Wilson, 1987) and the markedness expectation that in acquisition, as in the world's languages, CV syllables will be structurally favored over onsetless syllables.

Wauquier and Yamaguchi (2013) emphasize the role of rhythm in early child word production, with stress and syllable type/weight guiding child attention and shaping what they retain in memory (their "intake" from the "input" they are exposed to). The frequent occurrence of child <VCV> forms in French (and <VCCV> forms in Finnish and Italian) provide support for these ideas, showing that the unifying effect of accent and rhythm contributes to the whole-word patterns observed in children acquiring those languages. That is, children learning languages with a distinction between strong and weak syllables appear to develop representations that reflect the accentual envelope of the word as a whole, better retaining some parts than others and experiencing in one part an effect of elements that occur in other parts (Vihman & Croft, 2007).

Further evidence for the impact of the rhythmic properties of target words can be found in child production of multisyllabic words in Japanese (Ota, 2013). The initial syllable of a disyllabic target is more likely to be omitted in production when it is light (i.e., has neither a long vowel nor a coda) and followed by a heavy second syllable (i.e., a syllable with a long vowel or coda). A similar word-level effect has been found for English, with post-tonic syllables far less likely to be deleted than pre-tonic syllables (Allen & Hawkins, 1980; Echols & Newport, 1992; Snow, 1998); the latter are also the most likely to be replaced by a "dummy syllable" (Gnanadesikan, 2004). The tendency for Japanese disyllabic forms to be truncated to just the second syllable only when the first syllable is light is also reminiscent of initial consonant omission in children learning French or Hebrew. Note, however, that in Japanese the effect of rhythm is expressed more often in omission of a whole syllable than in omission of the word-initial consonant alone.

The perceptual effect of the accented syllable has been demonstrated experimentally in studies of children learning French (Hallé & de Boysson-Bardies, 1996), English (Vihman et al., 2004) and Hebrew (Segal et al., 2020); the perceptual effect of medial geminates was demonstrated experimentally for Italian (Vihman & Majorano, 2017). Studies have shown that most children recognize (untrained) words familiar from use in the home by 11 months (Hallé & de Boysson-Bardies, 1994; Vihman et al., 2004), but recognition is blocked if a change is made to the accented first syllable or to the initial consonant in words with medial singletons. Crucially, infant recognition of familiar words is *not affected* when what is changed is an unaccented syllable or the initial consonant in a word with a medial geminate – presumably because in those cases child attention is deflected from the word initial to the accented syllable or to the lengthened medial consonant. These experimental findings support the idea, inferred from children's word forms, that not all parts of a word are represented equally robustly.

Finally, the idea that child word forms are constructed not only from elements of the target words but also from holistic representations of the full set of words in the child's active lexicon is supported by a range of crossword effects. The most striking evidence comes from child "templates," or patterns

[1]The Italian children are named with initials (in lieu of pseudonyms).

which the child "collects" by first preferentially selecting targets of a given word shape and subsequently adapting less accessible words to fit the pattern. Priestly's (1977) account of his son Christopher's systematic imposition of a < CVjVC> structure on disyllabic words illustrates this well. The child was correctly producing words like *lion* [lajən] and *whale* [wɛjəl] at about 22 months, but over a four-month period he adapted 70 words to fit this pattern. Those forms drew freely on the consonants of the target to arrive at [kajak] for *choc'late*, [kajal] for *candle*, [mejas] for both *medicine* and *music*, and so on. We can surmise that falling back on a well-practised motoric routine enabled the child to produce challenging words without necessarily recalling all of their consonants or their positions in the word. (Note that from the second month of observation the "templatic" forms alternated with relatively accurate forms of the same words, suggesting that momentarily gaining access to the lexical representation was the problem, not an inability to produce the particular segmental sequence.[2])

Similarly, in her exhaustive longitudinal analyses of the idiosyncratic patterns of a Mexican-Spanish -learning child "Si" Macken (1979) provides evidence of repeated dynamic restructurings, which reflect, over the last months of her second year, the child's growing familiarity with the language and her own expressive lexicon. For example, the child prefers the pattern <LABIAL – CORONAL> early on, producing long words like *manzana* "apple" as [mən:a], *Fernando* "brother's name" as [man:ə, wan:o], *zapato* "shoe" as [pwat:o], *elefante* "elephant" as [bat:e], and even *telefono* "telephone" as [fəntonno] and *sopa* "soup" [p'wæt'a] (both with metathesis). Although these target words do each contain both a labial and a coronal somewhere in the word, the child imposes her preferred sequence on each of them, regardless of the target syllable sequence or the precise distribution of place and manner features; her preferred production pattern also dictates the choice of syllables to omit (e.g., for *manzana*, the unstressed word-initial syllable, with labial onset, is retained, in *zapato*, that syllable is omitted, and in *Fernando*, the labial of the word-initial consonant is retained but the syllable as a whole is not). Only words that included a labial and a coronal in the target form were produced with this pattern; other variegated targets – or target words composed of at least two different supraglottal consonants – were produced with harmony (see Fikkert & Levelt, 2008, for a similar pattern observed in Dutch-learning children).

The literature on patterning in children's early word forms has so far focused largely on a few language types: Germanic (English), Romance (Catalán, French, Italian, Spanish), Baltic-Finnic (Estonian, Finnish) and Semitic (Arabic, Hebrew). Here we turn our attention to Japanese and Mandarin, with their dramatically different word-level prosodic properties, to ask whether data from children learning those languages provides evidence for holistic word-form representation. Accordingly, we address the Research Question, does analysis of the Japanese and/or Mandarin data reveal any of the following?

(i) word-internal dependencies (e.g., movement or copying of features within the word, including consonant harmony),
(ii) effects of salient aspects of the adult target forms on the child form (e.g., retention of prominent prosodic properties alongside loss or omission of less salient aspects),
(iii) imposition of non-target elements that derive from the child's larger database of word forms.

It should be noted that some of these observations would be consistent with a mapping account as well. For example, consonant harmony can be explained as spreading or copying of a place feature in the target word (Fikkert & Levelt, 2008; Goad, 1997; Menn, 1978; Pater & Werle, 2003; Rose, 2000; Stoel-Gammon & Stemberger, 1994). Omission of non-prominent prosodic properties from the target has

[2]Priestly (1977) lists the child's 17 "bisyllabic ordinary forms" (such as *whale* [wɛjəl]) that were in use before the novel "experimental" forms began to be noted, as well as all 70 "bisyllabic experimental forms" (in alphabetic order) that he observed over four months. Vihman (2019, with thanks to Florence Oxley) provides a chronologically ordered list, which better reveals the dynamics of the child's production over time.

also been addressed in analyses that appeal to the idea that the child form initially permits only the minimal structure that will meet the well-formedness requirements of a prosodic word (i.e., a single foot, which, in English, consists of a stressed syllable followed by an optional unstressed syllable) and materials falling outside this template are not realized in the output (Demuth, 1995; Fikkert, 1994; Ota, 2003; Pater, 1997). However, other observations predicted by the whole-word model pose challenges for the mapping model. For example, the deletion of the word-initial consonant in CVCCV targets discussed above is difficult to explain through a mapping model, because there are no phonological processes by which the presence of a syllable coda (i.e., C_2 in $C_1VC_2C_3V$) can trigger the removal of the onset (C_1). It is even more difficult to see how a mapping model could account for imposition of non-target elements that originate in the child's pool of remembered words. We are therefore testing the overall weight of evidence for the whole-word model rather than a set of unique predictions associated with it.

Structure of Japanese and Mandarin

The inclusion of Japanese and Mandarin is motivated by a number of specific factors. First, Japanese and Mandarin differ in ways that are informative for the purpose of our investigation. Japanese has contrastive length in vowels (like Finnish) and consonants (like Finnish and Italian), while Mandarin has neither. Japanese is a pitch accent language, which marks a syllable in some words with a high pitch followed by a low pitch. Like stress, pitch accent creates a syntagmatic contrast between the more prominent (i.e., accented) and the less prominent syllables in a word (Beckman, 1986). In contrast, Mandarin is a tone language that assigns to each syllable one of four tones (conventionally marked in phonemic transcription as 1 high level, 2 rising, 3 fall-rise, 4 falling) or a neutral tone. Lexical tones involve paradigmatic contrasts, in the sense that each syllable is distinguished by its tonal quality rather than prominence. The implication of these differences is that we expect Japanese, but not Mandarin, to exhibit prominence-induced phenomena in early word production. Indeed, as noted above, Japanese-learning children are less likely to omit syllables that are accented or prosodically "heavy," suggesting that these properties of the language create differing levels of salience within the target word. What remains to be investigated is whether they also lead Japanese learners to show whole-word phenomena such as the deletion of onset consonants in the presence of geminates or accented syllables, as reported for languages such as Finnish, Hebrew, French and Italian .

Second, the two languages both differ from other languages previously investigated from the perspective of whole-word phonology in that they have much simpler syllable structures. The only onset clusters permitted are consonant + glide (Cj in Japanese, Ci(V)/Cu(V) in Mandarin) and the only non-geminate coda allowed is a nasal (in Japanese, this is a "placeless" nasal that agrees with the next syllable onset, if any, in place of articulation, or that is otherwise realized as a uvular; in Mandarin, a coda nasal is alveolar or velar). The upshot is that both languages have noticeably smaller syllable inventories compared to Germanic and Romance languages, for example. Japanese is estimated to have just over 600 unique syllables (Oh, 2015), while Mandarin reportedly has an inventory of some 400 unique syllables only, if tone is disregarded (Deng & Dang, 2007); this is in sharp contrast with English, which has some 9000 (Huff, 2017), or over 12,000 (according to Farrell & Abrams, 2011). This raises the possibility that syllables may play a more important role in the representation of early words in these languages, as repeated exposure to a smaller number of distinct syllables is likely to better support access to them.

Another important consequence of the different prosodic properties of Japanese and Mandarin is that in Japanese (and also in languages with stress accent) the status of individual syllables is defined by the word – i.e., one syllable per word receives an accent or (primary) stress – whereas in Mandarin (and other languages with lexical tone), the individual syllable is representationally independent in prosodic structure. A key factor that contributes to this independence is that the syllables of a multisyllabic Mandarin word are also often morphemes in their own right, most of which can also occur, with the same or a similar meaning, as monosyllables. Additional evidence of the independence

of the syllable is that Mandarin exhibits smaller between-syllable coarticulatory effects than are observed for French (L. Ma et al., 2015; see also, Mok, 2010 for a similar contrast between Thai, another tone language, and English) and weaker priming effects for initial syllables than are found for English (Chang et al., 2022). All of these observations suggest that the syllable may play a more important role relative to the word in Mandarin compared to Japanese. Despite these differences, the whole-word model, which assumes that infants learn words, not sounds, would predict that children learning both Mandarin and Japanese should show the kinds of evidence of whole-word representation that we indicated above.

Method

We draw here on longitudinal data from children acquiring Japanese (N = 7) and Mandarin Chinese (N = 5). Our Japanese data was collected in California and in Washington D.C., our Mandarin data in Yorkshire, England. Although the communities where the Japanese and Mandarin data were collected are English-dominant, the children recorded come from largely monolingual homes and none of them produced more than four English words in the recorded sessions sampled here. The data were audio- and video-recorded in free-play situations in the children's homes. A native speaker of each language carried out the recordings and subsequently transcribed the child's vocalizations. In all three sets of recordings word identification was based on close observation of the video to ascertain the situational context at the point when the vocalization was produced, along with the direction of the child's attention and any accompanying gestures, following the principles laid out in Vihman and McCune (1994). (For more detail, see De Boysson-Bardies & Vihman, 1991; Lou, 2021; Ota, 2003)

We established a set lexical level for the children to be included in the sample, to ensure that they would be developmentally comparable across the two language groups. We chose the half-hour recording session in which the child first produced 25 or more words spontaneously, corresponding to a cumulative vocabulary, by maternal report, of 50 words or more (Vihman & Miller, 1988). We consider this to be the earliest point in word production when we can make reliable generalizations based on a sufficient number of words spontaneously produced by children (in type count). These lexically-defined sessions have also generally been taken to mark the end of the single-word period, as word combinations occur in small numbers for the first time at this point or soon thereafter (Vihman, 2019; Vihman et al., 2022).

For three of the Japanese children – Hiromi, Kenta and Takeru – sessions were over an hour long. Accordingly, we selected the earliest sessions in which the child produced no more than about 25 words in 30 minutes, but we included in our analyses all the words produced in the session. (We also include imitations in our analyses, but these do not count toward our estimates of vocabulary size.) Our analyses cover both the target words for the children's forms and those forms themselves. For the language samples, child ages and numbers of words produced, see, Table 1. As a basis for analysis we focus on disyllables, which occur in roughly the same proportion in Japanese and Mandarin (64% and 61%, resp.).

Analysis and results

In what follows we look for whole-word phenomena in both Japanese and Mandarin, based on the criteria set out in the introduction. This will include, first, forms of intra-word repetition: We consider the nature of repetition in the two languages, looking for evidence of w*ord-internal dependency*, which would suggest holistic representation, in which some parts of the word affect the child's treatment of other parts. Secondly, we evaluate the effects on the children's forms, if any, of *perceptually salient aspects of the adult target forms*. And, finally, we ask whether there is evidence of *imposition of non-target elements that originate in other words in the child's repertoire*, an effect that would be difficult if not impossible to account for under the mapping model.

Table 1. Language groups and child names and ages.

Language group; child pseudonym	Child age in months	Total words produced	Spontaneous words produced
Japanese			
Haruo	19	38	30
Hiromi	21	37 (+ 2 Eng.)	37*
Kazuko	15	37 (+ 1 Eng.)	27
Kenji	19	25	24
Kenta	27	72	67*
Takeru	21	48 (+ 1 Eng.)	46*
Taro	23	32	28
Mandarin			
Didi	17	33 (+ 1 Eng.)	33
Keke	18	34 (+ 2 Eng.)	34
Shi	17	39	39
Xinyu	17	36 (+ 3 Eng.)	36
Yiyi	17	34 (+ 4 Eng.)	34

Note. Data sources: (De Boysson-Bardies & Vihman, 1991; Ota, 2003; Lou, 2021).
*Twenty-five spontaneously produced words identified in the first 30 minutes of hour-long recording; total spontaneous words recorded indicated here.

Word-internal dependencies: repetition of elements of the word in Japanese and Mandarin

We begin our analysis of possible evidence of holistic representation by reviewing how frequently and through what processes children learning each of the languages produce output forms that contain identical consonants ("consonant harmony") and/or identical syllables ("reduplication"). The distinction between consonant harmony and reduplication depends on the operationalization of these structures. For example, Japanese [bap:a] for *rap:a* "horn" fits the criteria for "reduplication,"[3] if it is based solely on the output form, but only that for "consonant harmony" if it is based on evidence of assimilatory processes, given that the two vowels are the same in the target as well as in the child form. As a first step to navigate these complications, we considered only child forms produced for *variegated* disyllabic target forms – disyllables composed of at least two different supraglottal consonants – and identified outputs that contain identical consonants (consonant harmony) or identical syllables (reduplication). For the reduplicated output forms, we identify reduplication as the process operating to create them only when the vowel has also undergone change (see, Tables 2 and 3, below). Thus, every form classified as either consonant harmony or reduplication has at least undergone consonant assimilation.[4]

Figure 1 presents the proportions of child forms of variegated disyllabic target words produced with consonant harmony (CH) or reduplication (RED). Japanese-learning children produce more outputs with CH (mean: 21.5%) than with RED (9.6%), while Mandarin-learning children produce more outputs with RED (29.0%) than with CH (4.7%). When these two types of structure are combined, 30 to 35% of the variegated disyllabic targets are produced as either CH or RED in both groups, which means that a sizable portion of the child forms produced by Japanese and Mandarin children contain identical consonants even when the targets do not. As we shall show below, however, these output forms have qualitatively different characteristics in the two languages.

Japanese. Table 2 presents all the Japanese child forms produced with either CH or RED for a variegated target. Here we show both the output-based classification and a process-based classification. That is, we indicate, for reduplicated forms, whether the process can be seen to reflect whole-

[3]Note that differences in voicing are disregarded here, as control of voice onset time is not generally mastered at this age (based primarily on studies of English and Spanish: Macken, 1980). In fact, we see child forms varying here between voiced and voiceless stop for the same target word in both the Japanese and the Mandarin data.

[4]We disregard tone differences in identifying reduplicated Mandarin forms as there is evidence from studies of both children (Choo, 2022) and adults (Chang et al., 2022) that Mandarin tonal and segmental sequences are independent in word processing and production.

Table 2. Reduplication and harmony for variegated disyllables in Japanese child forms.

Child name	No.	Target and gloss	Child form	Output form: CH or RED?	Process: CH or RED?	Non-target element (or other process)
Haruo	1	*do:zo "please"*	dʲɔ:dʲɔ:	RED	CH	palatalization
	2	*dʑɯ:sɯ "juice"*	dʲʊtʲʊ:, tʃɯtʃɯ:	RED	CH	
	3	*tatɕ:i (im.) "stand up!"*	tʰɐ̟:tʰʊ̟	CH	CH	V change
Kazuko	1	*dak:o "hug (me)"*	gʌkːɔ:	CH	CH	
	2	*do:zo "please"*	dɔ:dɔʔ	RED	CH	
Kenji	1	*batɕ:i (im.) "dirty"*	babɪ, bʌpːɪ	CH	CH	
	2	*biribiri "electric buzz"*	bɪbːɪ	RED	N/A	truncation
	3	*dak:o "hug (me)"*	gʌʔkʰɔ	CH	CH	
	4	*denwa "telephone"*	biba	CH	N/A	**syllable substitution**
	5	*genki "feeling good"*	gekɪ	CH	N/A	coda omission
	6	*kokokara "from here"*	kɔka	CH	N/A	truncation
	7	*kirekire "pretty"*	kɪekʰɛ	CH	N/A	liquid omission, truncation
	8	*pinpɔ:n "sound of doorbell"*	biba	CH	N/A	coda omission
Kenta	1	*arigato "thank you"*	do:do:	RED	RED	truncation
	2	*arigato "thank you"*	dad:o:	CH	CH	truncation
	3	*ɕo:bo:ɕa "fire engine"*	bo:m:bo:	CH	CH	truncation, coda insertion
	4	*gjɯ:njɯ: "milk"*	di:di	RED	CH	velar fronting? or **syllable substitution?**
	5	*naɕi "pear"*	dad:e	CH	CH	stopping
	6	*naɕi "pear"*	nane:	CH	CH	
	7	*nanda "what is it?"*	nainai	RED	RED	rime change, cluster simplification
	8	*nanda "what is it?"*	na:no, nan:o	CH	CH	vowel change, cluster simplification
	9	*raiõɴ "lion"*	na:jon	CH	CH	
	10	*rap:a "horn"*	bap:a	RED	CH	
	11	*ri:sɯ "wreath"*	gig:n	CH	CH	liquid stopping, backing
	12	*riŋgo "apple"*	dʏ:di	CH	CH	velar fronting, coda omission, **V metathesis**
	13	*sato: "sugar"*	dado	CH	CH	
	14	*tɕi:zɯ "cheese"*	di:dn	CH	CH	deaffrication or stopping
	15	*toire "toilet"*	de:de	RED	RED	liquid stopping
	16	*ton:erɯ "tunnel"*	nɪ:nːɪ	RED	RED	truncation
Takeru	1	*dare "who?"*	dedæ	CH	CH	**V metathesis**
	2	*dekita "did it"*	deta:	CH	N/A	truncation
	3	*gat:ai "combine"*	dat:ai	CH	CH	
Taro	1	*denɕa "train"*	dʑɛ:dʑaʔ	CH	CH	coda omission
	2	*dʑɯ:sɯ "juice"*	dɔ:tʲə	CH	CH	vowel change, stopping, **palatal metathesis**
	3	*dʑɯ:sɯ "juice"*	ɟiɟi	RED	CH	
	4	*riŋgo "apple"*	niŋgɔ	CH	CH	

Note. Im. imitated. **Bold face** marks the processes that point to non-target material in the child form (syllable substitution) and also (long-distance) metathesis, a whole-word process. Differences in VOT are disregarded in ascertaining match to onset C. Only vowel changes involving at least two features are marked as such.

syllable reduplication or, instead, simply harmonization of the target consonants. For example, where target [do:zo] gives the child form [dʲɔ:dʲɔ:] (Haruo, 1), we classify the process as CH, even though the output form contains two identical syllables, as no assimilatory process need be invoked to account for the vowels. We also note additional processes that are required to relate child form to target.

In addition to other regular processes, such as coda omission, cluster simplification or stopping of fricatives or liquids, most of the forms in Table 2 involve straightforward consonant assimilation, as is typical in other languages. Some of the Japanese reduplicated or harmony forms select a portion of the adult target that includes harmonizing consonants, omitting the rest (*biribiri, kokokara, kirekire* [Kenji, 2, 6, 7] and *dekita* [Takeru, 2]); others adapt the variegated target to repeat a particular place (e.g., labial or velar harmony in Kenji's forms) or manner of articulation (most of the others). Also, the

Table 3. Reduplication (and harmony) for variegated disyllables in Mandarin child forms.

Child name		Target and gloss	Child form	Output form: RED or CH?	Process: RED or CH?	Non-target element (or other process)?
Didi	1	*tsai4ʈʂe4* "it's here"	**tu**4**tu**4	RED	RED	**syllable substitution**
	2	*tsai4ʈʂe4* "it's here"	ta3**tu**4	CH	CH	deaffrication, rime change, **syllable substitution**
	3	*tʊŋ4tʊŋ4* "hurt"	**tu**4**tu**4	RED	N/A	coda omission
Shi	1	*pa1tiɛn0* "8 o'clock"	pa1pa0	RED	RED	
	2	*pən4tɛn4* "silly"	pai4**pai**4	RED	RED	**syllable substitution**
	3	*ɕi1kua1* "watermelon"	**p'ɤ**1p'ɤ1	RED	RED	cluster simplification (ku(V) > p)? Or **syllable substitution?**
Xinyu	1	*piŋ3kɛn1* "biscuits"	**ka**1ka0	RED	RED	**syllable substitution**
	2	*piŋ3kɛn1* "biscuits"	pei1**pei**1	RED	RED	**syllable substitution**
	3	*ʈʂɤ4kɤ0* "this"	ku4ku0	RED	CH	vowel change

Note. **Bold face** marks the inclusion of non-target material in the child form (syllable substitution). Only vowel changes involving at least two features are marked as such.

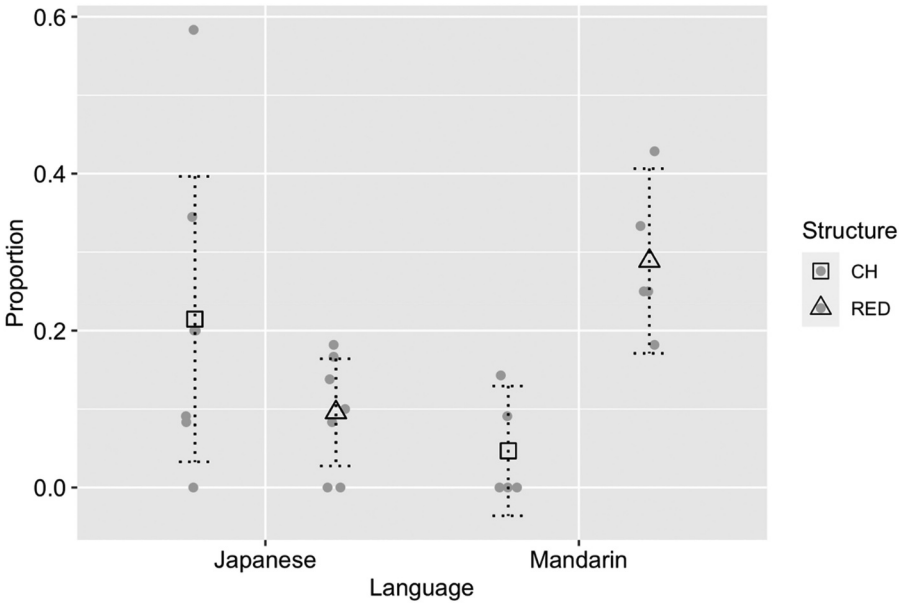


Figure 1. Proportion of consonant harmony and reduplication in child productions for variegated targets. Gray circles represent individuals, squares are means for consonant harmony and triangles are means for reduplication. Error bars show 95% confidence intervals.

Japanese children sometimes show (inconsistent) palatalization of alveolar consonants (e.g., Haruo, 1 and 2 but not 3). However, four or five of the 36 child forms in Table 2 include more idiosyncratic departures from the target: palatal metathesis, vowel metathesis and one or two instances of what appears to be wholesale syllable substitution. These processes all suggest holistic representation, as we discuss below; we will give separate attention to the imposition of whole non-target syllables.

Mandarin. Table 3 shows the CH and RED forms produced for variegated targets in Mandarin; there are only two cases that could be taken to show application of a consonant harmony process. (Two children produced neither reduplicated nor harmony forms for any variegated target.) The data look different here: Although many child forms are reduplicated, only two of the target words – *tʊŋ4tʊŋ4* (Didi, 3) and *ʈʂɤ4kɤ0* (Xinyu, 3) – have the same vowel in both target syllables.

The only apparent cases of application of harmony to a variegated word in Mandarin are Didi's [ta3tu4] for *tsai4ʈʂɤ4* and Xinyu's (reduplicated) [ku4ku0] for *ʈʂɤ4kɤ0*. To treat Didi's form as harmony would be an overinterpretation, however, as Didi substitutes [t] for most of the target affricates that he attempts and tV syllables occur in 12 of the 35 words he produces in the session (34%). In fact, retroflex /ʈʂ/ occurs four times in Didi's data, in monosyllabic *ʈʂʰi1* "eat" [ʈʂʰi1] and *ʈʂuo1* "table" [ʈʂu1, ʈʂu4] and as a second-syllable onset for a first-syllable target occurrence in *ʈʂuo1tsi0* "table" [tɤ1ʈʂai4] (in a rare instance of apparent metathesis in Mandarin). Assessing the extent of non-target elements in these data, we find that wholesale syllable substitution affects five or six of the nine variants listed in Table 3.

Thus, children learning both Japanese and Mandarin sometimes make heavy use of repetition of all or part of the word in producing variegated words, but they differ in the type of repetition they resort to, the Japanese children tending to create harmony forms while the children acquiring Mandarin rely mainly on reduplication. Furthermore, a possible non-target syllable replacement occurs in only one or two Japanese forms but is observed in most of the Mandarin forms. We discuss these non-target additions to the child forms below.

Other whole-word effects: Japanese

Palatalization. In Table 2 we saw at least one variant in which a palatal element migrates over the word, suggesting whole-word representation (Taro, 2 *ʤɯ:sɯ*). Two additional occurrences of palatal metathesis (where neither reduplication nor harmony is involved) can be seen in Kenji [tuʔʃiʔ, zɪˌʃɪ] for *ʤɯ:sɯ* "juice," Takeru [ciga, cikæ, cigʲa] for *kiɕa* "locomotive." These examples suggest that the palatal feature may be learned as part of the word as a whole rather than in association with a single target segment.

Whole-word template. One Japanese child uses a melodic template of the kind Macken (1979) described for Spanish, but with the reverse order, <CORONAL – NON-CORONAL> (Table 4). Note that in most cases, the second element of the child form is a labial, consistent with the observation that there is a bias in favor of coronal-labial sequences in Japanese infant-directed speech (Gonzalez-Gomez et al., 2014), but contrary to what might be expected from the tendency of Japanese children to target labials in higher proportion in word-initial position than overall (De Boysson-Bardies & Vihman, 1991, Fig. 3). In Table 4 the child's preferred manner-based sequence fits some targets, which may be considered "selected" for the opportunity they give the child to produce a sequence he finds articulatorily accessible (*e.g., denwa*); in other cases the target is adapted more or less radically to fit the child's pattern (e.g., *banzai*), sometimes in wholly unpredictable ways ("anomalous"). These forms encompass eight of the 39 words Haruo attempts (20%).

Other whole-word effects: Mandarin

We have failed to identify whole-word effects in the Mandarin data (although one isolated Mandarin child form seemed to show production, in the second syllable, of a consonant associated with the first syllable in the target: Didi's *ʈʂuo1tsi0* "table" [tɤ1ʈʂai4], mentioned above; see, Table 7, below); no instances were observed in which a fixed word-size pattern was imposed on a child form.

Table 4. Haruo's template, < CORONAL – NON-CORONAL>.

Target and gloss	Child form	Adapted or selected?	Process involved
banzai "cheers"	dɑba:ɪ, dæba:ɪ	adapted	coda omission, metathesis
bebi "baby"	ᵻdɛ:bɪ:	adapted	(anomalous)
bɯbɯ "car"	zigɯ:	adapted	(anomalous)
denwa "telephone"	daba:	selected	coda omission, glide stopping, vowel change
doko "where"	doko	selected	
dʒampɯ "jump"	dʲap:ɯ:	selected	deaffrication, coda omission
doN "falling-down sound"	dʌm	adapted	(anomalous)
nɯgɯ "take off clothes"	nɯgɯ	selected	

Effects of accent or duration on segmental sequences

Despite the fact that Japanese target words include medial geminates, a pattern found to result in initial consonant omission or glottal-stop substitution in several languages (Arabic: Khattab & Al-Tamimi, 2013; Estonian: Vihman, 2016; Finnish: Savinainen-Makkonen, 2007; Hindi; Vihman & Croft, 2007; Italian, Vihman & Majorano, 2017), Japanese children make little use of the <(ʔ)VC: V> pattern. Out of 107 disyllabic target words with supraglottal word-initial consonants, we find omission of that consonant in eight child word forms: Haruo *dak:o* "hug (me)" [ʔak:a], *tatɛ:i* "stand up!" [ʔʌᵗʧʊ̥], Kenji *mot:o* "more" [ʔatʌ, at:a:, hʌ̥ta:], Kenta *fo:ku* "fork" [oʔn], *zo:sãɴ* "elephant" [odon], Takeru *ɕɯɛ:ɯ:* "choochoo" [içʔʑi], *tot:e* "take (it)" [ʔotæ:], Taro *pakuɴ* "biting sound" [hapŋ, apŋ] (with metathesis); four of these have a medial geminate, all but one have a long segment. In addition, Takeru produced a range of different initial consonants for the target word *mo:ik:o* "one more" – [dok:o], [lak:o], [lok:o], [lɔk:o], [mok:o] (twice) and [nɔk:o], which also suggests an effect of the geminate on attention to or representation of the word-initial consonant. This provides some evidence of target geminates deflecting attention from the word-initial consonant in Japanese.

For Mandarin, none of the children in our database show omission of word-initial consonants. However, Choo (2022) reports a considerable proportion of word-initial consonant omissions in Yan Min, a Mandarin-dominant child she followed from age 1;7 to 2;3 in Singapore. In the first recording session Yan Min omitted 10 word-initial consonants out of the 30 disyllables she produced (33%). However, 6 of these words involved a consonant type that she was not yet producing (fricative, affricate or /l/), so that the articulatorily unexplained omissions amount to only 13%. None of the other five Mandarin-dominant children Choo recorded in Singapore showed even that extent of omission of word-initial consonants. Choo notes that Yan Min's real name was of the form <VCV>, which may have provided a personal template for omitting C_1 in her case. There is no evidence that some kind of prominence in other parts of these words could be contributing to Yan Min's consonant omission.

Imposition of non-target syllables

To summarize our findings so far, in Japanese we identified some weak evidence of whole-word effects: We saw consonant harmony and one case of template use and we noted cases in which segmental length elsewhere in the word affected children's response to (or representation of) the word-initial consonant. In Mandarin we found no evidence of template use or of word-initial consonant deletion in relation to other parts of the word. We did see intra-word repetition, in the form of relatively frequent use of reduplication, although not consonant harmony. However, when we looked at the reduplicated syllables in more detail, we found that although the consonant of the repeated syllable is in most cases related to one of the target consonants, the syllable as a whole does not always originate in the target form. We will now consider whether this type of reduplication can or should be seen as a whole-word effect. We begin by discussing the status of syllables in Japanese and Mandarin.

We suggested in the Introduction that the small syllable inventories of Japanese and Mandarin might enable children learning these languages to use frequently occurring or motorically accessible syllables as phonological "building blocks" to represent words. In this section we evaluate evidence for this possibility by examining cases in which target syllables are unpredictably substituted by an idiosyncratic but recurrent syllable in the child form. Note that we identify whole-syllable substitution only where we find at least one case of adaptation, or in other words, in cases where, in at least one word, neither onset nor rime matches the target; we then look for evidence that the syllable itself occurs repeatedly in the child's other word forms, whether the syllable is present in the target (i.e., the forms are "selected" for the syllable) or not.

Japanese. In the Japanese dataset we find some instances that look like syllable substitution, but each of them can readily be attributed to other processes. We find one such case in Haruo's data, two in Kenta's and one in Kenji's (see Table 5, which shows all the Japanese output forms in which the child appears to have

imposed non-target syllables on the target word). Haruo's substitution of [zi] for the first syllable of *bɯbɯ* is an anomaly, as there are no other occurrences of [z] in his word production; instead, the substitution fits in with that child's larger whole-word pattern or template (see Table 4). Kenta includes the syllable [di, de] in targets where it is not licensed, but he includes the coronal stop more generally in a number of words and followed by different vowels (e.g., *ana* 'hole' [dana], *haɛi* "bridge" [dadi], *baikīmmãɴ* "Baikinman (cartoon character)" [badi]). This suggests that the instances in Table 4, like the ones above, actually reflect segment substitution, with (inconsistent) fronting of velars (/g/ > [d]), liquid stopping (/ɾ/ > [d]) and vowel lowering (/i/ > [e]).[5]

The inclusion of [bi] in Kenji's production of *denwa* might be considered different from the examples given for Haruo and Kenta, as there is no "whole-word pattern" here and Kenji does produce the syllable [bi] or [pi] in four other words (Table 6), indicating its motoric familiarity for him. For comparison, only four of

Table 5. Non-target syllable substitution: Japanese.

Child	*Target and gloss*	*Child form*	*Other variants of these words*
Haruo	*bɯbɯ* "car"	**zi**gɯ:	bʊb:ɯ:, βɯβɯ:, bʊ, ʊ:::::::
Kenji	*denwa* "telephone"	**bi**ba, **bi**b:a	bi, ba
Kenta	*gjɯ:ɲjɯ:* "milk"	**di:di**	du:ʒə
	pampɯkĩɴ "pumpkin"	ba:**de:**	badi:
	ri:sɯ"wreath"	**de**gn (im.)	gig.gn, gik.kn

Note. ***Bold face*** marks the putative non-target syllables.

the other Japanese children produce [bi] at all, three of them in their form for the target word *bebi* "baby"; other targets are *pipi* "bird" and *pitɕapitɕa* "splashing sound"; the one remaining [anomalous] child form is Kazuko's *[o]meme* "eyes," produced as [**bɪ**:mɪʔ] alongside [memeʔ] and [jɔʔmehɪ:]. (The rarity of occurrence of the syllable is consistent with De Boysson-Bardies & Vihman, 1991, who found – in a study that included four of the seven children whose word forms make up our database – that Japanese children produced significantly fewer labials than English and French learners and that their production of labials steadily decreased over the course of the single-word period.) On the other hand, the second-syllable labial glide in *denwa* could also account for the production of word-initial [b] (with consonant harmony as well as stopping), followed by a one-feature vowel change. In short, Kenji's 18% use of the syllable [bi/pi] (five of 28 word types produced in the session) provides the only evidence of imposition of a non-target syllable in our Japanese data; the case is debatable. In short, on close inspection even the examples cited from Kenta and especially Kenji – infrequent instances of imposition of a non-target syllable on target words that lack the syllable – provide weak evidence at best of idiosyncratic use of familiar syllables as generalized building blocks for word forms.

Mandarin. In contrast with Japanese, imposition of non-target syllables is relatively common in our Mandarin data, as is evident from Table 3, where most of the Mandarin forms involve syllable substitution. For example, Didi produces the reduplicated sequence [tutu] for *tʊŋ4tʊŋ4* "hurt" (with coda omission) but also for *tsai4ʈʂɤ4* "it's here," along with the partially reduplicated form [ta3tu4]. Here, as in the template described for Haruo, a preferred child pattern or production routine may reflect (implicit) "selection" of targets that conform to it (here, the reduplicated target) or, more strikingly, "adaptation" (or assimilation) of targets to fit the routine (as in *tsai4ʈʂɤ4*). In the latter case the evidence of a child's representational bias toward the well-practised pattern is particularly clear and suggests an idiosyncratic dependence on it that we can term "templatic."

The Mandarin-learning children make between 13% and 28% use of one particular syllable, allowing for variation in VOT and minor shifts in the vowel (Table 7). Table 7 shows, for each child, the syllable that the child most often uses to replace adult target syllables (and the number of

[5]Note that these forms are not included in Table 2 because (i) neither *ana* nor *haɛi* are variegated targets and (ii) [badi] is neither a reduplicated nor a harmony form.

Table 6. Kenji's uses of [bi] or [pi] in target-related forms (selected).

Target and gloss	*Child form*	*Other variants of these words*
batɕ:i (im.) "dirty"	ba**bi**, bʌ**p:ɪ**	
biribiri "electric sound"	**bɪbɪ, bib:i**	
kop:ɯ "cup"	kʌ**p:ɪ:**	
pinpɔ:n "sound of doorbell"	**bi**ba, **p**ɪpa:	

Note. ***Bold face*** marks the putative non-target syllables [bi/pi].

disyllabic words attempted and proportion accounted for by use of the syllable). We distinguish "selected" and "adapted" uses according to whether only part of the target syllable or the entire syllable is replaced by non-target material. We also list any additional child variants for the target words involved in the putative syllable substitutions.

As can be seen in Table 7, all the children provide one or more examples of full syllable substitution ("adapted" forms); these account for 15 of the 40 words in question (38%). In each case the putative templatic syllable occurs in a reduplicated form as well as in other forms.

Some of these non-target substitutions are more radical than others. Keke replaces the fricative-onset syllable [ɕiɛ], in a reduplicated target, by what we can call her templatic syllable

Table 7. Non-target syllable substitution: Mandarin.

Child name; syllable (N target words affected; % over all targets)	*Target and gloss*	*Match onset or rime (Select)*	*Match neither onset nor rime (Adapt)*	*Any other child variants of these words*
Keke kɤ, kʰɤ (8; .22)	*kai4* "build"	**kɤ**1		kɤɑ1, kɤɑ4, kɤə4
	kai4kai0 "build"	**kɤ**1**kɤ**1		
	kʰan4 "look"	**kʰɤ**4		
	kʰɤ3kʰɤ3 "child's name"	**kɤ**1**kɤ**4, **kɤ**3kɤə3, **kʰɤ**1tɕʰiɛ4		tʰi1ɕiɛ4, ti1tɕiɛ0, tɛ1tɕiɛ0
	ɕiɛ2ɕiɛ0 "shoes"		**kʰɤ**2ɕiɛ2, **kʰɤ**2ɕi2	tʰi1ɕiɛ4, ti1tɕiɛ0, tɛ1tɕiɛ0
	ɕiɛ4ɕiɛ0 "thanks"		**kʰɤ**2tɕʰiɛ4	ɕiɛ1ɕiɛ4,tʰi1ɕiɛ4, ti1tɕiɛ0, tɛ1tɕiɛ0
	xɤ2tsi0 "box"		**kʰɤ**3kɤ1	
	ʈʂɤ4 "this"		**kɤ**4	*ʈʂɤ*4
Didi tɤ, tʰɤ (8; .23)	*ti4ti0* "younger brother"	**tɤ**1**tɤ**4		
	ʈʂʰɤ1 "car"	**tʰɤ**4		
	ʈʂʰɤ1ʈʂʰɤ1 "car"	**tɤ**1**tɤ**4		
	ʈʂʰɤ4li3 "here"		**tɤ**2ji2	
	ʈʂʰɤuo1tsi0 "table"		**tɤ**1*ʈʂ*ai4	
	ʈʂɤ4 "this"	**tʰɤ**4		tsɤ1te4, tsɤ4kɤ0
	ʈʂʰɤ4kɤ4 "this"	**tɤ**1kɤ4		
	tsui3pa1 "mouth"		**tʰɤ**1pa1	xa3pa4
Shi tɕi, tɕɛ, tɕe (12; .28)	*ʂou3* "hand"		**tɕɛ**4	
	tɕʰi3 "take"	**tɕi**2		
	tɕʰy4 "go"	**tɕi**4		
	tɕi1 "chicken"	**tɕi**1, **tɕi**4		
	tɕɛ3tɕɛ0 "older sister"	**tɕɛ**1**tɕi**4, **tɕɛ**2**tɕi**4, **tɕɛ**3**tɕi**4		
	tɕin1tʰiɛn1 "today"	**tɕi**4ɕi0		
	ti4ti0 "younger brother"	**tɕi**4		
	ti4ti0 "younger brother"	**tɕi**1**tɕi**0, **tɕi**4**tɕi**0		

(*Continued*)

Table 7. (Continued).

	tui4 "right"		**tɕi**4	te4
	ʈʂɤ4ʂi4 "this"		**tɕi**1te1	
	ʈʂʰɤ1 "car"		**tɕe**1	
	ʈʂu1 "pig"		**tɕɛ**1	
Xinyu	*pa1tiɛn3* "8 o'clock"	pa1**tia**4		pa1ta4
ti(a, ao, u), tʰi (5; .13)				
	ti4ti0 "younger brother"			
	tɕiɛ3tɕiɛ0 "older sister"	**tʰi**1tɕi1		te1tɕiɛ4
	ʈʂɤ4 "this"		**tiu**4	tɤ4, ta1, tou4
	ʈʂɤ4kɤ0 "this"		**tiao**4kʂɤ4, **tiao**4ku4	ku4ku0, ta4kɤ0
Yiyi	*pʰin1* "put together"		**tʰɛn**1	tsʰɛn1
tɛn, tʰɛn, təŋ (7; .18)				
	təŋ1 "light"	**tɛn**1		
	tɛn4tɛn0 "egg"	**tɛn**4**tɛn**4		
	tʰəŋ2 "hurt"	**tʰɛn**1		
	tiau4lə0 "drop"	**tɛn**2nə4		
	ta2ni3 "hit you"	**tɛn**2ne4		
	taŋ1mu3 "Tom (cat)"	**təŋ**1na1		

Note. The templatic syllable is in **boldface** in child forms. Differences in VOT are disregarded in ascertaining match to onset C.

[kʰɤ], in two words that take the segmental form *ɕiɛɕiɛ*, "shoes" and "thanks" (but that differ in their tone sequence); she also replaces both syllables of the variegated word *xɤ2tsi0* "box" with the sequence [kʰɤ] + [kɤ] and she makes a similarly radical substitution for *ʈʂɤ4* "this." Didi's substitution of [tɤ] or [tʰɤ] for the first syllable of the words *ʈʂe4li3* "here," *ʈʂuo1tsi0* "table" and *tsui3pa1* "mouth" could be taken to involve regular segmental processes (deaffrication, vowel change), but we see in the other columns that Didi does produce affricates on occasion; the imposition of his templatic syllable appears to reflect ongoing reliance on a well-practised routine in addition to possible deployment of a regular phonological process that he may no longer need. (Note that Didi also produces the syllable [tu] in 11% of the words he attempts, sometimes as a replacement for an entirely different target syllable, but [tɤ, tʰɤ] occur even more frequently.) Complementarily, Shi imposes an affricate-onset syllable on three words – but here a straightforward articulatory account is more plausible: The retroflex affricates are late-learned (J. Ma et al., 2022) and Shi provides no variants that suggest that she is able to produce them at this developmental point. Finally, Yiyi's templatic syllable has the form CVN, at a time when most of the children learning Mandarin are avoiding codas: Yiyi is the only one of the five Mandarin learners to attempt more than one target word with a coda and the only one to produce a target coda. Most of the targets on which she imposes the syllable share with it the onset /t/ if not the CVN structure; the only "adapted" form is her replacement of *pʰin1* "put together" by [tʰɛn1] (yet she also produces *pən3* "notebook" as [pəŋ2]). This is good evidence of a representational rather than an articulatory motive; the "hook" (or "trap") here appears to be the CVN shape of the target.

The evidence that these Mandarin-learning children make repeated use of particular syllables to substitute for target syllables that share neither their onset nor their rime supports the idea that they may be representing words in terms of individual syllables rather than – or in addition to – representing them as whole-word units; it also suggests a database of remembered syllables that the child may draw on when access to the target form is difficult. However, we lack information as to the frequency of these syllables in the input to which the children have been exposed or in their cumulative vocabulary (or output patterns) to date.

Discussion

This paper was designed to evaluate the fit of the whole-word model to early word production in Japanese and Mandarin. We have considered three types of evidence of holistic lexical representation in child data from each of these languages. First, we looked for instances of *word-internal dependencies*. For Japanese, we found not only considerable use of repetition (resulting in child forms with reduplication or consonant harmony) but also the selection of repeated patterns "hidden" within the target to create such a form. Each of these ways of dealing with variegation provides evidence that the whole word may serve as a representational unit for these Japanese learners, as each of them involves manipulation not of a particular segment, syllable or segmental sequence but of the word form as whole.

Second, we considered possible *effects of prominent aspects of the adult target forms on the child's form*. For Japanese, the few isolated cases of initial consonant omission, or unstable production, in apparent relation to the presence of medial geminates provided weak support for whole-word representation. In Mandarin we saw no evidence of any such effects.

Finally, we considered evidence of the *inclusion of non-target elements* in child forms in either language. For one Japanese child, Haruo, we saw production of a preferred word pattern or template, < CORONAL – NON-CORONAL >, in 20% of the words he attempted and with one use of metathesis to arrive at the output form; this can be considered the use of a whole-word pattern supported by the child's output lexicon as a whole. In addition, we identified non-target syllable substitutions in data from three of the seven Japanese children. When these were considered in more detail, one proved to be part of a whole-word pattern (Haruo) while another appeared to involve segmental aspects of the target form (Kenta, [di, de]), despite the absence from the target of the syllable as such. In one case (Kenji, [bi]) we found, arguably, weak evidence of the imposition of a whole (well-practised) syllable on a target whose segments were not as closely matched to it as is usual; a whole-word pattern is an equally plausible interpretation in that case, bearing in mind that a motoric practice effect might also have influenced the child's form.

For Mandarin, in contrast, we found that all five children make use of whole syllables to substitute for target syllables that share with them neither onset nor rime, although the children differed in the extent to which they made such substitutions. These instances of non-target syllable production by all five children, apparently drawing on a well-practised production routine in each case, suggest that these children may be representing words in terms of individual syllables rather than – or in addition to – representing them as whole-word units. This is unlike what has been reported for children learning English, Spanish or other languages for which such analyses have been carried out (see, Vihman, 2019).

Before further discussing these observations, we note that the generalizability of our findings is limited by the relatively small number of children (5 to 7 per language) and word items (24 to 67 per child) in the data. One consequence is that we cannot arrive at a fair assessment of the extent of variability within each language. For example, only one of the seven Japanese children (Haruo) exhibited what could be taken to be a word-size template; it is difficult to determine how anomalous this case is. Similarly, it is not immediately clear whether it is due to sampling or individual variation that we found no examples of reduplicated outputs imposed on variegated targets in two of the Mandarin children. Such variability may reflect differences in the linguistic input the children were exposed to. For example, parental speech to infants varies in the extent to which it contains register-specific lexical items, such as diminutivised and reduplicated forms (Berko Gleason et al., 1994; Ota et al., 2018), which may affect the global phonological profile of the child's lexical exposure. Whether this type of input variability can explain the attested individual differences in our analyses of Japanese and Mandarin data is a question that goes beyond the scope of the current study. Nevertheless, it is clear that our data robustly present overall differences between Japanese and Mandarin, and between these languages and previously examined languages (e.g, English, French and Finnish), in terms of the prevalence of phenomena that can be interpreted as evidence for whole-word phonology and syllable-size templates.

Whole-word phonology reconsidered

The whole-word phonology model rests primarily on the types of evidence mentioned above – word-internal dependencies, effects of prominence later in a word affecting production of the word-initial consonant and the presence of word-size templates. These departures from the target reflect aspects of child perception of, attention to and memory for the target form. We have argued further that those representational discrepancies are likely rooted in the limited child repertoire of motor routines onto which remembered words may be mapped. Thus both perceptual aspects of input speech and the vocal constraints of beginning talkers contribute to the holistic representations with which we have been primarily concerned.

In addition to effects of accentual prominence in input speech children may show more idiosyncratic effects in their treatment of adult targets when they deploy already established articulatory routines to produce the elements most familiar to them, not always sequencing segments in conformity with the adult model (i.e., showing metathesis or other restructuring of the elements of the target word). Such adaptation of adult targets to a consistent child pattern or template likely reflects articulatory limitations but also points to representational insecurity, suggesting that the children are implicitly groping for a representation that may be more or less complete and robust, depending in part, no doubt, on the frequency with which they have previously heard the word (Ota, 2013) or produced it themselves (Vihman, 1993, 2022), or both (see also, Aitchison & Chiat, 1981).

The idea that children reuse existing motoric routines once they have begun to expand their vocabulary, going from relatively accurate production of a small number of words to less accurate but more systematic production of many more words, is basic to the whole-word model. The idea that existing vocal patterns and lexical knowledge support new word learning has received experimental support from studies of nonword repetition in both older children (e.g., Cychosz et al., 2021; Dollaghan et al., 1995) and two-year-olds (Keren-Portnoy et al., 2010). Complementarily, Faytak (2018) has provided ultrasound imaging evidence from adults of the grounding of phonologically systematic production in phonetics, or more specifically, in the reuse of motoric patterns. Faytak succinctly summarizes the idea of "uniformity of speech articulation" as it relates to phonological development:

> Early speech production experience, most likely acquired using trial-and-error learning, aggressively generalizes to new words once word production becomes a major goal of the child learner, with learner-internal consistency often winning out over resemblance to the adult form. The child's idiosyncratic language-learning experience systematically affects uptake of lexical material from the environment: words that contain mastered articulatory patterns are learned at a greater rate than words that do not . . . That child language learners generalize a handful of successful motor routines in this fashion in the course of building a larger lexical (and, presumably, gestural) repertoire suggests that a "good-enough" regime dominates at this time period. (p. 32)

The "winning out" of "learner-internal consistency . . . over resemblance to the adult form" would be the outcome of what we call here "adaptation," while the more rapid learning of "words that contain mastered articulatory patterns" is "selection."

Whole-word phonology in relation to Japanese and Mandarin

We have provided various bits of evidence converging on the suggestion that children learning Japanese are representing whole word forms in the single-word period. This evidence – the frequent reliance on consonant harmony in producing variegated target words, the occurrence of metathesis at a distance in several forms, the occasional omission or replacement of word-initial consonants and the finding of at least one apparent template (the < CORONAL – NON-CORONAL > pattern) – had no parallel in Mandarin. Here, instead, we observed clear instances of children making idiosyncratic overuse of whole "templatic syllables" (Table 7); such imposition of whole non-target syllables has to our knowledge not previously been reported. We also note an absence of consonant harmony from our

Mandarin data, a characteristic previously reported for Cantonese (Vihman, 1978). And we see no "spilling over" of features from one part of the word to another and no signs of attention being shifted from one part of the word to another.

This is not to say that children learning Mandarin do not also draw on the adult word as a whole in their productions, as evidenced by the partial resemblance to the targets of many words with "templatic syllable" substitutions or, for example, the one instance we noted of possible metathesis. However, the primary type of discrepancy between adult and child word forms in our Mandarin data involves wholesale replacement of an entire syllable rather than features, sub-syllabic elements or, to the contrary, whole-word-sized patterns. This leads us to conclude that children learning Mandarin may focus their representations primarily on syllable-sized units rather than on the whole word. In short, we suggest that the two syllables of a disyllabic word may not make up a whole for Mandarin learners in the same sense as for children learning Japanese and the other languages we have referenced.

The status of syllables in Japanese and Mandarin

Given the finding that Japanese and Mandarin learners targeted relatively fewer variegated words (compared to children learning English, Finnish or French), yet managed to produce proportionately more of them with variegation, Vihman et al. (2022) suggested that these languages, with their simpler phonotactic structure, might present a lesser challenge to emergent representational capacity: Given fewer syllables to learn, children might more easily be able to retain different sequences, within their individual articulatory competence. If this were the case, it should mean higher syllable reuse in Japanese and Mandarin than in English, Finnish and French. To test the idea Vihman et al. tallied the number of unique syllables in the target words used by each group. They found that children learning either of those languages need access only about 10 different or "unique" syllables to produce 10 different disyllabic words (9.88 for Mandarin, 10.74 for Japanese); in contrast, children learning any of the European languages sampled in that study must recruit or choose between 12.23 and 15.53 unique syllables per 10 disyllabic words, a somewhat greater representational challenge.

In other words, children acquiring either Japanese or Mandarin may be able to gain familiarity with individual syllables relatively quickly and become sensitive to their repeated use in diverse target words. However, for children learning Japanese we have uncovered evidence of whole-word representation. In contrast, Mandarin-learning children's response to the smaller syllable inventory of Mandarin may lead them to represent words using syllable-sized units. Vihman et al. (2022) also found that Mandarin, although not Japanese, learners made greater use of reduplicated forms than did the children learning English, Finnish or French (see also, Figure 1, above). One might argue that reduplicated words serve as a training ground for word production, creating for each Mandarin-learning child a sufficient database of highly familiar syllables to support the mapping of new words onto emergent lexical representations. This would facilitate production by permitting the child to retain new words by mapping them onto existing representations syllable-by-syllable rather than as prosodic wholes; the suggestion is in line with the evidence that Mandarin learners specifically draw on their experience with syllable production as their lexicon expands.

Thus the phenomenon of a child repeatedly imposing non-target syllables on the word forms he or she produces may be related to the status of syllables in the adult language; the fact that, in Mandarin, individual syllables typically occur as meaningful words in their own right as well as combining to create independent disyllabic forms may also play a role. At the same time, the presence of tonal marking on each syllable of a Mandarin word and the concomitant absence of a "rhythmic envelope" is also likely to be a factor behind the absence of whole-word effects.

Interestingly, in a recent experimental study of Tip-of-the-tongue (TOT) priming with Mandarin-speaking adults Chang et al. (2022) found that priming with the first syllable of a four-syllable idiom had only a weak effect on prompting memory for the idiom in Mandarin, in contrast to English, where such priming has been found to be effective in resolving TOT states. Furthermore,

in cases where the TOT failed to be successfully resolved, syllables other than the first were more likely to be recalled. The authors mention two factors that might account for the difference. First, due to the limited number of Mandarin syllables (especially if tone is disregarded: It was found not to matter whether the tone of the priming stimulus matched the to-be-remembered syllable or not), all syllables are of relatively high frequency, contrary to English, where only lower frequency syllables have been found to serve as effective primes in such experiments. Secondly, the fact that syllables are "individually meaningful" might detract from the value, for priming, of hearing the first one. In this psycholinguistic study, then, we see that the distinct status of the syllable in Mandarin affects adult processing. In the case of children, we presume that only a small subset of the 400 syllables available will have become familiar by the developmental point of interest here, but it is plausible that the child's existing database of syllables serves them better as a means of accessing the words they know in Mandarin than can be the case in English, for example, with the far larger database of syllables available to a child.

It should be noted that the phenomenon of recurrent use of idiosyncratic syllables is not unknown in English. Smith (1973) reports that, from about age 3;4, his son frequently replaced pretonic phonological material with the syllable [ri] e.g., *attack* [riːˈtæk], in a process he considered a type of "grammatical simplification" (pp. 171ff.). Gnanadesikan (2004) uses the term "dummy syllable" for her daughter's similar replacement of (whole) pretonic syllables by [fi]: *umbrella* [fi-bɛyə],*potato* [fi-tero] and even, with movement of the target onset to the stressed syllable, *balloon* [fi-bun], *koala* [fi-kɑlɑ] and other words with liquid onsets in the second syllable. However, these cases seem to be qualitatively different from the pattern found in Mandarin in that the syllable with fixed segmental content predictably occurs in prosodically weak position in English. In that sense, these cases are more akin to the use of reduplication to replace prosodically weak syllables reported from Catalan and Estonian, as discussed above.

We tentatively conclude that the critical unit for early phonological and lexical learning in Mandarin – and potentially in other languages of similar phonological structure, such as Burmese or Thai, for example – may be the syllable rather than the word. This would make it easier to represent (retain and access) target words as a combination of familiar units, rather than as a single unit involving a complex and unpredictable sequence of articulatory changes. This means that while the whole-word model can provide useful insight into some languages, it may be less well adapted to a language like Mandarin, in which representation of the syllable may take precedence over the word. Further study of early phonological development in languages of a range of different types is needed to test this conclusion.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Marilyn May Vihman http://orcid.org/0000-0001-8912-4840
Mitsuhiko Ota http://orcid.org/0000-0003-3693-8490
Tamar Keren-Portnoy http://orcid.org/0000-0002-7258-2404
Rui Qi Choo http://orcid.org/0000-0002-2804-0237

References

Aitchison, J., & Chiat, S. (1981). Natural phonology or natural memory? The interaction between phonological processes and recall mechanisms. *Language and Speech*, *24*(4), 311–326. https://doi.org/10.1177/002383098102400402

Allen, G. D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology, 1: Production* (pp. 227–256). Academic Press.

Beckman, M. E. (1986). *Stress and non-stress accent*. Foris publications.

Berko Gleason, J., Perlmann, R. Y., Ely, R., & Evans, D. W. (1994). The baby talk register: Parents' use of diminutives. In J. L. Sokolov & C. E. Snow (Eds.), *Handbook of research in language using CHILDES* (pp. 50–76). Erlbaum.

Chang, K. L., Hu, P., & Abrams, L. (2022). The tip-of-the-Mandarin tongue: Phonological and orthographic priming of TOT resolution in Mandarin speakers. *Language, Cognition and Neuroscience*, *37*(7), 925–938. https://doi.org/10.1080/23273798.2022.2033803

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

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*(1), 31–54. https://doi.org/10.1017/S0305000920000094

de Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, *67*(2), 297–319. https://doi.org/10.1353/lan.1991.0045

Demuth, K. (1995). Markedness and the development of prosodic structure. In J. N. Beckman (Ed.), *NELS 25: Proceedings of the North East Linguistics Society* (pp. 13–25). Amherst, MA: GLSA (Graduate Linguistic Student Association).

Deng, L., & Dang, J. (2007). Speech analysis: The Production-Perception perspective. In H.-Z. Li & C.-H. Lee (Eds.), *Advances in Chinese spoken language processing* (pp. 3–32). World Scientific.

Dollaghan, C. A., Biber, M. E., & Campbell, T. F. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics*, *16*(2), 211–222. https://doi.org/10.1017/S0142716400007098

Echols, C., & Newport, E. (1992). The role of stress and position in determining first words. *Language Acquisition*, *2*(3), 189–220. https://doi.org/10.1207/s15327817la0203_1

Farrell, M. T., & Abrams, L. (2011). Tip-of-the-tongue states reveal age differences in the syllable frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 277–285. https://doi.org/10.1037/a0021328

Faytak, M. D. (2018). *Articulatory uniformity through articulatory reuse: Insights from an ultrasound study of Sūzhōu Chinese. Unpublished* [PhD thesis], University of California.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*(2), 419–439. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 93-132. https://doi.org/10.2307/412864

Fikkert, P. (1994). *On the acquisition of prosodic structure*. Holland Institute of Generative Linguistics.

Fikkert, P., & de Hoop, H. (2009). Language acquisition in optimality theory. *Journal of Linguistics*, *47*, 311–358.

Fikkert, P., & Levelt, C. (2008). How does Place fall into place? The lexicon and emergent constraints in children's developing grammars. In P. Avery, E. Dresher, & K. Rice (Eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 231–270). Mouton.

Gnanadesikan, A. (2004). Markedness and faithfulness constraints in child phonology. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in phonological acquisition* (pp. 73–108). CUP.

Goad, H. (1997). Consonant harmony in child language: An optimality-theoretic account. In S. J. sh & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 113–142). John Benjamins.

Gonzalez-Gomez, N., Hayashi, A., Tsuji, S., Mazuka, R., & Nazzi, T. (2014). The role of the input on the development of the LC bias: A crosslinguistic comparison. *Cognition*, *132*(3), 301–311. https://doi.org/10.1016/j.cognition.2014.04.004

Grunwell, P. (1982). *Clinical phonozlogy*. Croom Helm.

Hallé, P., & de Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development*, *17*(2), 119–129. https://doi.org/10.1016/0163-6383(94)90047-7

Hallé, P., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, *19*(4), 463–481. https://doi.org/10.1016/S0163-6383(96)90007-7

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

Ingram, D. (1976). *Phonological disability in children*. Elsevier Press.

Keren-Portnoy, T., & Segal, O. (2016). Phonological development in Israeli Hebrew-learning infants and toddlers: Perception and production. In R. Berman (Ed.), *Acquisition and development of Hebrew: From infancy to adolescence* (pp. 69–94). John Benjamins.

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.

Khattab, G., & Al-Tamimi, J. (2013). Early phonological patterns in Lebanese Arabic. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole word approaches, cross-linguistic evidence* (pp. 374–414). Cambridge University Press.

Lleó, C. (1990). Homonymy and reduplication: On the extended availability of two strategies in phonological acquisition. *Journal of Child Language*, *17*(2), 267–278. https://doi.org/10.1017/S0305000900013763

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.

Macken, M. A. (1979). Developmental reorganization of phonology: A hierarchy of basic units of acquisition. *Lingua*, *49* (1), 11–49. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 133–167. https://doi.org/10.1016/0024-3841(79)90073-1

Macken, M. A. (1980). Aspects of the acquisition of stop systems: A cross-linguistic perspective. In G. Yeni-komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, I: Production* (pp. 143–192). Academic Press.
Ma, L., Perrier, P., & Deng, J. (2015). Strength of syllabic influences on articulation in Mandarin Chinese and French: Insights from a motor control approach. *Journal of Phonetics*, *53*, 101–124. https://doi.org/10.1016/j.wocn.2015.09.005
Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. In U. H. Frauenfelder & L. K.Tyler (Eds.), *Spoken word recognition* (pp. 71–102). Elsevier.
Ma, J., Wu, Y., Zhu, J., & Chen, X. (2022). The phonological development of Mandarin voiceless affricates in three- to five-year-old children. *Frontiers in Psychology*, *13*, 809722. https://doi.org/10.3389/fpsyg.2022.809722
McAllister Byun, T., Inkelas, S., & Rose, Y. (2016). The A-map model: Articulatory reliability in child-specific phonology. *Language*, *92*(1), 141–178. https://doi.org/10.1353/lan.2016.0000
Menn, L. (1971). Phonotactic rules in beginning speech: A study in the development of English discourse. *Lingua*, *26*(3), 225–251. https://doi.org/10.1016/0024-3841(71)90011-8
Menn, L. (1978). Phonological units in beginning speech. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments* (pp. 157–171). North-Holland.
Menn, L. (1983). Development of articulatory, phonetic, and phonological capabilities. In B. Butterworth (Ed.), *Language Production* (Vol. 2, pp. 3–50). Academic Press. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 168-214.
Mok, P. K. P. (2010). Language-specific realizations of syllable structure and vowel-to-vowel coarticulation. *Journal of the Acoustical Society of America*, *128*(3), 1346–1356. https://doi.org/10.1121/1.3466859
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*. 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). CUP.
Ota, M., Davies-Jenkins, & Skarabela, B. (2018). Why choo-choo is better than train: The role of register-specific words in early vocabulary growth. *Cognitive Science*, *42*, 1974–1999. https://doi.org/10.1111/cogs.12628
Pater, J. (1997). Minimal violation and phonological development. *Language Acquisition*, *6*(3), 201–253. https://doi.org/10.1207/s15327817la0603_2
Pater, J. (2004). Bridging the gap between receptive and productive development with minimally violable constraints. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in phonological acquisition* (pp. 219–244). Cambridge University Press.
Pater, J., & Werle, P. (2003). Direction of assimilation in child consonant harmony. *Canadian Journal of Linguistics/ Revue Canadienne de Linguistique*, *48*(3–4), 385–408. https://doi.org/10.1017/S0008413100000712
Priestly, T. M. S. (1977). One idiosyncratic strategy in the acquisition of phonology. *Journal of Child Language*, *4*(1), 45–66. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 217-237. https://doi.org/10.1017/S0305000900000477
Rose, Y. (2000). *Headedness and prosodic licensing in the L1 acquisition of phonology* [Unpublished PhD dissertation]. McGill University.
Savinainen-Makkonen, T. (2007). Geminate template: A model for first Finnish words. *First Language*, *27*(4), 347–359. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 362-373. https://doi.org/10.1177/0142723707081728
Segal, O., Keren-Portnoy, T., & Vihman, M. (2020). Robust effects of stress on early lexical representation. *Infancy*, *25*(4), 500–521. https://doi.org/10.1111/infa.12340
Smith, N. V. (1973). *The acquisition of phonology: A case study*. The University Press.
Snow, D. (1998). A prominence account of syllable reduction in early speech development: The child's prosodic phonology of tiger and giraffe. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1171–1184. https://doi.org/10.1044/jslhr.4105.1171
Spencer, A. (1986). Towards a theory of phonological development. *Lingua*, *68*(1), 3–38. https://doi.org/10.1016/0024-3841(86)90021-5
Stampe, D. (1979). *A dissertation on natural phonology*. Garland.
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). Singular Publishing Group.
Vihman, M. M. (1978). Consonant harmony: Its scope and function in child language. In J. Greenberg, C. A. Fergu-son, & E. A. Moravcsik (Eds.), *Universals of human language* (Vol. 2, pp. 281–334). Stanford University Press.
Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, *21*(1–2), 61–82. https://doi.org/10.1016/S0095-4470(19)31321-X
Vihman, M. M. (2016). Prosodic structures and templates in bilingual phonological development. *Bilingualism: Language and Cognition*, *19*(1), 69–88. https://doi.org/10.1017/S1366728914000790
Vihman, M. M. (2019). *Phonological templates in development*. Oxford University Press.
Vihman, M. M. (2022). The developmental origins of phonological memory. *Psychological Review*. https://doi.org/10.1037/rev0000354
Vihman, M. M., & Croft, W. (2007). Phonological development: Toward a 'radical' templatic phonology. *Linguistics*, *45* (4), 683–725. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 17-57. https://doi.org/10.1515/LING.2007.021

Vihman, M. M., & Greenlee, M. (1987). Individual differences in phonological development: Ages one and three years. *Journal of Speech, Language, and Hearing Research*, *30*(4), 503–521. https://doi.org/10.1044/jshr.3004.503

Vihman, M. M., & Keren-Portnoy, T. (Eds.). (2013a). *The Emergence of Phonology: Whole word approaches, cross-linguistic evidence*. Cambridge: Cambridge University Press.

Vihman, M. M., & Keren-Portnoy, T. (2013b). Introduction. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole word approaches, cross-linguistic evidence* (pp. 1–14). Cambridge University Press.

Vihman, M. M., & Kunnari, S. (2006). The sources of phonological knowledge: A cross-linguistic perspective. *Recherches Linguistiques de Vincennes*, *35*(35), 133–164. https://doi.org/10.4000/rlv.1467

Vihman, M. M., & Majorano, M. (2017). The role of geminates in infants' early words and word-form recognition. *Journal of Child Language*, *44*(1), 158–184. https://doi.org/10.1017/S0305000915000793

Vihman, M. M., & McCune, L. (1994). When is a word a word*? Journal of Child Language*, *21*(3), 517–542. https://doi.org/10.1017/S0305000900009442

Vihman, M. M., & Miller, R. (1988). Words and babble at the threshold of language acquisition. In M. D. Smith & J. L. Locke (Eds.), *The Emergent Lexicon* (pp. 151–183). Academic Press.

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*(3), 336–353. https://doi.org/10.1016/j.jml.2003.11.004

Vihman, M. M., Ota, M., Keren-Portnoy, T., Lou, S., & Choo, R. Q. (2022). Child phonological responses to variegation in adult words: A cross-linguistic study. *Journal of Child Language*, 1–28. https://doi.org/10.1017/S0305000922000393

Waterson, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics*, *7*(2), 179–211. Reprinted in Vihman & Keren-Portnoy, 2013a, pp. 61-92. https://doi.org/10.1017/S0022226700002917

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). CUP.