Influence of geminate structure on early Arabic templatic patterns

Ghada Khattab, Jalal Al-Tamimi

Influence of geminate structure on early Arabic templatic patterns

Ghada Khattab and Jalal Al-Tamimi
Newcastle University

This paper reports on the early development of phonology during the one-word stage in five Lebanese children, paying particular attention to the influence of the adult phonology as well as to the children's individual journeys towards adult-like patterns. The study contributes to two of the main aims of this volume: First, it shows that early word shapes by Lebanese-speaking children do not follow a straightforward developmental track from simple to complex structures; rather, individual preferences in early productions and the frequency or prominence of particular structures in the adult phonology play a major role in shaping the phonological structure of words in the second year of life. Second, the study sheds more light on the so-called U-shaped curve in development whereby children may have accurate forms in their production at an early stage of development but later 'regress'; in this study this is scrutinised from the point of view of the acquisition of phonological length in consonants and it is suggested that accurate forms before and after a 'regression' stage may be qualitatively different, with only the latter showing 'real' acquisition of adult-like phonological structure.

14.1 The emergence of phonology and the role of cross-linguistic differences

As shown by various contributions to this volume, children build their phonological knowledge from an initially small repertoire of words that may occur frequently in their input, attract their attention and contain sounds that are part of their babbling and early word practice; subsequently, their attempts at producing these words gain the attention of caregivers who potentially repeat the words to the children. The phonological structure of these words may influence the children's subsequent selection of adult targets, as well as lead to adaptation of phonologically distant targets to that same structure, resulting in productive template use. The structure of words in the child's own first lexicon, together with segmental and phonological patterns of the adult language, are jointly responsible for the shape of the templates and for individual differences in children's templatic shapes (Vihman and Croft 2007: 707). Below we unpick some of these seminal ideas and look at cross-linguistic effects on children's early words.

In a whole-word account, the child may group phonetically-related words together and acquire word shapes or word patterns as the basic units (Ferguson and Farewell 1975; Menn 1983; Macken 1979). Children are typically highly variable in their word production in the early stages, suggesting a lack of command over individual sounds within these words and/or a lack of abstract categorical knowledge of the sounds within them (Vihman and Croft 2007: 689). Adaptations of adult words to the child's preferred templatic shapes reveals the relationship between groups of words in the child's lexicon and offers a window into the way children deal with challenges with respect to particular sounds or sound sequences. Evidence for the emergence and development of templatic behaviour in a child's lexicon includes: a) consistency of patterning in several of the child's words produced over several sessions, b) occurrence of unusual phonological correspondences between adult and child forms due to the influence of the template and c) a sharp increase of words that fit the template (Vihman and Croft 2007: 694-5). More phoneme-like categorisation may appear, when re-organisation of word shapes and units takes place, though the child may still use some of the preferred sounds from their early prosodic units (Macken 1979: 34).

While early words may be similar cross-linguistically, the phonology of each adult language – that is, the ambient language shapes – influence the first phonological templates that

emerge out of these shapes and that start to be applied to new words which are beyond the child's range (Vihman and Croft 2007: 692). For instance, examples of English-speaking children's templates include monosyllables with final nasals ([CVN]) or trochaic disyllables with child-specific consonant or vowel components, e.g. [C_1VC_2V], [CVjV], or [$CV_{low}CV_{high}$] (Macken 1979; Priestly 1977; Vihman et al 1994). French-speaking children's templates, on the other hand, tend to follow a language-specific prosodic shape, [$_{l}\sigma\ (\sigma)_{n}{}'\sigma$], with a final stressed syllable, a counter-stress on the initial syllable, and up to two optional syllables in between (Wauquier & Yamaguchi this volume), e.g. [a-o] template as in [ato] for /bato/ 'ship'; [afo] for /elefɑ̃/ 'elephant'; and [abalɔ̃] for /ɛ̃ balɔ̃/ 'one balloon'. In Estonian, Vihman and Vihman (2011) find a C_0Vi/jV template (where C_0 represents an optional C), with a palatal medial glide that is more consistent than the initial C, which is often omitted. Note that while medial glides have been reported to occur as part of English templates as well (Priestly 1977), the prosody of each language influences the way other segments in the templatic structure are realised, as exemplified by the initial consonant omission in Estonian but not English. The prominence of medial position in Estonian has been discussed in studies on languages with a quantitative length distinction, where gemination may further attract the child's attention to medial consonants at the expense of initial ones. These studies are discussed next.

14.2 The role of geminate structure in shaping early words

Waterson (1971: 181) suggests that in the early stages of production, the child may produce only those features of the adult target that they can perceive and easily reproduce. Long consonants fall into the category of sounds which must at the same time be salient in the input, due to their prominent duration (alongside non-durational cues, e.g. Al-Tamimi & Khattab 2011; forthcoming), and relatively easy to produce, since children's early articulations are slow (e.g. Stoel-Gammon & Cooper 1984). While the child's long phonetic durations in the early stages of production do not necessarily translate into contrastive acquisition of segmental length, that early practice must provide a stepping stone for later internalisation of length as a phonological feature. Whereas for languages like English the starting point for the child's production pattern is often considered to be CV(CV), a phenomenon referred to as the 'core syllable' stage (Demuth 1995; Fee 1996; Fikkert 1994), children acquiring languages with quantitative medial contrasts have been shown to exhibit different early patterns. For instance, while English has a dominant trochaic pattern with a louder, higher-pitched first syllable, Finnish (which is also consistently trochaic) has many medial geminates which may be inherently salient for children, as mentioned above, and which may also attract their attention due to their frequency in child-directed speech, as can be seen from the relatively high number of medial geminates the children aim for and produce, regardless of target (Vihman and Velleman 2000).

The prominence of the geminate structure in the language has led some researchers to suggest that CVC:V, rather than CV(CV), is the starting point for Finnish children. For example, Savinainen-Makkonen (2007: 346) looks at data from a Finnish child, Joel, between the ages of 1 and 1;6 and finds that the majority of his utterances (47 out of his first 50 words) have a disyllabic structure. Furthermore, it is medial gemination rather than stress that seems to govern what is deleted and what is retained in Finnish children's production, who tend to omit initial consonants in trochaic shapes while showing more accurate production of medial consonants. Similar results are reported by Vihman and Velleman (2000), who were surprised to find that the second most common pattern in their Finnish data (after consonant harmony) was onset deletion (31%, both selected and adapted), a pattern considered to be a sign of

deviant phonology in English. Similar patterns have also been found for a child acquiring Hindi (Bhaya Nair 1991), where deletion of onsets is present in many disyllabic Hindi words with medial clusters or geminates.

Finnish and Arabic share common phonological patterns in the adult language, including phonemic consonant and vowel length and rich morphophonology, leading to multisyllabic words being frequent in the input due to the addition of various suffixes to stems. One notable difference relates to the status of initial consonants, with the phonology of Arabic disallowing onsetless syllables (Watson 2002: 56). We were therefore interested to determine whether children acquiring Arabic show similar patterns to children acquiring Estonian, Finnish, Hindi, and other languages with gemination. What we found surprising was that, within the scarce literature on phonological development in Arabic, the acquisition of gemination has not been dealt with in any detail. In the next two sections we present an overview of relevant aspects of Arabic phonology before exploring findings from cross-linguistic acquisition studies which challenge the Anglo-centric claims about the salience of initial consonants and the typical patterns of acquisition of syllable structures.

14.3 Gemination and other relevant characteristics of Arabic phonology

Arabic has a complex root-and-pattern (or non-concatenative) morphology (Watson 2002; McCarthy & Prince 1990ab). In Arabic linguistics this system is coincidentally also referred to as 'templatic' (although this usage should not be confused with the terminology used in this volume to refer to developmental processes). The stem of a content word in Arabic has three discontinuous morphemes:

a) the consonantal root (e.g. k, t, b), which is the underlying lexical unit of the language that conveys semantic information (in the example here k, t, b relates to 'writing');
b) the templatic pattern into which the consonantal root is inserted, adding morpho-syntactic and phonological information to the root (e.g. the word pattern referred to by Arabic linguists as 'faʔal' expresses the past tense, whereby /f/, /ʔ/, and /l/ are placeholders for each of the consonants in the root, thus /ˈkatab/ 'he wrote');
c) the interpolated vowels, which signal changes in voice (active or passive in verbs), agent relations in nouns derived from verbs, and singular-plural relations in nouns (e.g. /ˈkutib/ 'it was written'; /ˈkaatib/ 'writer'; /ˈkutub/ 'books').

Whether Arabic speakers acquire and store whole stems or individual components of their lexicon (roots, templates, and melodies) is matter of debate (e.g. Boudelaa & Marslen-Wilson 2001; 2004; Ravid 2002) and is beyond the scope of this study. What is of interest here is the wide range of resulting word shapes that Arabic-speaking children are exposed to, many with a final coda. Although a lot more work is needed on deriving frequency of occurrence of various templatic shapes in both the adult lexicon and child-directed speech, the three most commonly occurring shapes tend to be: 1) CVCV(C) for nouns (e.g. /ˈʒabal/ 'mountain'; /ˈdawa/ 'medicine') or form I present perfect 'faʔal' verbs (e.g. /ˈdaras/ 'he studied'; /ˈkatab/ 'he wrote' etc.), 2) CVC:V(C) for form II causative 'faʕʕal' verbs (e.g. /ˈdarras/ 'he taught'; /ˈkattab/ 'he made someone write' etc.) or nouns (e.g. /ˈbatˤtˤa/ 'duck'), and 3) CV:CV(C) for nouns and form III active participles or nouns (e.g. /ˈwaadi/ 'valley'; /ˈsaadid/ 'having blocked'). Out of the ten triliteral verb templates in Arabic, form II with the geminate consonant is the most productive and the most common in modern Arabic dialects (Watson

2002: 134). Medial gemination is also used in the derivation of nouns of profession from form II verbs, resulting in an iambic CVC:ˈV:C shape, e.g. /xabˈba:z/ ‘baker’.

In terms of syllable structure, disyllables are much more common than monosyllables, with nine out of the ten triliteral verb forms having a disyllabic structure and the majority of nouns having disyllabic or trisyllabic structure (Watson 2002; 134-165). The majority of disyllabic verbs have a trochaic stress pattern, while nouns can be iambic or trochaic. Arabic is also a quantity-sensitive language, with the mora playing an important role in syllable weight (McCarthy & Prince 1986; Hayes 1989). The minimal word is thought to be bimoraic, i.e. either consisting of a monosyllabic word with two vowels or a coda (CV: or CVC), or a disyllabic CVCV word (Broselow 1992; McCarthy & Prince 1986; 1990ab). Syllable types in Lebanese Arabic include: CV (in non-final position, e.g. /ˈʔalam/ ‘pen’); CVC (e.g. /sin/ ‘tooth’); CV: (e.g. /laa/ ‘no’); CV:C (e.g. /be:b/ ‘door’); CVCC (e.g. /nahr/ ‘river’); CV:CC (e.g. /ʕaamm/ ‘public’) (Khattab 2007; Nasr 1960; 1966; Obrecht 1968). CV is light and does not occur in monosyllabic words; CVC and CV: are heavy; CV:C, CVCC, and CV:CC are superheavy. Gemination and vowel length are two main characteristics of syllable structure, and their weight is unaffected by syllable position. Each vowel or geminate consonant has one mora, while singleton consonants acquire weight by position, with onset consonants being weightless and final consonants extra-metrical (Watson 2002: 54).

At the segmental level, all Arabic consonants can be geminated, which for Lebanese Arabic means at least 28 consonants (Table 14.1). Many disyllabic French loan words are also pronounced with a long medial consonant in the French accent in Lebanon (e.g. *tape* /ˈtapə/ [ˈtappø] ‘clap’; *papa* /paˈpa/ [papˈpa] ‘daddy’, etc.), contributing to the high frequency of words with long medial consonants in the adult phonology. Vowel length is also contrastive, with the following impressionistic set for LA (there are no experimental studies of LA vowels): /iː, ɪ, eː, e, æː, æ, ɑː, ɑ, uː, ʊ, oː, o, æɪ, æʊ/. Geminate consonants are about twice as long as their singleton counterparts, and the same applies to phonologically long vowels in comparison to short ones (Khattab 2007; Khattab & Al-Tamimi 2008; forthcoming). Non-durational cues also play a secondary role in the singleton-geminate contrast (Al-Tamimi & Khattab 2011; forthcoming).

Table 14.1: Consonant inventory of Lebanese Arabic (adapted from Khattab 2007)

	Bilabial	**Labio-dental**	**Dental-alveolar**	**Post-alveolar**	**Palatal**	**Velar**	**Uvular**	**Pharyngeal**	**Glottal**
Plosive	(p) b		t d tˤ dˤ			k (g)	(q)		ʔ
Nasal	m		n						
Trill			r						
Tap			ɾ						
Fricative		f (v)	s z sˤ zˤ	ʃ ʒ		x ɣ		ħ ʕ	h
Approximant (+ lat. app)	w (lab-vel.)		l lˤ		j				

Note: Three of the sounds in brackets occur only in loan words (/p/, /v/, and /g/), while /q/ is normally realised as [ʔ] in most Lebanese dialects but retained as [q] by the Druze community and in the Standard variety.

14.4 Studies on the acquisition of Arabic

Given that exposure to frequent prosodic structures in a language may explain earlier acquisition of these structures in that language, the properties of Arabic prosodic structure described above suggest the following predictions: Arabic-speaking children may: a) produce disyllables early in the acquisition process; b) show coda production early; and c) acquire gemination and complex syllables early. These patterns can indeed be found in the data from studies on the acquisition of Arabic phonology, but they are seldom highlighted or discussed in any detail, perhaps because it is difficult to reconcile these results with the often assumed universal sequence of syllable structure acquisition. Moreover, most studies on phonological acquisition in Arabic, whether large-scale cross-sectional or small and longitudinal, have mostly looked at the order of acquisition of consonants and the phonological processes exhibited by Arabic-speaking children (e.g. Amayreh, & Dyson 1998; Ammar & Morsi 2006; Dyson & Amayreh 2000; Saleh, Shoeib, Hegazi, & Ali 2007; Shahin 1995; 2003), though more recent studies have looked at syllable structure as well (e.g. Abdoh 2011; Ammar 2002; Salem 2000). Here we review relevant findings from some of these studies.

In two studies looking at the acquisition of Jordanian Arabic consonants by children aged 2;0 to 6;4 (across the two studies), Amayreh and Dyson (1998) and Dyson and Amayreh (2000) found that medial consonants are much more accurate than initial and final consonants, with no significant difference between initial and final position. The authors wondered whether this result was influenced by the stress pattern in the words they elicited (Amayreh and Dyson 1998: 651), but a look at the word list in their appendix shows a balanced number of iambic and trochaic stress patterns. In a parallel study on Egyptian children aged 1;0 to 2;6, using naturalistic data, Saleh et al (2007) surprisingly found final position the most accurate in terms of consonant realisation, followed by medial and lastly initial position, which showed the highest degree of errors in production (substitutions and deletions). This was echoed in a study on the acquisition of consonants in all word positions in 21 Palestinian children aged 1;4 to 2;10 by Shahin (2003), who notes that final codas were highly accurate (Shahin adopts a phonologically-driven explanation, suggesting that final codas are representationally onsets, see Harris & Gussman 1998). Out of all four word positions, initial, medial onsets, and final codas were deemed to be acquired early by the children, while medial codas were acquired late. While this is not explicitly discussed in the study, the cross-sectional data showed more accuracy for final consonants and medial onsets than initial word onsets, especially in the youngest age group (Shahin 2003: 9-17), and development followed a non-linear progression, with dips in accuracy at all ages and a lot of individual variation.

Abdoh's (2011) study is among the few Arabic acquisition studies focussing more on word shapes than segmental acquisition. The author looked at first words in 22 Hijazi-speaking children aged 1;0-1;9 within Prosodic and Moraic Theory approaches to phonological structure (e.g. McCarthy & Prince 1990; Hayes 1981). Despite the fact that her data do not fully support the presumed universal order of acquisition of word structure, Abdoh maintains that the children in her study follow that order, albeit with a starting point that skips the monomoraic core syllable stage. The children are said to start at the minimal word stage where the maximal word size is a single binary foot and their outputs display bimoraic forms (ages 1;1 to 1;6); at later stages (1;6-1;9) they are reported as going beyond this stage and producing forms showing disyllabic words with a trochaic (SW) or iambic (WS) foot, and more complex structures exceeding the maximal size, i.e. structures with two feet. However, looking at the children's most frequent word shapes in the early stages (Abdoh, 2011: 149-155), the data show that disyllables constituted 60.9% of the children's production, followed by monosyllables at 38.2% and then trisyllables at 0.9%. When these three word shapes are

combined, the frequency of word types produced is the following: CVCV (29.1%) > CVC > CVC:V > CV:CV > CV (10%). Note that coda production is present from an early age (e.g. /dub/ 'bear', /ba:b/ 'door', etc.), despite reported cases of coda deletion. Gemination is also reported to be acquired early, particularly in medial position (Abdoh 2011: 149). The author points out that one reason for this might be that medial geminates often appear in baby talk (e.g. /ˈdubba/ 'teddy bear', /ˈdadda/ 'grandma' etc.). More interestingly, children's truncation patterns seem to preserve final syllables regardless of stress, e.g. /fusˈta:n/ 'dress' realised as [ta:n]; but also /ˈʔarnab/ 'rabbit' realised as [nab] and /ˈsamaka/ 'fish' realised as [ka].

Similar results regarding the early acquisition of complex syllable structures were reported by Ammar (2002), whose study of syllable structure in the speech of 10 Egyptian children aged 2;0 to 3;0 found that 90% of the children had acquired all syllable types. Ammar also reports on final consonant deletion being accompanied by lengthening of the preceding vowel. Furthermore, although she and other authors note cluster reduction in all the children up to age 4, she notes that clusters in CVCC are acquired earlier by Egyptian children than by English speaking children (Ammar 1999).

In sum, the results from these studies highlight the influence of the adult phonology on Arabic children's early words in terms of the early acquisition of medial and final consonants, complex syllable structures, and the predominance of disyllables in early words. However, very little mention is made of the potential role of gemination in shaping Arabic children's early words and influencing their attention to non-initial word positions. Moreover, with most of the above studies being cross-sectional in design, very little attention has been paid to individual children's development of phonology from the earliest stages of production. The present study therefore aims to fill this gap.

14.5 Current study

The data presented here are part of a longitudinal study of 10 Lebanese children, five based in Beirut and five in London (only the Beirut data are presented here). The study was carried out as part of an investigation of the acquisition of gemination by Lebanese-speaking children exposed to Lebanese Arabic alone and in conjunction with English and/or French. The Beirut-based families were recruited from the Greater Beirut area, but no further control for dialect was imposed. The emphasis was on locating families who were mainly Arabic-speaking (the use of French and/or English alongside Arabic is very common in Lebanon). The children were primarily cared for by their mothers and none had started attending nursery in the first two years of life. The children were recorded once a month from around 9 months of age until their third birthday. The recordings used for this paper are for the sessions where the children were deemed to be at the 4-word-point (4wp, i.e., when they produced 4 different word types spontaneously in a session) and all subsequent sessions leading up to the 25-word-point (25wp, when the children produced 25 different word types spontaneously in a session and had around 50 words in their vocabulary). Their ages ranged between 1;1 and 1;6 at the 4wp and 1;9 and 2;2 at the 25wp. The number of months that elapsed between the two points ranged between 4 and 9 (Table 14.2).

Table 14.2: Overall data. Number of recording sessions required from the 4wp (session 1) to the 25wp (final reported session) for each of the children. The table shows each child's age at the 4wp and the number of word types and tokens (in brackets) produced in each session.

child name	age at start	session 1	session 2	session 3	session 4	session 5	session 6	session 7	session 8	session 9	total

Rama	**1;6.11**	5 (17)	17 (50)	15 (27)	10 (23)	35 (86)					**82 (203)**
Martin	**1;3.06**	11 (75)	9 (54)	16 (57)	19 (63)	22 (41)	29 (174)	46 (140)			**152 (604)**
Lina	**1;3.25**	7 (27)	9 (28)	2 (11)	10 (31)	15 (28)	19 (57)	19 (59)	64 (206)		**145 (447)**
Hiyam	**1;1.05**	4 (15)	12 (56)	13 (48)	17 (52)	12 (22)	47 (130)	45 (105)	55 (149)		**205 (577)**
Mohamed	**1;6.02**	5 (20)	8 (42)	19 (70)	25 (70)	19 (90)	21 (97)	16 (101)	48 (153)	89 (389)	**234 (1032)**
All	**mean age: 1;4**					**Total: 818 (2863)**					

14.5.1 Procedure

The children were recorded at home while engaged in 30-40-minute spontaneous interactions with their mothers, and occasionally with grandparents or older siblings. The mothers were instructed to engage in play sessions with the children as they normally would, using familiar toys, picture books and other household items, while at the same time trying to elicit words/utterances they knew the children were able to produce.

Recordings were made in mono, 16-bit, 44.1 KHz sampling rate, using an Edirol R9 solid-state recorder with high quality wireless Sennheiser UHF microphones, one worn by the mother and one hidden in a baby vest worn by the children. Simultaneous video recordings were also made using a Sanyo camcorder and both audio and video recordings were used for the word identification process, while phonetic transcription relied mainly on the audio. The files were transferred onto a computer and the child's utterances segmented, labelled and transcribed using narrow IPA transcription for all segmental material. Both Praat v.5.1.10 (Boersma and Weenink 2009) and PHON v.1.5.2 (Rose 2012) were used for processing the audiovisual files (Praat allowed easier segmentation and labelling of speech while PHON allowed transcription using both audio and video outputs).

The children's utterances were categorised as 'babbling' (vocalisations with no identifiable target or communicative function), 'words' (utterances with identifiable target, using Vihman & McCune's 1994 word identification procedure) or 'unidentifiable' (utterances that were either unintelligible or where a word target was suspected but could not be established even after going through the word-ID test). Sessions in which the children had 4 to 25 identifiable spontaneous words were included in the analyses. Imitations were also recorded and analysed separately to determine whether they showed different patterns.

As can be seen from Table 14.2, the children vary in how quickly they get to the 25wp, the fastest being Rama, who reached criterion within 5 months, and the slowest Mohamed, who took twice as long. Interestingly, age at the 4wp does not predict how quickly the children will accumulate a vocabulary of around 50 words, since at the 4wp Rama and Mohamed are coincidentally the same age and the oldest children in the group. Both had been followed from an early age (around 11 months), and the differences between them were obvious right away: Rama was voluble from the start, but her utterances in the early recordings mostly consisted of babbling and lengthy unanalysable jargon (often 'monologues') that neither her mother nor the fieldworker could identify as words. Her 4-word session at age 1;6 marked the beginning of a change in her vocal behaviour, as she became less vocal (mostly due to producing less jargon) but began producing utterances that had identifiable targets and were fairly accurate. This remained the trend up to and including the 25wp. Mohamed, on the other hand, was a much more cautious and quiet child at the beginning. His mother noted that his

speech was developing more slowly in comparison with that of his older brother. He was a lot less vocal than Rama in the sessions leading up to the 4wp and then had several sessions with no noticeable increase in vocabulary (based on the recordings from sessions 3 to 7 and on his mother's observations).

On average, the children's age at the 4-word point (mean 1;4) is older than what is sometimes reported for US English (Vihman, Ferguson & Elbert 1986; Vihman & McCune 1994); all of the children experience a spurt in their production at some stage around the 25wp, in terms of either the overall number of tokens (Martin, session 6; Lina, session 6) or both word types and tokens (Rama, session 5; Hiyam, session 6, Lina session 8, Mohamed, sessions 8-10). This tends to coincide with either the session identified as the 25wp or the session immediately before that.

14.5.2 General patterns

As expected, Arabic words constituted the majority of utterances at 65%, followed by English (18%) and then French (8%). Words which could belong to more than one language were labelled as 'multilingual' and constituted the remaining 9% of the data (Table 14.3 and Fig. 14.1). Note that our interest in categorising the utterances into the three languages here was driven by the need to examine the influence of the language of origin on the syllable and word structure of the utterances that the children heard and produced. While the majority of the utterances that were labelled English and French in this study had commonly used translation equivalents in Arabic, they were not necessarily code-switches on the part of the children; an account of code-switching behaviour would require a different type of discourse analysis in order to establish whether the utterances were part of the Arabic CDS that the children heard or genuine switches to French or English discourse by mother and/or child, which is beyond the scope of this study.

Table 14.3: Language of origin for the utterances targeted by the children

	Arabic	**English**	**French**	**Multilingual**	**Total**
Total	1864 (65 %)	514 (18%)	217 (8%)	268 (9%)	2863

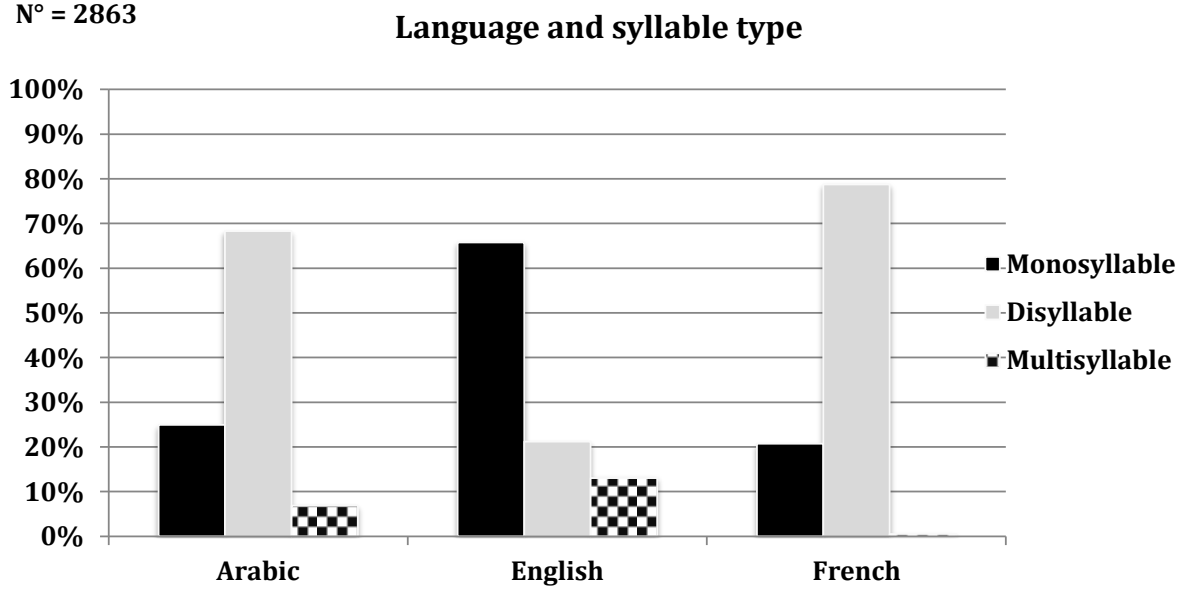

Figure 14.1: Distribution of target word shapes as a function of utterance language.

The distribution of early word shapes in Figure 14.1 reflects the differences in frequencies of mono- and disyllables in the three languages (e.g. Menn 1971; Rose & Wauquier-Gravelines 2007; Stoel-Gammon 1987), with the majority of early Arabic and French targeted words being disyllabic (68% and 79% respectively) while the majority of English words are monosyllabic (66%). Here, multisyllabic word frequency cannot be compared across the three languages because of the small numbers involved; as the children's productive abilities increased over the sessions, the emergence of multisyllabic words (with more than two syllables) was most prominent in Arabic, their dominant language. The difference in word shapes across the three languages was also reflected in the syllable structure within each word shape. For instance, within monosyllabic words targeted by the children, the most frequent syllable structure for Arabic words was CVC: (with a final geminate consonant), e.g. /baħħ/,[1] 'all gone', that for English words was CVC, e.g. *cat*, and for French words it was CV, e.g. *deux* /dø/ 'two' (Fig. 14.2). The same applies to disyllables (Fig. 14.3), with the most frequently targeted disyllabic shapes in Arabic being ˈCVC:V e.g. /ˈnanna/ 'food', ˈCV:CV e.g. /ˈbaːba/ 'daddy', and ˈCVCV e.g. /ˈtaʕa/ 'come here'; the most frequent targeted shapes for English were ˈCVVCV(C) e.g. *baby* /ˈbeɪbi/, followed by ˈCVCV, e.g. *teddy* /ˈtɛdi/ and ˈCCVCV e.g. *story* /ˈstoɹi/. French disyllable shapes showed a much more skewed pattern towards a single structure, which was CVˈCV (90%), e.g. *chapeau* /ʃaˈpo/ 'hat'.

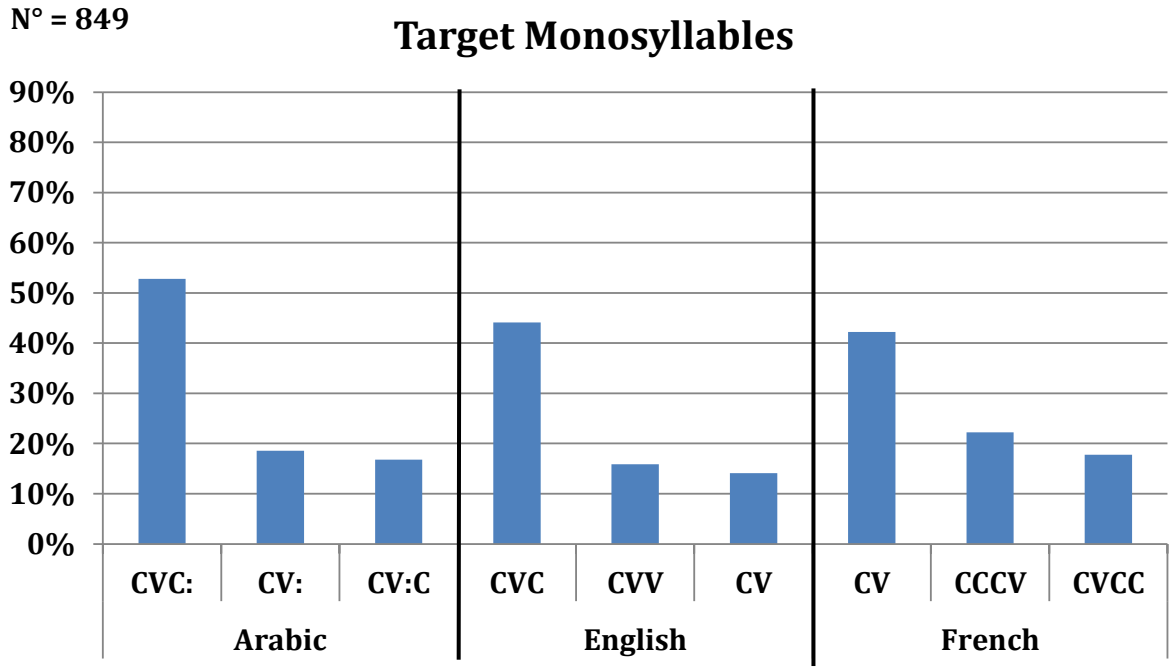

Figure 14.2: Distribution of the three most frequent syllables shapes for monosyllabic words targeted by the children in Arabic, English, and French utterances. Other, less frequently used shapes are not shown.

[1] Geminates are transcribed as double consonants in the IPA transcriptions throughout, but as 'C:' in syllable structure notation in order to separate them from consonant clusters, which are denoted as 'CC'. Long vowels are denoted as V: and diphthongs as VV.

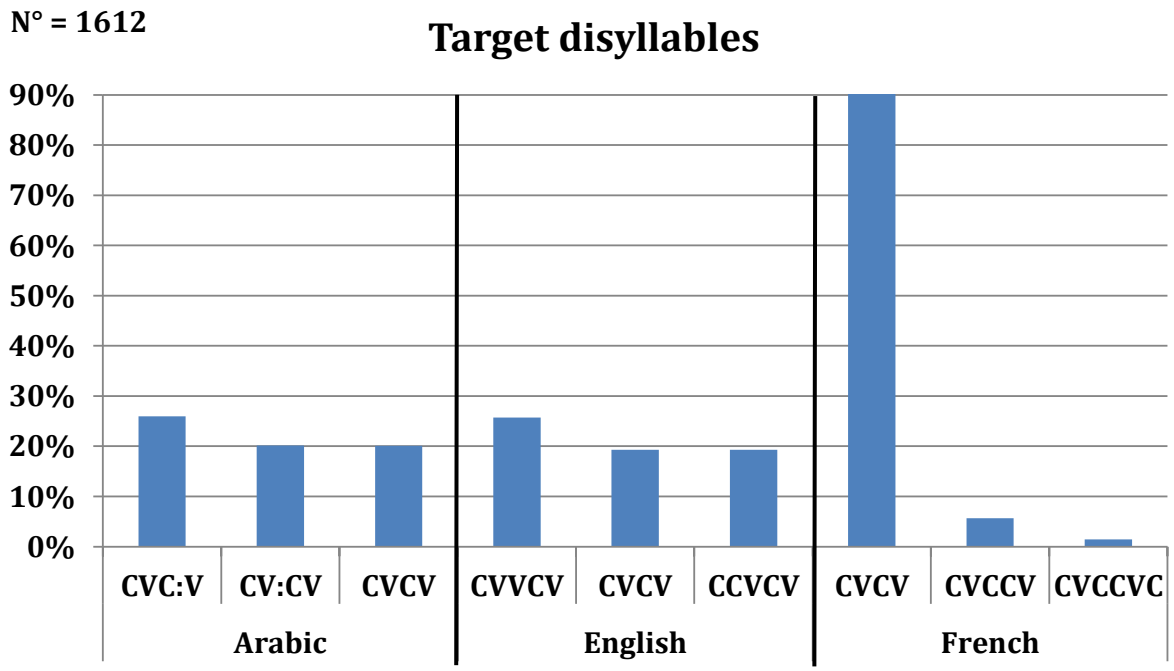

Figure 14.3: Distribution of the three most frequent syllables shapes for disyllabic words targeted by the children in Arabic, English, and French utterances. Other, less frequently used shapes are not shown.

On the whole, disyllables constitute a large part of the children's early word shapes (59%), which is expected, given the predominance of disyllables in Arabic and French. Moreover, due to the high frequency and salience of the medial geminate pattern in Arabic, all children both aim for disyllabic shapes with medial geminate or 'heavy' targets (with clusters or affricates) and adapt other shapes to the CVC:V(C) pattern. Baby words (/buwwa/ 'water'; /nanna/ 'food'; /bɪsse/ 'pussycat') and nicknames (e.g. /kitto/ 'Christopher'; /lillo/ 'Lina') also contributed to the high number of disyllabic words with medial geminates. Fig. 14.4 shows the distribution of disyllables that were targeted (left) by the children in terms of whether the medial consonant was a single consonant (e.g. /ˈʔana/ 'I'), a geminate (/ˈbaddo/ 'he wants') or complex (e.g. /mɪfˈteːħ/ 'key'), and how they were realised (Fig. 14.4, right); the complex category included clusters (e.g. /fʊtbɔl/ 'football') and secondary articulations (e.g. /sˤuːsˤ/ 'chick') in targets, but also affricated ([bobβõ] for French *balon* 'ball') and other doubly-articulated consonants ([ʔæ͡lnæʔ] for /ˈʔana/ 'I') in the realisations. Long C: and complex realisations by the children (54% and 30% respectively) are around 1.5 times as frequent as geminate and complex targets, suggesting that the medial consonants of many words with singleton targets were lengthened or produced with complex articulation.

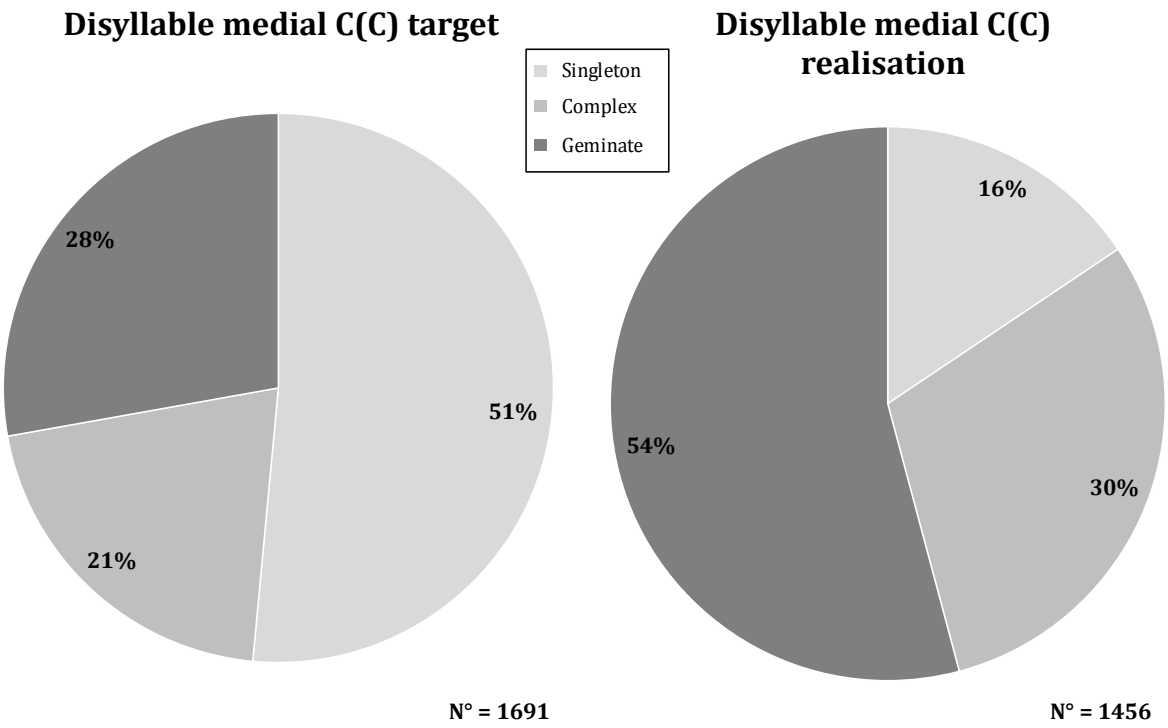

Figure 14.4: Distribution of medial consonant type (single, geminate, complex) in disyllabic words targeted by the children (left) and their realisation (right).

14.5.3 Developmental patterns

On the whole, the children target similar word structures in the early (4wp) and later (25wp) stages of production (Figs. 14.5-6), with a wider range of word shapes at the more advanced stage and an emergence of more complex shapes (not all listed in the figures below due to their very low frequency). One notable difference is a 14% drop in disyllabic CVC:V targets at the 25wp (Fig. 14.6), but *not* in realisations; in fact, lengthening of singleton consonants is still prominent and actually increases at the more advanced stage (Table 4, Figs. 14.9). The structure of the realisations for the most frequent target word shapes does not change very much as the children progress to the 25wp (Figs. 14.7-8); this is due to the fact that the children produce target-like structures from an early age, if phonological length is set aside. What they seem to take some time to acquire is phonological length, and their patterns of acquisition seem to involve experimenting with adding phonetic length to all elements of the target syllable structures rather than just to the phonologically long ones, or strengthening consonants (denoted as 'Cs' in the figures below). For instance, a target CVC: can be produced not just with a long coda, but also with a long onset and/or a long vowel, e.g. /baħħ/ 'all gone' realised as [baħħ], [bbaːħ], [baːħ], etc. Similarly, a target CVCV can be realised with varying lengths for all segments, e.g. /baːba/ 'daddy' realised as [babbah], [baːbbaħ], [bbaːbam], and [baːbaː], etc.; the realisations of disyllables with open final syllables in the target frequently contained a final coda, often a guttural sound (glottal stop, glottal fricative or pharyngeal fricative) but occasionally also other consonants with supraglottal places of articulation. While variable phonetic length may apply to all children's early productions regardless of their native language, the fact that Arabic has phonological vowel and consonant length may increase the salience of contrastive duration for the children, leading to their extensive experimentation with segment length and the production of syllables with

heavy rhymes and/or codas. Acoustic analysis is currently under way in order to obtain a clearer picture of the relationship between phonetic and phonological length in the children's productions.

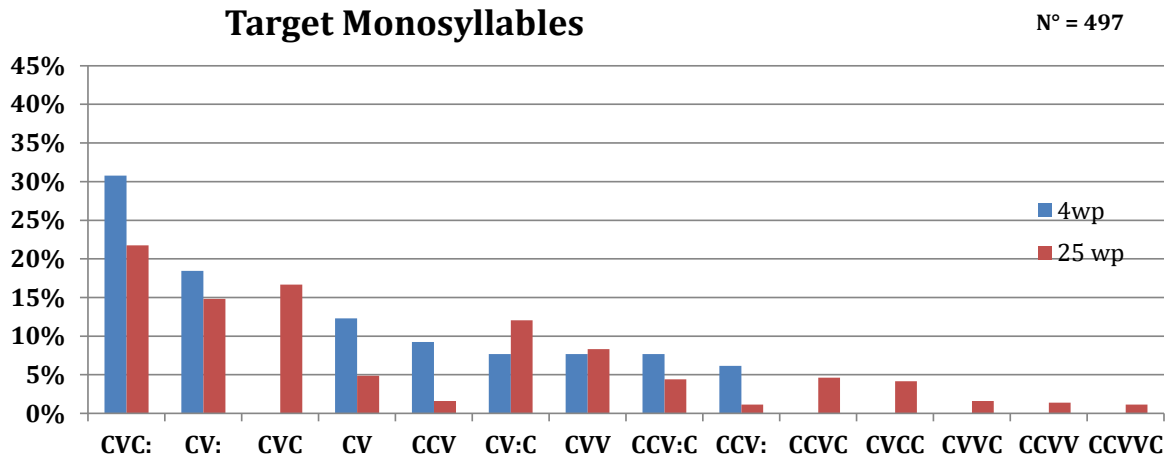

Figure 14.5: Most frequent types of word structures targeted in monosyllabic word shapes at the 4wp and the 25wp. Shapes constituting less than 1% of the data are not included.

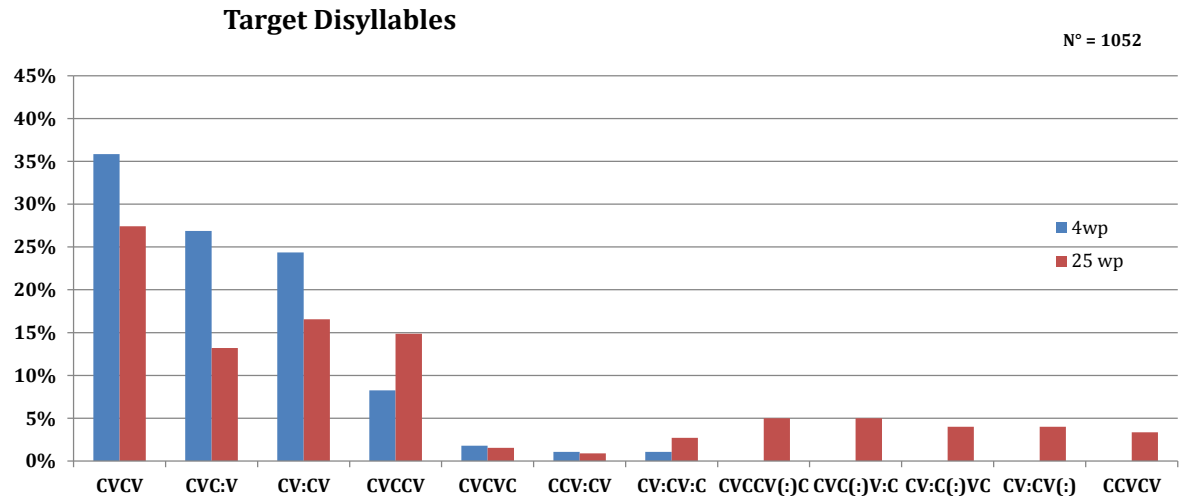

Figure 14.6: Most frequent types of word structures targeted in disyllabic word shapes at the 4wp and the 25wp. Shapes constituting less than 1% of the data are not included.

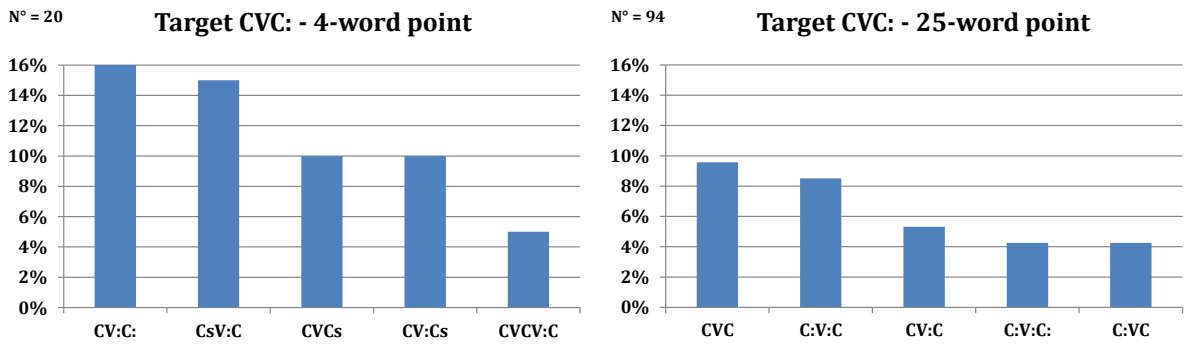

Figure 14.7: Range of realisations for the most frequently targeted monosyllabic word shape, CVC:, at the 4wp (left) and 25wp (right). Here and elsewhere, Cs refers to a consonant that is articulated with extra strength/tenseness.

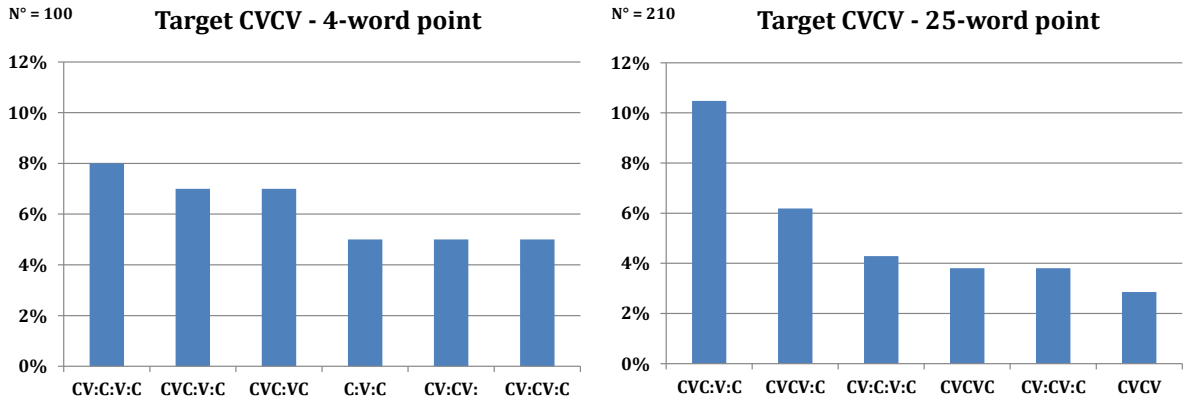

Figure 14.8: Range of realisations for the most frequently targeted disyllabic word shape, CVCV, at the 4wp (left) and 25wp (right).

Table 14.4: Proportions of CVCV shapes being realised with a singleton or a geminate consonant at each developmental stage

	CVCV realisation	
	Singleton	Geminate/Strong
4wp	46%	54%
25wp	35%	65%

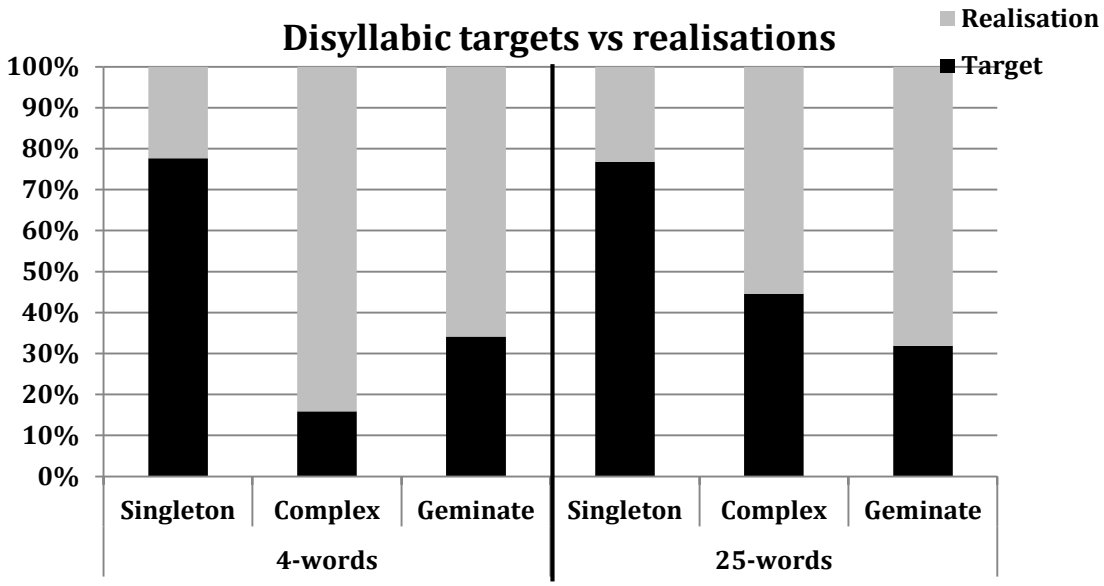

Figure 14.9: Target medial consonant type and realisations in disyllabic productions at the 4- and the 25wps. N = 1142.

The prominence of monosyllabic CV(:)C(:) and disyllabic CV(:)C(:)V shapes in the targets that the children are aiming for throughout the single-word period can also be seen at the individual level, though with interesting differences connected to each child's starting point

(the structure of their earliest words), the relative frequency of each language that they hear, and their individual journey towards the 25wp. The next section looks at longitudinal data from three of the children whose data are presented here in order to explore the interaction between language-specific and individual differences in the development of early phonological structure. In the data presented below only one token per lexical item is presented, chosen from the most frequent and/or most adult-like realisations.

14.5.4 Individual paths and templatic behaviour

14.5.4.1 Martin

Martin was exposed mostly to Arabic, often mixed with French, and his production in the seven sessions that were analysed reflects that exposure (77% of his utterances are Arabic, followed by French at 13% and English at 6%). His 4wp was identified at age 1;3, which is close to the mean age at the 4wp for the children studied here. He is the most systematic of the children in that his earliest productions fell mostly in the CVC:V(C) pattern, and this remained his favourite structure throughout. Below is a more detailed account of Martin's phonological patterns across the one-word-stage.

At the 4wp all but one of Martin's word types have the CVC:V(C) shape, and the majority of these (88% of tokens) show consonant harmony either in the target or the realisation or in both (Table 14.5). Martin adapts both mono- and multisyllables to the disyllabic shape with a long medial C:, e.g. French *train* /tʀɛ̃/ [tɪtta:h] 'train'; /habbu:ba/ [βɐbbæh] 'Habbouba' (nickname for Hiba). He reduces initial consonants more frequently than medial ones, and experiments with the lengths of all segments involved, e.g. /nanna/ 'food' is realised as [nɐ̃nnɐ̃ˑh], but also [nˑɐ̟ːɲɲɐ̟̃ˑh], [θ̟ennɐ̟̃͡ɐ̟̃h], [j̃ɐnnɐ̃ːh], [h̃ɐ̟̃nnɐ̟̃͡ɐ̟̃h], [nɐ̃ˑnnɐ̟̃n], etc. An initial 'anchor' syllable is often used as a form of support for initial consonants, lengthening the initial consonant, e.g. [ʔmmɐ̃nnɛ̃ˑh] for /nanna/ above. As expected, Martin's consonant inventory at the 4-word point is relatively small, mainly consisting of bilabial and alveolar nasals and stops, along with glottal stops and fricatives (Appendix 1).

Over the next two sessions, Martin maintains this pattern but also produces relatively target-like words with disyllabic CV(:)CV(C) and monosyllabic CV(:)C(C) patterns, e.g. /mamɑ̃/ [mɛ̃ˑmɛ̆h] 'mummy'; /ʕaww/ [ʔa̟ːww] 'woof'; /teːta/ [tæ̟ˑtæˑ] 'grandma'. His consonant inventory remains relatively stable, with some attempts to target new stops and glides (/p(p)/ and /k(k)/, /ʕ(ʕ)/ and /w(w)/). There is also an emergence of glottal and vowel-like codas for target codas (e.g. /daħħ/ [dˑæːʔ] 'nice'; /ʔajj/ [ʔaːːh] 'ouch'; /ʕaww/ [ʔa̟ːww] 'woof'), alongside the frequent addition of the glottal and supraglottal codas that Martin and the other children in this study exhibit (e.g. /baʔʔa/ [ʔæʔʔəːm] 'peek-a-boo'; /nanna/ [nɐ̃ˑnnɐ̟̃n] 'food'; /bʊbbo/ [βʊbbuˑh] 'baby'). Sessions 4 and 5 contain a large number of imitated and relatively few spontaneous utterances; since their phonological profile is similar, these have been combined in Table 14.5. These sessions also exhibit the prominence of the CVC:V(C) pattern with consonant harmony, particularly in adaptations (e.g. /faːdja/ [dɐ̟ːddɐ̟̃ː] 'Faadia' (Proper Noun, henceforth PN); French *Jesus* /ʒezy/ [dyddy̟ː] 'Jesus'). Medial consonants in Martin's imitated utterances are more target-like than initial ones, which are more variable. Words with target fricatives and liquids are targeted in imitations for the first time, with frequent stopping and other adaptations, adding to the prominence of consonant harmony.

Table 14.5: Martin's selected and adapted form over the one-word stage. Words were considered *selected* if the adult target matched the pattern of interest and *adapted* if they were modified to fit the child's pattern(s). Shaded grey is used for imitations. Here and elsewhere, the half-length symbol following a consonant was used both for half-long and/or noticeably strong/tense articulation.

1;3	**N (type) = 11**	**N (tokens) = 75**	**% figures are out of all tokens**		
Main Pattern:	**CVC:V(C)**	79%			
Select	32%		**Adapt**	47%	
medial C:, CH			**medial C:, CH, disyllable**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
nanna	nɐ̃nnɐ̃ˑh	food	mamɑ̃	mˑɐ̃mmɐ̃ˑh	mummy
gɪtta	ʔn̪ɐ̃nnɛ̃ːh	Gitta (PN)	baːba	bæˑb̪b̪æ̤ː	daddy
baʔʔa	ʔɐ̃ʔʔəːm	peek-a-boo	teːta	tɪˑttæˑh	grandma
ʔoʔʔo	mʔæʔʔəm	night-night	**medial C:, CH, di-/multisyllable**		
			habbuːba	βɐbbæh	Habbouba
			tɾɛ̃	tɪttaːh	train
			no	nˑɐ̃h**n̪ɪ̃nn̪ɑ̃ˑh**	No
01;05	**N (type) = 9**	**N (tokens) = 54**			
Main Pattern:	**CVC:V(C)**	48%			
Select	28%		**Adapt**	20%	
medial C:			**medial C:**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
buwwaa	bŭβ̞β̞aːh	water	teːta	teˑttɐˑh	grandma
baʔʔa	ʔə**ɯɻ**aʔʔa**ː**	peek-a-boo	waːwa	wɐwwɐˑh	poorly
			maɾijja	ʔɪjjɻ̤ˑ	Maria (PN)
Other Patterns:	**CV(V)CVC**	13%			
tuːtuːt	dæitʉːt	choo-choo			
mamɑ̃	mɛ̃ˑmɛ̆h	mummy			
	CV(V)C(C)	28%			
ʕaww	ʔa̤ːww	ouch	baʔʔa	ʔæː	peek-a-boo
dad(i)	tɑːd	dad(dy)	maɾijja	jɐh	Maria (PN)
1;06	**N (type) = 16**	**N (tokens) = 57**			
Main Pattern:	**CVC:V(C)**	63%			
Select	46%		**Adapt**	17%	
medial C:			**medial C:**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
buwwa	buwwɐˑʰ	water	tapø	tæ̤ˑppʉ̤ˑ	clap
kakka	d̪ɐʰccæˑh	poo	taʕa	d̪aʕʕɑˑh	come here
			baːba	**bɵ**ɲ̊æːβ̞β̞ɑ̃ːh	daddy
medial C:, CH			**medial C:, CH, disyllable**		
nanna	ɲɛ̃ˑnnɑ̃ˑh	food	dˤaww	dɐˈddɐˑʰ	light

bʊbbo	bŏbbuˈh	baby			
dɪdde	dɪddæ̰ˀʔ̰	don't do that			
Other pattern:	**CVVC**	28%			
daɦɦ	d̰a͡ːɛɦ	nice	marɪjja	ɟa̰ːh	Maria (PN)
ʔajj	ʔaːːh	ouch	teːta	tæˈh	grandma
1;07-1;08	**N (type) = 41**	**N (tokens)= 104**			
Main Pattern:	**CVC:V(C)**	58%			
Select	26%		**Adapt**	32%	
medial C:, CH			**medial C:, CH**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
buwwa	βʊwwæː	water	fuːfuː	βʉβ̞βuːh	nickname
wawwaː	β̞ᵇʉˈwwðʉaː	poorlie	mamɑ̃	mᵇõmmõh	mummy
bʊbbo	b̰ʊ̰ˈβ̞β̞ʊ̰ː	baby	teːta	d̰ɛˈtˈæh	grandma
gɪtta	ʔᵉtettæːh	Gitta (PN)	sɪntja	tˡeʊ̈ttʰɐæʰ	Cynthia (PN)
dɪdde	dˡæddɪdᶾe̞ˈh	don't do that	tˤuːtˤuːt	~~tʉtɨtʊ̰ttʊ̰t~~	choo choo
ʒiddo	dᶾɪddoːh	grandpa	marˌtin	tˈæˈtˈɐh	Martin (PN)
kakka	kˀɐˈkkɐ̞ʰ	poo	faːdja	dẹːddæ̰ː	Fadya (PN)
ʔoʔʔo	ʔɵʔʔɵː	night night	ʒezʏ	dyddy̰ː	jesus
			dɔlɪn	g̰ɛg̰g̰ɛ̰ːn	Dollen (PN)
medial C:			**medial C:**		
baʔʔa	bɛ̆ʔʔæ̰ˈh	peek a boo	maːma	ʔɛ̃̃mmæ̰ˈɦ	mummy
			papa	t̪ɪppʰɐːh	daddy
Other patterns:	**CVCV**	16%	ʕɪlke	kẹ̈ttɕ̈h	gum
dodo	dɵdɵː	night night	bitɪmmak	wʉˈddɪˈm	in your mouth
	CV(V)(C)	21%	balõ	ʔəd̰d̰oːʊ͡	ball
taʕ(a)	tˡæˤʕ	come (here)	koːko	ʔăkkoː	nickname
po	pʰʊˈəˈ	Po	marɪjja	dˡæːjjæˈh	Maria (PN)
1;9	**N (type) = 29**	**N (tokens) = 174**			
Main Pattern:	**CVC:V(C)**	51%			
Select	17%		**Adapt**	34%	
medial C:, CH			**medial C:, CH**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
ʒiddo	tʰë̞ttʰɵː	grandpa	baːba	pˈɐ̞̈ppʷæː	daddy
kakka	g͡ʏɐkk͡ʔɐˈ	poo	balõ	bobbo	ball
ʔoʔʔo	ʔɐˀʔʔɑ̰ːh	night night	teːta	tʰë̞ttæ̰ːh	grandma
			ʃaːtˤiɾ	d̰ˈa̰ːttiːˌ	good boy
			koːko	koˈkko	nickname
			trɪkha	gɪkkæˈh	leave her
medial C:			**medial C:**		
bʊbbo	hʊββʊ	baby	marˌtin	mmɐ̃ʔʔi̪ːɦ	Martin (PN)

maɾijja	mẹ̃ɦĭjja:	Maria (PN)	mamɑ̃	w̃ɑ̃mmæ̃:	mummy
lillo	ðˡɛ̰l͡ʒɵ̞:h	nickname	sɪntja	tˑĭs͡ʃt͡ʃɐ:	Cynthia (PN)
Other patterns:	**CV(V)(C)**	19%			
ðats	dˑɜa:t͡s	that's	nanna	næ::n	food
ʒoɾʒ	d̥ˑɵˑe͡tʰ	George (PN)			
nja:ʔ	njɛ::ʔ	yuk			
dɪʔʔ	dˑɛ:ʔ	knock			
naʔʔ	næ̞ʔ	nagged			
ʔe:	ʔˑa̠h	yes			
	CVCV	20%			
ba:ba	pʰɐˑβʷæˑ	daddy	ji:	dˑijɛˑʔ	oh
ko:ko	ko:kɵ̞ˑʰ	nickname	maɾtin	ʔæ̰:ʔʉ̃:	Martin (PN)
tu:tu:t	tʉtʉ:t	choo choo			
1;10	**N (type) = 46**	**N (tokens) = 140**			
Main Pattern:	**CVC:V(C)**	71%			
Select	28%		**Adapt**	43%	
medial C:, CH			**medial C:, CH**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
bʊbbo	bɵbbu:h	baby	ħaki:m	pʰɨˑˈppʰĩ͡ˑĩm	doctor
			balõ	b͡βʷɵˑbbʷɵ͡ɑ̝:h	ball
			mɪʃteːħ	t̠ʷə̹ttʰæ:t˺	key
			mamɑ̃	mæ̰ˑmmæˑʕ	mummy
			noɾma	mɞ̃:ʱʷmmæ̃͡:ɛ̰h	Norma
medial C:			**medial C:**		
nanna	ʲæ̃nnɐh	food	ħali:b	mɐ̰ppʰḭ:pˀ	milk
ʕammo	ɲæ̃:nnm̆ɣ̃ʔ	uncle	te:ta	ç̥ɛ̤:tta̤:h	grandma
buwwa	gʊwwɑ:ʊ	water	tu:tu:t	hʉ̆ttʉˑt˺	choo coo
jalla	hæ̰ˑlla͡ˑɛx	come on	ʒɪzɛl	hɪttʰa:tʰ	Gizelle (PN)
laʔʔa	l͡ð̞ɑ̟ˑʔʔɑ:h	no	mabadde	cɐd̥d̥ɪh	I don't want to
			θaŋk#ju	t͡θɦæ̠k̠ku͡:ɐh	thank you
			ekol	kˀo̞qqˀɔ:p˺	school
			nu:nu:	mə̥̃nnɣ̃:h	potty
			ʔana	ʔæ̞nnæ:h	me
			malʕʔa	mæ̰̃ʔʔæ̤̃:h	spoon
			wi	ʰw͡ɥi:jjɐh	yes
medial C:, multisyllable			**medial C:, multisyllable**		
tattu:na	tʉttʉ:næh	nickname	ʕafa:f	t̠ʰæʰppˑɐpʔˑɐp	Afaf (PN)
batˤtˤa:ɾɪjje:t	d̥ʲe̞ˑt̠t̠ʰɛ̞ˑtʰɛ͡:ĩh	batteries	bobi	he̞**bɵ:bbi:h**	doggy
maɾijja	ma̠ˑɦĭjjæ:h	Maria (PN)	wa:wa	ʔʊwɐwwa:h	poorli
			dodo	ʔɪ̤̆ddˡɵ͡:ɪddø:h	night night

			pardõ	lɐ̪̈ʔ̪̈ɪd̪̈d̪̈ɛ̃ːɐ̪̈ḧ̪	pardon
Other patterns:	**CV(V)(C)**	12%			
wuff	bʷʊ͡u	woof	ʔoʔ̈ʔo	ʔɞˈʔ̈ʔ	night night
daħħ	ⁿ͡tˈæˌħ	nice			
taʕ(a)	t̪ʰˈæ̤ˌh	come (here)			
jiː	ɟˈiːh	oh			
majj	mmq̰ːjj	water			
weːn	ⁿ͡wɛˌn̆	where			

Martin's last two sessions show a marked change in the frequency of words produced as well as a growing consonant inventory (Appendix 1), but Martin's preferred CVC:V(C) pattern is still prominent, with adaptations that are twice as frequent as the selected words with this pattern in session 6. These stand out compared with the other minor word shapes that Martin produces, which tend to be more accurate, e.g. CV(V)C(C): English *that's* /ðats/ [dˈɜaːt͡s], Arabic /weːn/ [ⁿ͡wɛˌn̆] 'where'. The majority of words that Martin targets are still disyllables, and despite his increased phonetic and phonological inventory his productions still exhibit frequent consonant harmony. In the last session the CVC:V(C) shape rises to 71% of all of Martin's productions, the highest since his first session, which suggests that the medial long consonant template is at its most productive for Martin as he approaches the 25wp. Consonant harmony is not as prominent in this session, as new consonants are attempted and coda consonants are more frequent. Session 7 also sees the geminate/long pattern being applied to longer words as Martin starts producing multisyllabic words; multisyllabic words with medial geminates like /tattuːna/ [tʉttʉːnæh] 'nickname for Martin' and /batˤtˤaːrɪjjeːt/ [t̪ɛ̈ˌcc̟ʲɛːtt̪ɛ̈ːĭt̪˺] 'batteries' are selected, while disyllabic words are sometimes adapted to the multisyllabic shape with one or two internal long consonants, e.g. /bʊbbo/ [bɐ̤ˌb̪ˌɐ̤ˌb̪b̪ɐ̤ːh] 'baby' and /nanna/ [jæɳɐɲɲq̰ːh] 'food'. Although Martin's consonant inventory is expanding, variation in the realisation of some consonants is higher than in earlier sessions, especially in initial position e.g. for /b/, /k/ and /m/ (Appendix 1). Medial codas are targeted but are often assimilated to the next onset, adding to the geminate pattern, e.g. /malɪʕʔa/ [ʔɛ̈mmæ̤ˌʔʔæ̤ːh] 'spoon', /min hoːn/ [xɜ̃ˌnñ̪ːn̆h] 'who's there?' and /mɪfteːħ/ [cɛ̤ˌt̪t̪aːt͡θ] 'key', but awareness of medial codas is noticeable and some disyllables are adapted to that pattern, e.g. /mamɑ̃/ [bɐmd̪̊æːh] 'mummy', /koːko/ [kˀɵˌl̆kˌɐ̤ːh] 'nickname'.

14.5.4.2 Rama

Rama was exposed to more English than Martin, and 30% of her utterances were English. Arabic still constituted the majority of her utterances at 60%. Rama's first two sessions are combined, due to the small number of spontaneous utterances in her first session (Table 14.6). Her profile at this early stage of production is strikingly different from Martin's, mainly due to the high frequency of monosyllabic words that she produces (46%). The disyllabic geminate pattern is prominent as well (40% of utterances), with adaptations such as /baːba/ [mɐbbɑ̤ˌh] 'daddy'; /kɪle/ [kɪððe] 'eat!'; /hajda nɔːz/ [ʰd͡ᶻænnɯ̲ˌz] 'that's (a) nose'. Perhaps due to Rama's jargon practice and older age at this stage, her consonant inventory is more varied than Martin's at the 4wp (Appendix 1), with a small number of fricatives and laterals alongside stops and nasals as well as final consonants and occasional two-word

utterances (e.g. /jalla kıle/ [jakıððe] ‘come on eat’; /pʊtɔn mını/ [pʊdõ mıdĩŋ] ‘put on Minnie’).

Table 14.6: Rama’s selected and adapted forms over the one-word stage; shaded grey is used for imitations

1;5-1;6	**N (type) = 22**	**N (tokens) = 67**	**% figures are out of all tokens**		
Pattern 1:	**CV(:)(C)**	46%			
Select	29%		**Adapt**	17%	
V = i(:)			**V = i(:)**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
jiː	hˈiːh	oh	tivi	mĩ	TV
miːn	mˈĩʰːh	who			
θɾi	t͡nĩːh	three			
V other			**V other**		
ʔeː	ʔˈɛːh	yes	happi#bɜɾθdeɪ	bˈæˈħ	happy birthday
daħħ	ɟˈæħħ	nice	hajda	d͡jɐʔ	this
mbuː	mʙʉˈ	water			
fuː	fˈʉːh	crasy			
Pattern 2:	**CVC:V(C)**	40%			
ʔoʔʔo	ʔaʔʔo̤ˈh	oh-oh	mamɑ̃	mˈæ̰̆mmæ̃ˈh	mummy
			baːba	mɐbbo̤ˈh	daddy
			mını	mınniː	Minnie
			hajda#nɔːz	ⁿd̪ˀænnu̟ˈz	that's a nose
			kıle	kıððe	eat!
1;7-1;8	**N (type) = 25**	**N (tokens) = 50**			
Pattern 1:	**CV(:)(C)**	38%			
Select	22%		**Adapt**	16%	
CV(:)glide					
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
hajj	ɦæ̟j	this			
baɪ	bˈæjj	bye			
heɪ	hĭeːjj	hey			
waːw	bˈo̤ʷˈww	wow!			
V = front mid-close to mid-open			**V = front mid-close to mid-open**		
dɑ̃s	tˈeh͡s	dance	mijaw	ᵐnɛ̃ˈm̥ʰ	miaw
daħħ	n͡dðæħ	nice	mɛɾsi	βɛːh͡s	thanks
			ʔeɾjɛl	ʔĕe̤ːh	Ariel
Pattern 2:	**CVC:V(C)**	48%			
ʔoʔʔo	ʔɔʔʔɔh	oh-oh	mamɑ̃	mˈæmmɐh	mummy
			teːta	tˈa̟tˈa̟h	grandma
			haj#daħ	hæʲddæ̟ˈħ	this (is) nice

1;9	**N (type) = 35**	**N (tokens) = 86**			
Pattern 1:	**CV(:)(C)**	40%			
Select	39%		**Adapt**	1%	
C(C)V(:)glide					
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
majj	mɛɪh	water	noʊz	nɛɪ:	nose
aɪz	ʔai:	eyes			
aɪ	aɪjj	I			
miːn	kbiːə	who?			
CV(:)(C)					
baɦɦ	bæːɦ	all gone	hajda	ddə	this
bɔl	baːːh	ball			
ʔaː	ʔa̤	A (letter)			
haʊs	hæː	house			
mbuː	mbuːu	water			
bʊk	buːh	book			
ʒus	ʔuːṳs̤	juice			
po	poː̤	Po			
noʊz	noːːz	nose			
beːt	be̤ː	house			
leːk	lək	look			
Pattern 2:	**C(:)V(:)CV**	29%			
teːta	teta	grandma	banana	næːnaːh	
beɪbi	beːbiː	baby	baɪ	ba̤vɪiˀ	
hajda	hæɪda	this	ʔoʔʔo	ʔæːʔʉː	
haʔa	haʔə̰	no			
maːma	mmaːma̤h	mummy			
laːla	llæːlæ̰	Lala			
Pattern 3:	**CVC:V(C)**	20%			
baddo	bɛddo	he wants	tiktak	tiːttih	sweet
ʒiddo	ʒ̤ɪddo̤	grandpa	hajj	ʔəhhaiːh	this
(ʔɪ)dˤaww	ʔɪddaʊh	(the) light	ovən#glʌvz	ʔaʔʔuː	oven gloves

In the next two sessions (age 1;7 and 1;8) the two patterns identified at the 4wp still make up the majority of utterances, though the prominence of the disyllabic geminate pattern is due more to frequency of use (48% of utterances) than to type (28% of the total of different words). Monosyllabic CV(:)(C) is the most varied and productive shape, showing a final glide pattern (e.g. /hajj/ [ɦæ̤j] 'this'; /baɪ/ [bˈæjj] 'bye'; /waːw/ [bˈb̤ˈww] 'wow') and a front mid-high to mid-low vowel pattern (e.g. French *danse* /dɑ̃s/ [tˈeh͡s]; *merci* /mɛrsi/ [βɛːh͡s]; /mijaw/ [ᵐnɛ̃ˈm̰] 'miaow').

In the final session the monosyllabic CV(:)(C) shape becomes the most prominent, accounting for 40% of all utterances. Within this shape a subset of productions still have the

final glide pattern, as in previous sessions (Table 14.6), but others include other consonants as well and a rich variety of vowels (e.g. /baħħ/ [bæːħ] 'all gone'; /mbuː/ [mbuːu] 'water'; English *Po (name of TV character)* /po/ [poː]). The influence of words of English origin is obvious in the frequency of monosyllabic words in Rama's sessions, with words like *Po*, *bye*, *wow*, *ball*, *eyes*, *book* and *nose* making up a large proportion of her productions, especially in the last session. The second most frequent pattern in this session is a disyllabic C(:)V(:)CV shape (29% of utterances), which takes over from the medial long C(:) as the second most frequent shape (e.g. /teːta/ [teːta] 'grandma'; /maːma/ [mmaːmạh] 'mother'; English *baby* /beɪbi/ [beːbiː]). These and all but one of the monosyllabic words are 'selected' and, apart from expected developmental features, they are fairly accurate. In fact, most of Rama's productions in the final session are essentially accurate; in comparison with Martin, she produces fewer utterances and fewer repetitions of words (98 types and 203 tokens over 5 sessions for Rama, compared with 179 types and 604 tokens over 7 sessions for Martin) but the words tend to be more accurate and her production exhibits no large-scale adaptations to any preferred shape. The only pattern that still shows more adaptation than selection is the disyllabic long/geminate pattern (e.g. /tiktak/ [tiːttih] 'sweet'; /baħħ/ [bæħħaʊ] 'all gone'; English *oven gloves* /ovən glʌvz/ [ʔaʔʔuː]), though the frequency of occurrence of this pattern is now down to 20%.

14.5.4.3 Lina

Lina was exposed to both French and English on a regular basis, and her production reflects that, with Arabic productions constituting only about half of her overall utterances at 48%, followed by French (28%) and English (21%). Her first three sessions, between 1;3 and 1;5, had similar patterns and no increase in the number of spontaneous words produced, so they are combined here for analysis. In these sessions, Lina's utterances can be grouped into the three word shapes identified for the other two children, with the disyllabic pattern with a long medial C: being the most frequent (e.g. French *Oui Oui* /wiwi/ [ʔɪwwih]; *maman* /mamɑ̃/ [mæːʰmmæːː]; English *thank you* /θaŋk ju/ [ʔʰættʊ]). Interestingly, these early words are all disyllabic French or English targets with lengthened medial consonants. Lina also produces disyllabic words with short medial consonants (e.g. /ʕalam/ [ʔɐðæˑh] 'flag'; /ʔalo/ [ʔævʉː] 'hello') and monosyllabic shapes which consist of either monosyllabic targets (mostly English and French numbers and letters of the alphabet [Table 14.7]) or reductions of disyllabic words (e.g. /ˈsabat/ [bæˑt] 'it stayed still'; /ˈʔaʕtˤi/ [ʔɐˑtʂ] 'give'). In terms of her consonant profile, over half of Lina's words have selected or adapted initial glottals, while in medial position she targets and produces a collection of bilabial and alveolar sounds; these include /l/, which is advanced relative to her age but which might relate to her own name having an /l/ in it (see data from Laurent in Vihman 1993). Other relatively advanced sounds that she produces include dental, labiodental, and alveolar fricatives (Appendix 1).

Table 14.7: Lina's selected and adapted forms over the one-word stage; shading is used for imitations

1;3-1;5	**N (type) = 20**	**N (tokens) = 66**	**% figures are out of all tokens**		
Pattern 1:	**CVC:V(C)**	46%			
Select			**Adapt**		
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
ʔɪjja(ʔɪjjaʔoː)	ʔɪjjæːh	song	wiwi	ʔɪwwih	Oui Oui

			θaŋk#ju	ʔʰæ̤ttʊ	thank you
			mamɑ̃	mæːʰmmæ̤ː	maman
			pepa	b̥æ̤ppæh	Peppa
			ʔeːjbiːsiːdiː	ð̥iˈd̥aːtti:	ABC song
Pattern 2:	**CV(:)CV**	34%			
ʕalam	ʔɐðæˈh	flag			
ʔalo	ʔævʉː	hello			
papa	θɐ̝ˈtæˈʰ	papa			
Pattern 3:	**CV(:)(C)**	20%			
ʔeː	ʔeːʲ	letter A	wabat	bæˈt	stayed still
di	diːʲ	letter D	ʔaʔtˤi	ʔɐˈtʂ	give him
			buːse	ʰp͡siˈh	kiss (her/him)
1;6-1;7	**N (type) = 25**	**N (tokens) = 59**			
Pattern 1:	**CVC:V(C)**	34%			
Select	5%		**Adapt**	29%	
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
lallo	**ʔ͡ɪlæˈlləˈh**	nickname	ʔɛ̃#døː	ʕæˈttː͡øːɜ̃	un deux
tˤajjaːra	dɐ̝ˈllɛːh	plane	ʔaʔtˤi	ʔæʰˈttˢiː	give him
ʔɪjjaʔɪjjaʔoː	ʔɪ̈jjæːiːʰ	song	θaŋk#ju	ʔæ̤ˈttiː	thank you
			haːti	hɐtti	give me
			tneːn	ʔɪnnɛˈn	two
			tɾwa	t̪ɐ̝ˈjjeh	trois
			ʃokola	kollɛ̃ᵉʕ̃	chocolat
			doːɾa	dˈʊˈwwæ̤ˈh	Dora
			maːma	mæ̤mmæ̤ˈh	mummy
Pattern 2:	**CV(:)(C)**	53%			
laʔʔ	ʔ͡læ̤ʔ	no			
ʔɛ̃	ʔɛ̃	un			
fɪʃ	pɪs	fish			
dø	d̥oah	deux			
tu	d̥ˈɵːh	tu			
no	nˈo̤ˈʔ	no			
Pattern 3:	**C(C)V(:)CV**	14%			
doːɾa	tˈæ̤ˈⁿ̃ɔ̃ɛh	Dora (PN)	katɾ	ʰt͡cæˈɟiːh	quatre
liːna	ɬʰɪn̪ɛ̆	Lina (PN)	dø	n͡d̪oɽiːh	deux
hɛloʊ	hɐloˈ	hello			
1;8	**N (type) = 19**	**N (tokens) = 57**			
Pattern 1:	**CVC:V(C)**	51%			
Select	22%		**Adapt**	29%	
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**

ʔoʔʔo	ʔeʔʔo̞ˑh	oh-oh	mamɑ̃	mæ̰̃mmæ̃ˑʔ	maman
ʔɪlbʊbbo	ʔɪl**b̥iˑbbɵːh**	the baby	bebe	bi̞ˑd̪d̪ᵇe̞ˑh	bebe
			laːlo	hæ̤llɘ̝ˑh	nickname
			tɾwa	tˑɻ̩wwæːh	trois
			multisyllabic/across word boundary		
			laːla	θ̪ellɐ̠llæˑh	Lala
			wiwi	lɪʷvˑiːwwĭ	Oui Oui
			do̤ːɾa#leʕbe	ðəllæ̰͡dvɛ̤ˑh	Dora toy
			laʔmabadde	ða̠llaβæ̰ˑddḭ̤ːh	I don't want to
Pattern 2:	**C(C)VV(C)**	30%			
ʔɛ̃	ʔɐæ̠h	un	tˤlaʕe	tˑæ̤ˑħ	come up
dø	d̪ˑɵːh	deux			
katɾ	kʰæˑh	quatre			
bi	d͡ʝ̥iːh	letter B			
di	t͡ð̪iːh	letter D			
Pattern 3:	**CV(:)CV**	16%			
do̤ːɾa	tˑʊwɐh	Dora (PN)			
liːna	liːɲæ̤ˑh	Lina			
laːla	lælæ̟ˑh	Lala			
1;9	**N (type) = 19**	**N (tokens) = 59**			
Main Pattern:	**CVC:V(C)**	79%			
Select	8%		**Adapt**	71%	
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
nanna	næ̤̃ˑnnḭ̤	food	baːba	pæ̟bbʉ̟ˑ	daddy
			apəl	hæ̤ppʊl	apple
			mazbuːtˤ	mˑɐddʉˑt̚	right
			bajdˤa	bʷʊˑddæ̰	egg
			haːti	ʔæ̠ːttiˑ	give me
			teːta	tæ̟ːˤttæˑˤ	grandma
			kompjutəɾ	tʉːttæˑʰħ	computer
			mamɑ̃	mæ̃ˑmmæ̟̃ˑʰ	maman
			majj	mæ̃ˑjjæʲˑʰ	water
			laːla	ð̥ˡɑːɭɭæ̤ˑħħ	Lala
			fɪʃ	pʰɪs̪s̪ʉ	fish
			suːsuː	tsɪssʊˑ	nickname
multisyllabic			**multisyllabic**		
ʔoʔʔo	hɪʔʔɪjɜʔʔoh	oh-oh	dʒeːʒe	d͡z̪ə̝z̪z̪ʒ̟͡lzzɛːh	chicken
1;10	**N (type) = 64**	**N (tokens) = 206**			
Main	**CVC:V(C)**	57%			

Pattern:					
Select	12%		**Adapt**	45%	
Target	**Form**	**Gloss**	**Target**	**Form**	**Gloss**
tappø	tʰæˑpp͡vʉːʰ	shoes	lɪʕbe	he̝bbɛ̤ːħ	doll
waʔʔaʕ	wˑæ̠ʔʔæ̠ˤ	he dropped	ʔɪʔʕʊde	ha̟ddɛ̠ːh	sit down
ʔoʔʔo	ʔo̟ʔʔɵʲ	oh-oh	batˤaːtˤa	tæ̠ʲttæːh	potato
			hajdi	ħaːddɛˑh	this
			haːti	hittʰɛːh	give me
			kitti	ʔĕttᶴi̠ˑh	kitty
			teːta	ɟettʲʰæ̟h	grandma
			lili	lli̠lli̤h	Lili (PN)
			lilli#bebe#lilli	bˑḛlle̝ːh	Lili bebe
			ʃokola	k͡xɵllæˑ	chocolat
			hɛloʊ	hæˑll͡əʊ	hello
			maɾjam	bˑɪjjæ̤ːh	Marjam
			dawa	dˑɐ͡ˑʊʊæ̠h	medicine
			bɾa̤ːvo	b͡wæ̠vvɵːh	bravo
			neːme	nˑemme̤ħ	sleep!
			nanniː	l͡əɛːnnɪ̤ˑh	food
multisyllabic/across word boundary			**multisyllabic/across word boundary**		
ʔɪjjaʔɪjjaʔoː	ʔĭʒ̆ʔĭjɐʲʔʔɔ̝h	song	kompjutəɾ	ʔĩtʉˑttɛ̝ː	computer
			bɪskoːte	ʔəkkˀo͡ˑĭtˑɛ̤ːh	biscuit
			mamɑ̃	ʔæ̃mmẽmmẽ̝h	maman
			televizjõ	ʔeððiˑzˑø̠ˑh	television
			dɪŋdaŋdoŋ	d̥ˑɛˑzzḛˤloːh	ding dang dong
			ʔɪlʔalo	ĭlʔæ̠llɞˑh	the hello
			laːlo	ʔɪllæˑlo	nickname
			laːlapo	ð͡le̝llæ̠ːp͡vˑʉˑh	Lala and Po
			waʔʔaʕ#batˤatˤa	wɐʔʔɐ̟ˤt̠t̠ɐ̟ˑt̠ɐ̟ˑh	he dropped potato
			waʔʔaʕte	wĕʔʔæ̠ˤːddi̠ːʰh	you dropped
			ɾaːħ#kiti	jæ̤ħᵊç̟ˑet̠t̠ɛ̟h	Kitty's gone
Pattern 2:	**C(C)VV(C)**	35%			
with strong/long C1			**with strong/long C1**		
daħħ	d̥ˑɐ̟ʰħ	nice	ʒiddo	ddɐ̝	grandpa
dø	dʲˑœːh	deux	taʕa	tˑɐɐħ	come (here)
tneːn	nˑeˑʰh	two			
ʔeː	ʔˑe̝ːh	yes			
ʔajj	ʔˑæʲːjç	ouch			
si	ssiˑh	si			
ʔeːj	ʔˑɛ̰ː͡ɪjħˑ	letter A			

tɾwa	x͡wʊ̈ːa͡ːh	trois			
bi	bˈiːɦ	letter B			
gou	gˈɐ̟	go			
with heavier rhyme/coda			**with heavier rhyme/coda**		
fɪʃ	ɸe͡hsː	fish	bjuːʔaf	ʔæ^{h}ff	he stands
moxx	ɣeχχ	brain			
weːn	ɪ̃n͡l	where			
ʔɛ̃	ʔɐ̈ˑa͡h	un			
ʔoː	oː̟͡ɜh	letter O			
waːw	v^{w}æː^{h}	wow!			
hajj	hɐʔ	this			
xoːx	k˺͡xo̟ːx	plum			
Pattern 3:	**CV(:)CV**	12%			
ʔalo	ʔælɐ̈ˑ	hello			
nimo	mˈiˈməˈh	Nemo			
kɪle	k͡xːuɛ̟͡liːh	eat!			

Over the next two sessions Lina's production of monosyllabic CVC words increases, mostly due to her engaging in various games around counting and reciting the letters of the alphabet in French and English with her mum. Her disyllabic CVC:V(C) pattern is less prominent during those sessions, but interestingly it is the only pattern which shows active adaptations on Lina's part in comparison with the mostly selected words from the monosyllabic and disyllabic shapes with medial short C (e.g. /ʔaʕtˤi/ [ʔæʰˈttˢiː] 'give'; French *trois* /tɾwa/ [t̪ɐ̟ˈjjeh]; *chocolat* /ʃokola/ [kollɛ̰^{ɐ}ʕ̰]). This pattern becomes stronger and more prominent over the next two sessions until it makes up 79% of all of Lina's productions at age 1;9, the session immediately prior to her 25wp. Lina is actually very close to the 25wp at age 1;9, as in the last session she produces 41 spontaneous different words and has therefore moved beyond the first 50 words (she also produces many words in general during the last session, four times as many as in the previous session). Words like *fish*, which previously had monosyllabic realisations, now acquire disyllabic forms (1;6 [pɪs] but 1;9 [pʰɪs̪s̪ʉ]). Lina also starts applying the lengthening pattern to multisyllabic targets (e.g. /televizjõ/ [deˑðittɐ̟ˑh] 'telly'; /bɪskoːte/ [ʔəkkʔo͡ɪ̈tˑɛ̟ːh] 'biscuit'), multisyllabic realisations of disyllabic targets (e.g. /doɾa/ [ʔʊwwɐlɐ̟ˑl̪͡dᶻ] 'Dora'; /laːla/ [θ̪ellɐ̟ˑæˑh] 'Lala'), or across word boundaries (e.g. /laʔ ma badde/ [ðə̲llaβæ̈ˑddi̤ːh] 'no, I don't want to'; /waʔʔaʕ batˤatˤa/ [wɐʔʔɐ̟ˤt̪t̪ɐ̟ˑt̪ɐ̟ˑh] 'he dropped potato'). Therefore, despite the frequency of French and, to a lesser extent, English words in Lina's vocabulary, her disyllabic pattern with a medial geminate has become as strong as Martin's by the 25wp. This may be due to the lengthening of many medial consonants in French words by adults in the community. Of the monosyllabic words that Lina produces in that last session a large proportion (55%) now have a long or strong/heavy first consonant (e.g. /ʒiddo/ [ddɐ̟] 'grandpa'; /daħħ/ [d̥ˑɐ̟^{h}ħ] 'nice'; French *si* /si/ [ssiˑh]), perhaps showing influence from the disyllabic geminate pattern. The same applies to more than half of the disyllabic shapes with short medial consonants (e.g. /nimo/ [mˈiˈməˈh] 'Nemo'; /laːlo/ [lˈɐle] 'nickname for Elias'). The remaining monosyllabic and disyllabic productions are mostly selected and largely accurate.

14.6 Summary and discussion

This study looked at early production patterns in five Lebanese-speaking children between the beginning and end of the one-word-stage. The aim was to provide new data on early word shapes in Lebanese Arabic and to look for patterns in the children's production which may be indicative of the templatic behaviour reported in other languages. In terms of word shapes, we predicted that children's early words would show the influence of the frequent disyllables with medial gemination that are common in Arabic and that medial and final consonants would be acquired early, leading to the early emergence of relatively complex syllable structures. In terms of templatic behaviour, we predicted that disyllabic shapes with medial long consonants would dominate children's preferred patterns and lead to adaptations of other word shapes to the CVC:V(C) shape; moreover, given that template patterns are influenced by language exposure and the child's individual experience with early words, we predicted that individual differences and the children's varying exposure to English and French would also play a role in how early these patterns would appear and how systematic their productions would be. The findings support our predictions and highlight the special role of phonological length in Arabic in the child's acquisition of lengthening as a suprasegmental feature and the children's tendency to over-generalise this feature before achieving target-like production. Below we revisit some of these findings and discuss their implications for the relationship between accuracy and phonological advance.

14.6.1 The prevalence of disyllabic structures from an early age

The data presented here support the rich and minimally bimoraic word shapes in the Arabic language (Broselow, 1992; McCarthy & Prince, 1986; 1990; Watson, 2002) are exhibited in Arabic-speaking children's early word production. Furthermore, the difference in the distribution of word shapes from the three languages targeted by the children (Fig. 14.1) provides an insight into how the prosodic shapes of early words vary across languages. While the Arabic and French words targeted were mostly disyllabic, the majority of English words targeted were monosyllabic with codas. As a group the children produced Arabic the most, followed by English and then French. Disyllables were therefore targeted the most, and monosyllabic and multisyllabic words were often adapted to the disyllabic shapes. The children also frequently produced a filler syllable at the beginning of the word (Peters, 2001), which increased the percept of multisyllabic production. The use of initial filler vowels or syllables by children as a speech initiation strategy is not uncommon (see, for instance, Si's data in Macken, 1979). In this study, the most common filler used by the children was a CV syllable consisting of a glottal stop followed by a neutral vowel, but there were other CV shapes as well; our impression is that children often used these as a springboard for word production, as if to initiate articulation. Another possibility is that the children were producing dummy syllables based on the frequent occurrence of the definite article /ʔal/, which assimilates to coronal onset consonants in following nouns (e.g., /ʔal/ + /ʃams/ is realised as [ʔaʃʃams] 'the sun').

The children produce a wide variety of syllable structures from an early age, including syllables with final codas. Final consonant deletion, which is common in the production of children acquiring English and Spanish (Macken, 1979), was not found to be frequent in the production of the Arabic-speaking children in this study. In fact, these children were more likely to add final codas to words which would otherwise end in vowels than to delete them. These results therefore agree with other acquisition studies which have suggested that Arabic-

speaking children acquire a range of complex syllable shapes from an early age (e.g. Abdoh, 2011; Ammar, 2002).

14.6.2 The role of gemination in phonological advance

As a group, the children both target and produce more disyllables with geminate/long consonants than any other word shapes. Further work on the frequency distributions of word shapes in the adult language is needed, but the sparse literature on Arabic phonology suggests that ˈCVC:V(C) is a frequent and productive pattern in the language, being used in both nouns and form II verbs (Watson, 2002). A large part of the CVC:V(C) realisations were also adaptations of a ˈCV:CV(C) target, with the children shifting length from the preceding vowel to the medial consonant (e.g. /baːba/ realised as [babbah] 'daddy'). Lengthening was often applied to more than one segment in a word and was also variable. As Macken (1979: 29) points out, when words are treated as prosodic units the child may freely swap features within the unit, the feature being swapped here being segment length. Initial consonants were also occasionally preceded by filler syllables, which turned the original initial consonant to a medial one that was then lengthened (e.g. French *dodo* /dodo/ 'night night' realised as [ʔĩ̪dˈdɵːɪddø̞ːh]).

While the prosodic CVC:V(C) shape was consistent target that the children aimed for or adapted words to, their production of the segmental material in each word was quite variable, as evidenced by analysis of several repetitions of the same word. On the whole, initial consonants varied more than medial ones and were more often reduced, but interestingly there was hardly any case of the initial consonant deletion that is often reported for languages with medial geminates, where the geminate position diverts the child's attention to the medial consonant (Bhaya Nair 1991; Savinainen-Makkonen 2007; Vihman and Velleman 2000; Vihman and Vihman 2011). So while the children's higher accuracy for medial consonants and codas chimes in with findings on other Arabic dialects (e.g. Amayreh and Dyson 1998; Dyson and Amayreh 2000; Shahin 2003) and other languages (e.g. Bhaya Nair, 1991; Szreder, this volume), the importance of onsets in the phonological structure of Arabic words may have played a role in the maintenance of onset consonants by the children, even if their realisation was more variable.

Although the children were on the whole more accurate in their segmental productions towards the end of the 25wp, their realisation of phonological length became less accurate as they adapted more words to the geminate template (e.g. Figure 14.10, Table 14.4). This coincided with their vocabulary showing a quantum leap in terms of word types and/or tokens (Table 14.2). Apart from the children becoming more systematic in their production of the CVC:V(C) pattern, two factors contributed to the increased production of medial long consonants: a) targetting of medial codas, which were often assimilated to the following consonant (e.g. /xamse/ *five* realised as [θ͡sɐzzɛh]) and b) the emergence or, for some children, increase in the production of multisyllabic words, in which one or more medial consonants were lengthened in the same way as disyllables. We hypothesise that this U-shaped curve or decrease in accuracy, which is often reported in other studies, is the children's way of using a well-practiced and articulatorily accessible production routine, the CVC:V(C) shape, to aid them in aiming for and learning new and longer or more challenging words. The result was less accuracy in achieving target phonological length in the later recordings due to overuse of the medial long consonant pattern, even as children's phonetic and phonological inventories were starting to look more adult-like.

As Savinainen-Makkonen (2007) observes for Finnish children, we think that Lebanese-speaking children use the CVC:V(C) prosodic shape as an anchor to practice new words and adapt them if their target form does not fit that shape. The outcome may not resemble the patterns found in the adult phonology, but it is a sign of the children being actively involved in 'doing phonology'; this is evident both in the way that the children select groups of words that match the phonological structures that they hear in the input and that they are able to produce, and the way they adapt other words to fit the prosodic shape that they are familiar with producing. This comes at a time when the children have more articulatory control and a richer consonant inventory, and therefore have less maturation-related reasons to lengthen target singleton consonants. Given that two of the three children whose individual data we looked at here show regression in accuracy in terms of the realisation of phonological length towards the end of the 25wp, this calls into question whether their earlier sessions with target-like length show 'true' acquisition of the singleton-geminate contrast. We suspect that the early accuracy might reflect an item-learning phase when the link between singleton and geminate consonants has not yet been acquired. In the later recordings, the children's overuse of long durations suggests their growing attention to this salient phonetic and phonological characteristic of Arabic and their application of length as an active process in the production and learning of new words. We predict that the return to accuracy in the realisation of consonant length, which is to be expected in the third year of life – once templatic behaviour has receded and the children's productions are more adult-like, will represent 'real' acquisition of gemination.

14.6.3 Individual differences

As part of normal variation of language and linguistic use within Beirut, the data reported here show varying use of words from English and French across the five children, and this contributed to the individual differences that were evident in both their segmental development and their early word shapes (though only the latter was dealt with in detail in this paper); it also made a difference to whether or not the children showed any systematic patterns in the early recordings. For instance, one reason Martin appeared to be the most systematic from the start is because he targeted a higher proportion of Arabic words than any of the other children, and many of these were disyllabic. This, together with his frequent use of consonant harmony and over-reliance on the medial geminate pattern, made his productions look very systematic and template-like from the earliest recordings, when typically there are not enough productions for any patterns to stand out. The prevalence of consonant harmony and medial lengthening in the later recordings, at an age when his consonant inventory was expanding and he was beginning to produce multisyllabic words, cemented the conclusion that the $C_1VC:_1V$ pattern is a preferred shape for Martin rather than the consequence of articulatory constraints and/or a small consonant inventory. This early systematicity was not found for the other children and confirms findings elsewhere that not all children apply consonant harmony as an active phonological process (Macken 1979; Vihman 1978).

Rama, on the other hand, targeted more English words from the start than Martin or Lina and, as a result, produced many more monosyllabic structures than the other children. This, added to the fact that her productions were generally accurate and that she did not reach the 4wp till relatively late, made it more challenging to capture a systematic stage for her before her productions became target-like. Within her monosyllabic productions, a weak pattern for final glides could be pinpointed in the later sessions, similar to what has been reported for English children in final position in monosyllables and in medial position in disyllables (e.g.

Priestly 1977; Vihman et al 1994). The monosyllabic shape remains the preferred one for Rama and constitutes 40% of her productions at the 25wp, but what is interesting is that it consists more of selected than adapted words. Her disyllabic medial long consonant pattern, on the other hand, is in decline in the last session (constituting 20% of her productions) but shows more adaptation than selection, suggesting that even for a child like Rama, who produces more monosyllables, templatic behaviour is evident in her disyllabic productions.

Lina's profile too can be partly linked to her language exposure/use, with around half of her productions consisting of Arabic utterances, followed by French and then English. With French having frequent disyllables like Arabic, Lina's productions were, as expected, mostly disyllabic; but while the disyllabic shape with a medial long consonant emerged as the most frequent pattern in the early recordings, it was not as systematic as what was found in Martin's data, and Lina still produced many monosyllabic and disyllabic words with singleton consonants. Out of the datasets presented in detail here, Lina's longitudinal data provide the best example of a decrease in accuracy as a result of the application of a templatic pattern in the later sessions. Following the early sessions in which the disyllabic CVC:V(C) pattern was frequent in Lina's productions, her middle sessions were more accurate and more diverse in terms of the word shapes produced, with many fairly accurate monosyllabic productions. Towards the 25wp, however, the CVC:V(C) pattern again became more prominent, with many adaptations and a decrease in accuracy (for example, the decrease in accuracy in the production of *fish*), just as her vocabulary was rapidly expanding. We see a qualitative difference between the apparent systematicity of the early sessions, where Lina's lexicon is still small, and the later more active application of the medial geminate pattern, at a time when articulatory control is more advanced. More research is needed to look at individual differences in preferred word shapes and how their patterns evolve over time within a group of children with comparable language exposure.

14.7 Conclusion

This study is the first detailed investigation of Lebanese Arabic children's early word patterns with a focus on the transition that the child makes from the item-based production of the first few words towards more generalized learning and phonological systematicity. This is achieved both in the way children gradually move towards adult-like word shapes and segmental productions and in the way they form their own generalisations about word shapes and apply these to new incoming words so that, for a short time, their accuracy may decrease. The children in this study all produced many disyllabic word shapes with medial long consonants due to their frequency in the adult input. However, their individual preference for this pattern varied across sessions and between children, depending on the frequency with which they heard and produced other languages and on their individual preferences. Differences were also present in their segmental inventories and the degree to which they applied early developmental patterns such as consonant harmony. Despite the prevalence of onsets in the children's productions, syllables with heavy rhymes or codas were produced from an early age, and the children were more accurate in their production of medial than initial consonant position. Their data therefore adds to the growing number of studies on languages with quantitative contrasts that challenge the universal attention to initial consonants that is sometimes implied. Medial gemination was used by the children as an active process that enabled them to select words with a familiar rhythmic shape and to adapt other words to that shape. In the later stages of development, this was extended to multisyllabic word production.

Gemination has not received much attention in the literature on Arabic acquisition despite its high functional load and the discrepancy between the phonetic and phonological challenge involved in its acquisition. This study therefore constitutes a first step towards offering a detailed account of the acquisition of gemination in Lebanese Arabic. Current work is looking at the acoustic indices for gemination both in adult and child production in order to better understand the process by which children acquire the singleton-geminate contrast; data from later sessions are also being analysed in order to explore the influence of morphosyntax on the acquisition of this contrast.

References

Abdo E. (2011). *A study of the phonological structure and representation of first words in Arabic*. Unpublished PhD dissertation, University of Leicester.

Al-Tamimi, J. & Khattab, G. (2011). Multiple cues for the singleton-geminate contrast in Lebanese Arabic: Acoustic investigation of stops and fricatives. In W.S Lee & E. Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, 17-21 August, 2011. pp. 212-215.

Al-Tamimi, J. & Khattab, G. (submitted). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: the case of fricative consonants. *Journal of the Acoustical Society of America*.

Amayreh, M. & Dyson, A. (1998.) The Acquisition of Arabic Consonants. *Journal of Speech, Language and Hearing Research,* 41, 642-653.

Ammar, W. (2002). Acquisition of syllabic structure in Egyptian Colloquial Arabic. In F. Windsor, L.M. Kelly & N. Hewlett (eds.), *Investigations in Clinical Phonetics and Linguistics* (pp. 153-160). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Ammar, W. (1999). The acquisition of consonant clusters in Egyptian children from two to four years. *Language Sciences* 2(3), 10 – 37.

Ammar, W. & Morsi, R. (2006). Phonological Development and Disorders: Colloquial Egyptian Arabic. In Z. Hua & B. Dodd (eds.), *Phonological Development and Disorders in Children* (204-232). Clevedon: Multilingual Matters Limited.

Bhaya Nair, R. (1991). Monosyllabic English or disyllabic Hindi? Language acquisition in a bilingual child. *Indian Linguistics*, 52, 1-4.

Boersma, P. & Weenink, D. (2009). Praat: doing phonetics by computer (version 5.1.10) [Computer program]. Retrieved from http://www.praat.org/.

Boudelaa, S. & Marslen-Wilson, W.D. (2004). Abstract morphemes and lexical representation: the CV-Skeleton in Arabic, *Cognition* 92 (3), 271-303.

Boudelaa, S. & Marslen-Wilson, W.D. (2001). Morphological units in the Arabic lexicon, *Cognition* 81 (1), 65-92.

Broselow, E. 1992. Parametric variation in Arabic dialect phonology. In E. Broselow, M. Eid, & J. McCarthy (eds.), *Perspectives on Arabic Linguistics* (pp.7-45). Amsterdam, Philadelphia: John Benjamins.

Demuth, K. 1995. The prosodic structure of early words. In J. Morgan & K. Demuth (eds.), *From Signal to Syntax: bootstrapping from speech to grammar in early acquisition* (pp. 171-184). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Dyson, A. & Amayreh, M. (2000). Phonological errors and sound changes in Arabic speaking children. *Clinical Linguistics & Phonetics*, 14, 79-109.

Fee, J. (1995). Two strategies in the acquisition of syllable and word structure. In E.V. Clark (ed.), *Child Language Research Forum*, vol. 27 (pp. 29- 38). Stanford: Leland Standford Junio University.

Fikkert, P. (1994). O*n the Acquisition of Prosodic Structure*. PhD dissertation, University of Leiden.
Ferguson, C.A. & Farwell, C.B. (1975). Words and sounds in early language acquisition. *Language*, 51, 419-439.
Harris J. & Gussman, E. (1998). Final codas: why the west was wrong. In E. Cyran (ed.), *Structure and Interpretation: Studies in Phonology* (pp. 139-162). Lublin: Folium.
Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry,* 20, 253–306.
Khattab, G. (2007). Lebanese speech acquisition. In S. McLeod (ed.), *The International Guide to Speech Acquisition* (pp. 300-312). Clifton Park, NY: Thomson Delmar Learning.
Khattab, G. & Al-Tamimi, J. (submitted). Geminate timing in Lebanese Arabic. *Laboratory Phonology*.
Khattab, G. & Al-Tamimi, J. (2008). Durational cues for gemination in Lebanese Arabic. *Languages and Linguistics*, 22, 39-56.
Nasr, R.T. (1966). *Colloquial Arabic: An oral approach*. Beirut: Librarie du Liban.
Nasr, R.T. (1960). Phonemic length in Lebanese Arabic. *Phonetica,* 5, 209-211.
Obrecht, D.H. (1968). *Effects of the Second Formant on the Perception of Velarisation Consonants in Arabic*. Mouton: The Hague.
Macken, M.A. (1979). Developmental reorganization of phonology: A hierarchy of basic units of acquisition, *Lingua,* 49, 11-49.
Menn, L. (1983). Development of articulatory, phonetic, and phonological capabilities. In B. Butterworth (ed.), *Language Production*, vol. 2 (pp. 3-50). London: Academic Press.
Menn, L. (1971). Phonotactic rules in beginning speech. *Lingua*, 26, 225-251.
McCarthy, J. (1982). Prosodic templates, morphemic templates, and morphemic tiers. In H. van der Hulst & N. Smith (eds.), *The Structure of Phonological Representations, Part 1* (pp.190-223). Dordrecht: Foris.
McCarthy, J.J. & Prince, A. (1986). *Prosodic Morphology*. Ms, Massachusetts: University of Massachusetts, Amherst, and Brandeis University, Waltham.
McCarthy, J.J. and Prince, A. (1990a). Foot and word in prosodic morphology: the Arabic broken plural. *Linguistic Enquiry*, 8, 209-283.
McCarthy, JJ. & Prince, A. (1990). Prosodic morphology and templatic morphology. In M. Eid & J. McCarthy (eds.), *Perspectives on Arabic Linguistics: Papers from the Second Symposium* (pp. 1–54). Amsterdam: John Benjamins.
Ota, M. (this volume). Lexical frequency effects on phonological development: the case of word production in Japanese.
Peters, A.M. (2001). Filler syllables: What is their status in emerging grammar? *Journal of Child Language* 28, 229-242.
Priestly, T.M.S. (1977). One idiosyncratic strategy in the acquisition of phonology, *Journal of Child Language,* 4, 45-66.
Ravid, D. (2002). A developmental perspective on root perception in Hebrew and Palestinian Arabic, In J.Y Shimron (ed.), *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 293–319). Amsterdam: Benjamins.
Rose, Y. & Wauquier-Gravelines S. (2007). Acquisition of speech in French, in Mc Leod S. (ed), *International guide of Speech Acquisition* (pp. 364-385). Thomson Delmar Learning, USA.
Rose, Y. (2012). Phon (version 1.5.2). [Computer program]. Retrieved from http://phon.ling.mun.ca/phontrac/wiki/Downloads.
Savinainen-Makkonen, T. (2007). Geminate template: A model for first Finnish words. *First Language*, 27, 347-359.

Saleh, M., Shoeib, R., Hegazi, M. & Pakinam, A. (2007). Early phonological Development in Arabic Egyptian Children: 12–30 Months. *Folia Phoniatrica et Logopaedica*, 59: 234-240.

Salem, H. (2000). *Study of the acquisition of the syllable structure in sentence perspective in the speech of normal Egyptian children*. Unpublished PhD thesis, University of Alexandria.

Shahin, K. (2003). *Prosody-segmentism in the acquisition of Arabic: word-final onsets and no stress effects*. University of British Columbia Child Phonology Conference. Vancouver, 1-4 July, 2003.

Shahin, K. (1995). Child Language evidence on Palestinian Arabic phonology. In Eve Clark (ed.), *Proceedings of the 26th Child Language Research Forum* (pp. 104-116). Stanford, CSLI.

Stoel-Gammon, C. (1987). Phonological skills of 2-year-olds. *Language, Speech, and Hearing Services in Schools*, 18, 323-329.

Stoel-Gammon, C. & Cooper, J. (1984). Patterns of early lexical and phonological development. *Journal of Child Language*, 11, 247-271.

Vihman, M.M. (1993). Variable paths to early word production. *Journal of Phonetics*, 21, 61-82.

Vihman, M.M. (1978). Consonant Harmony: Its Scope and Function in Child Language. In J.H. Greenberg, (ed.), *Universals of Human Language 2: Phonology* (pp. 281-334). Stanford: Stanford University Press.

Vihman, M.M. & Croft, W. (2007). Phonological development: Toward a 'radical' templatic phonology. *Linguistics*, 45, 683-725.

Vihman, M.M. & Velleman, S.L. & McCune, L. (1994). How abstract is child phonology? Towards an integration of linguistic and psychological approaches. In M. Yavas (ed.), *First and Second Language Phonology*. San Diego: Singular Publishing.

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

Vihman, M.M. & Velleman, S.L. (1989). Phonological reorganization: a case study. *Language and Speech*, 32, 149-170.

Vihman, M.M., Ferguson C.A. and Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, 7, 3-40.

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

Watson, J. (2002). *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.

Wauquier, S. & Yamaguchi, N. (this volume). Templates in French.