

# THE ACQUISITION OF GEMINATION IN LEBANESE-ARABIC CHILDREN

Ghada Khattab and Jalal Al-Tamimi

Speech and Language Sciences Section, Newcastle University  
[ghada.khattab@newcastle.ac.uk](mailto:ghada.khattab@newcastle.ac.uk), [jalal.al-tamimi@newcastle.ac.uk](mailto:jalal.al-tamimi@newcastle.ac.uk)

## ABSTRACT

This is the first study on the acquisition of gemination in Arabic, a phonological aspect that is prominent in the adult phonology yet complex in terms of its implementation and interaction with the grammar. The study reports on the longitudinal development of five Lebanese children in the second year of life and enables the authors to trace the transition from phonetic to phonological acquisition in the child.

An acoustic investigation of consonant duration in the children's emerging lexicon shows very little distinction between short and long targets in the early stages, followed by rapid word learning and the beginnings of target-like durational patterns. However, the language-contact situation in Lebanon creates an interesting challenge for geminate acquisition; variable phonetic lengthening of medial consonants in French and English loan words in the input creates a fuzzy category which the children need to interpret alongside their unstable representations for Arabic.

**Keywords:** gemination, child phonology, Lebanese Arabic, phonetics-phonology interface.

## 1. INTRODUCTION

The acquisition of gemination (contrastive consonant length, e.g. /bana/ *he built* vs /ban:a/ *barista*) has occupied very little of the child phonology literature despite the major contribution it can make to the understanding of the relationship between phonetic and phonological acquisition. Cross-linguistically, children are typically reported to start with relatively slow moving articulators, leading to the production of phonetically long consonants (and vowels) regardless of whether or not length is phonological in the target language [11, 13, 16]. It is therefore interesting that geminates are considered marked consonants [4] - which in the child language literature translates as *acquired late* - given that their hallmark phonetic realisation can be found naturally in children's early speech. The key of course lies in understanding a) the prosodic patterns of gemination, i.e. how consonant duration interacts with that of surrounding vowels and syllables, and b) the child's ability to make a

phonetic as well as a phonological distinction between short and long consonants.

These issues have only been investigated in a small number of studies [1, 11, 13, 16, 17]; their results suggest that, alongside the need for a growing awareness of lexical (and grammatical) contrast, children's motoric ability needs to develop in order to enable them to shorten short targets, and if needed, lengthen long ones (depending on the language-specific duration patterns for long consonants). Below we review what we know about these emerging phonetic and phonological abilities in languages where the acquisition of gemination has been documented, before we turn the focus to Arabic. The influence of language contact in Lebanese Arabic is then used to explore the many facets of geminate acquisition.

## 2. THE ACQUISITION OF GEMINATION

A small number of studies have looked at the acquisition of medial gemination in languages such as Finnish and Japanese, and compared these with the acquisition of languages which make no use of contrastive consonant length on the one hand (e.g. French and English), and those where long consonants are prosodically conditioned (e.g. in Welsh, where medial consonants are long in trochaic disyllables) [1, 13]. The three-way comparison is interesting because it allows one to disentangle lexical from prosodic effects.

What these studies show is that in the initial stages of speech acquisition (around age 1;0), children typically have a small vocabulary, and for those who are acquiring a language with gemination, the consonantal make up of their early lexicon may not contain sufficient instances of phonologically short or long consonants for them to develop implicit awareness of the contrast. Children at this age also have little motoric control for durational contrasts. As a result of that, medial consonant duration patterns initially look very similar across children from different language backgrounds.

A few months later, as their motoric and linguistic abilities develop, children acquiring French and English exhibit shortening of their consonants, which brings them more in line with the adult targets. Crucially, however, children acquiring Welsh do not exhibit such shortening, which shows

that this phenomenon is not an automatic process that is linked to motor control, but one that is also connected with the development of linguistic ability and more targeted articulation. This becomes more obvious in children acquiring languages with gemination, as they typically show a unique pattern of both shortening of short targets and lengthening of long ones.

Children at this young age are still generally far from exhibiting adult-like durational patterns, but language-specific influences are already apparent. These relate to aspects of the phonology which may contribute to phonological length being more salient/noticeable to the child, and therefore acquired earlier. These include the number of consonants that can be geminated (salience of lexical contrast), the ratio of geminate to singleton duration (acoustic salience), whether vowels length is also phonological (salience of length as a phonological contrast), and whether gemination is important beyond the lexicon (grammatical salience). Within existing results, children acquiring Finnish have been found to have an advantage over children acquiring Japanese due to the majority of the factors above. In the next section, we explore how Arabic fares on these factors before we move on to the first set of data on the early acquisition of gemination in Lebanese Arabic.

## 2. GEMINATION IN ARABIC

Gemination in Arabic has lexical (e.g. /ħame:m/ *pigeons* vs /ħam:e:m/ *bathroom*) and morphosyntactic functions (e.g. /katab/ *he wrote* vs /kat:ab/ *he made someone write*). All 27 consonants in Arabic can be geminated, and vowel length is also contrastive [7, 8, 10]. This makes for a salient phonological feature in the language, and might predict early acquisition by Arabic-speaking children. However, there has been very little attention to this aspect of phonological acquisition in Arabic, with some studies making only a brief mention of early acquisition of gemination [2]. Other studies, however, point out that adult-like geminate patterns appear late, with Egyptian children reported to be only 80% target-like by age 6 [3]. None of these studies, however, employed acoustic analysis in their investigations.

Lebanese Arabic has been in contact with French since colonial times in the early 1900s and more recently English. While the two languages do not have gemination, the varieties spoken in Lebanon have been reported to exhibit long medial consonant durations in linguistic environments which are still poorly understood; while factors such as double letters in the spelling (e.g. English *apple*, French

'attend' *wait*) and iambic stress (French 'chateau' *castle*) seem to play a role, the patterns are by no means categorical and require further investigation.

## 3. CURRENT STUDY

### 2.1. Participants

Five children from Beirut were recruited for this study. Inclusion criteria were for the participants to be around age 1 (with a 3-months range), to have Lebanese Arabic-speaking parents who speak mainly Arabic at home (as determined by a language background questionnaire), and to have no reported history of hearing disorders. The children's language development was followed throughout the one-word stage (before children typically start producing sentences). This consisted of monthly 30-minute audio-visual recordings of mother-child spontaneous interactions starting when the child had a small productive vocabulary (referred to in the literature as the 4 word-point [14]), and ending around 6-months later, when the child's vocabulary had progressed to around 50 words (referred to as the 25 word-point [14]); this normally coincides with the period when children make the transition to longer utterances.

Sampling at this early age allows the researcher to capture the earliest emergence of a phonological length contrast, if it is to be manifested at this early stage; the challenge is the relatively small dataset that one needs to contend with, especially at the 4 word-point, and having to rely solely on naturalistic interactions to extract comparable tokens of interest. The children wore custom-made vests with Sennheiser radio mics hidden in a pocket, and the mothers wore separate mics. Both mics were connected to an Edirol R-9 recorder with a sampling frequency of 44.1 kHz, 16 bit quantisation in stereo.

### 2.2. Data and analyses

Disyllables with target medial short and long consonants and preceding short or long vowels were extracted from the children's and the mothers' speech (in order to obtain data on adult durations in child-directed-speech). Disyllables were chosen due their frequency in Arabic children's early words and the productivity of the medial singleton-geminate contrast. The data were phonetically transcribed using PHON [12] and a word identification procedure was followed in order to separate babbling from early words [14]. Durational measurements of the medial consonant and the surrounding vowels were made in PRAAT [5], both to look at proportional duration and to examine the interaction between phonologically short and long vowels and consonants [8]. A total of 1291 tokens

across sessions were acoustically analysed. A two-way ANOVA with stage (4- vs 25-word point) and consonant type (singleton vs geminate) was run. Each ANOVA was followed by Bonferroni post-hoc analyses and Cohen's *d* effect size measures to evaluate how robust the results are in distinguishing singleton vs geminate categories [6].

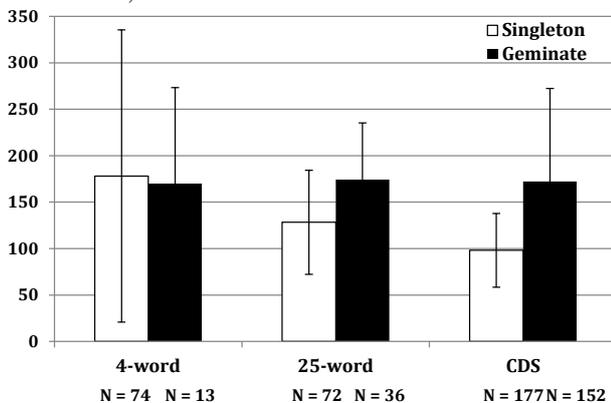
### 2.3. Results

#### 2.3.1. Durational patterns for Arabic medial singleton and geminate consonants

Figure 1 below shows durational measurements for medial singleton and geminate stops and nasals as produced by the children at the beginning and the end of the single-word stage (in keeping with existing literature, these are referred to as 4-word and 25-word). At the 4-word point, and when all vowel contexts are combined, the children's small lexicon shows no durational distinction between singleton and geminate consonants, with most durations falling within the geminate range of the adults' productions in child-directed-speech (CDS). The ranges are also wide for both targets, especially in the singleton context, suggesting a lack of motoric control over durations.

A contrast emerges at the 25-word point, with a significant difference between singleton and geminate consonants ( $p < 0.05$ ,  $d = 0.8$ ), but the durations for the singleton target are still relatively high when compared with those in CDS, leading to a smaller distinction between singleton and geminate targets for children than for adults.

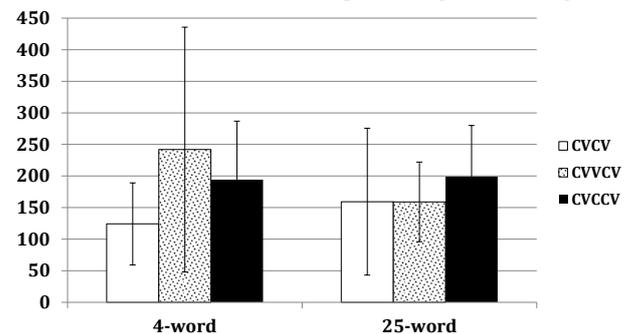
**Figure 1:** Absolute duration (in ms) of medial singleton and geminate consonants at the beginning and end of the single-word stage by the children, with the adult durations in CDS.



The noticeable change in the children's durational patterns at the end of the single-word stage coincides with a period of significant lexical development by the children, both in type and token frequencies [8].

A closer look at the vowel preceding singleton and geminate consonants (Fig. 2) reveals that the children mostly struggle with producing singletons when the preceding vowel is long (e.g. /ba:ba/ daddy), frequently choosing to shift the length onto the consonant ([bab:a]). This phenomenon reduces considerably at the 25 word-point, showing maturation in the production of both vowel and consonant durations<sup>1</sup>. When the preceding vowel is short, singleton consonant production is more target-like even at the 4-word point, with the wide variation reducing at the 25-word point.

**Figure 2:** Absolute duration (in ms) of medial consonants as a function of preceding vowel length



#### 2.3.2. Durational patterns for French and English loanwords

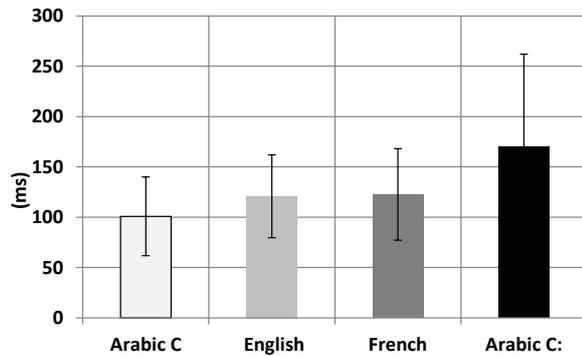
Given that code-switching and loan words are frequent in Lebanese Arabic, we further looked at the children's production of disyllabic French and English words in order to examine their medial consonant duration. This was prompted by an observation of variable phonetic lengthening of these consonants in the adults' speech, which required further investigation in order to establish whether the durational patterns fell within those of the geminate consonants in Arabic and what consequences this had for the child's interpretation of phonetic and phonological lengthening.

Table 1 shows examples of French and English words produced with noticeably long medial consonants by the adults. The actual durations vary a great deal within and across individuals, creating a medium category along the short-long durational scale (Fig.3). This is also reflected in the children's productions (Table 1 and Fig. 4), which fluctuate between the short and the long durational ranges; the overall mean for French (but not English) is significantly different from each of the singleton ( $p < 0.0001$ ,  $d = 0.5$ ) and geminate ( $p < 0.0001$ ,  $d = 0.5$ ) patterns at the 25-word point, suggesting that French medial target consonants are treated differently from Arabic singleton and geminate counterparts.

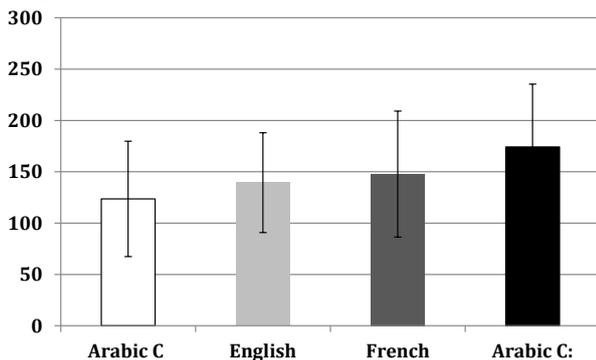
**Table 1:** Sample French and English words produced by Lebanese adults and children, with their impressionistic and acoustic profile.

Adults (CDS)	IPA target	Adult form	Medial C duration
French	papa	pəp:æ <sup>h</sup>	151
	ʃapo	ʃəp:ɔ:h	182
English	heloo	hæ:l:o:h	168
	kiti	kit:hi:h	160
Arabic mean C duration = 99; CC = 168			
Children	IPA target	Adult form	Medial C duration
French	bebe	həbɛ	117
	moto	tɛ:t:ɔ:h	521
English	beibi	pəp:ɛi:h	118
	kiti	kit:hi:h	214

**Figure 3:** Mean durations (in ms) for medial C (and CC) for Arabic, English and French words as produced by the adults in CDS.



**Figure 4:** Mean durations (in ms) for medial C (and CC) for Arabic, English and French words as produced by the children at the 25-word point.

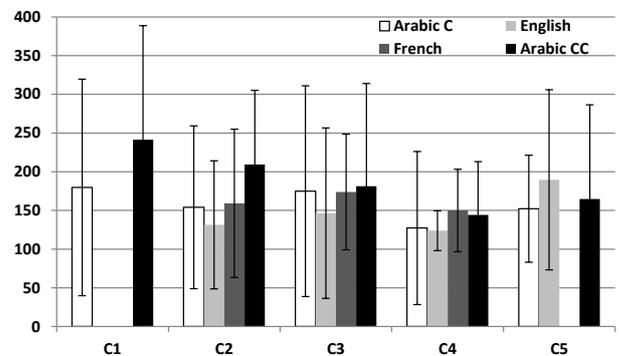


### 2.3.3. Individual differences in the development of a durational contrast

While generalisations are hard to make from a necessarily small dataset (longitudinal studies are very labour-intensive, especially when the analyses consist of transcriptions and acoustic measurements of child speech), the individual journeys are often as informative as group results, if not more so. In this study, individual differences are present throughout

the children's development (Fig. 5), with some (C1 and C2) already exhibiting a durational contrast between their singleton and geminate consonants at the 25-word point, while others showing no distinction (C3 to C5). Here, there is a tendency for children who mostly produce phonetically long French or English medial consonants to show no distinction between their singleton and geminate consonants in Arabic (C3 and C5), while C1, who produces no French or English, shows the sharpest contrast. While it is hard to generalise these results given the small sample, they raise interesting theoretical questions which are discussed in the next section.

**Figure 5:** Individual differences in the development of the singleton-geminate contrast.



## 4. DISCUSSION AND CONCLUSION

This study looked at the earliest signs of acquisition of the singleton geminate contrast by Lebanese Arabic-speaking children. The acquisition of gemination was shown to be quite complex, with both phonetic control and phonological representation required of the child before they can approximate adult targets. Children initially produce long consonants across the board, as has been shown in studies on other languages [11, 13, 16, 17], and over time gain enough control to shorten the consonants for short targets. This takes place with practice and a large enough vocabulary which draws the child's attention to the phonological contrast.

Lebanese Arabic children are also exposed to phonetically long consonants in French and English loanwords. The variability in the realisation of these consonants and the children's already unstable phonetic and phonological representations pave the way for individual differences; these manifest themselves as an apparent delay in the acquisition in the acquisition of the contrast by some children or a re-categorisation of the consonants in the loan words as long by others.

## 5. ACKNOWLEDGEMENTS

Many thanks to the reviewers for their comments on earlier versions of this paper. This work was supported by an ESRC First Grants RES-061-23-0053.

## 6. REFERENCES

- [1] Aoyama, K. 2001. *A Psycholinguistic Perspective on Finnish and Japanese Prosody*. PhD dissertation, University of Hawai'i.
- [2] Abdoh E. 2011. *A study of the phonological structure and representation of first words in Arabic*. PhD dissertation, University of Leicester.
- [3] Ammar, W., Morsi, R. 2006. Phonological development and disorders: Colloquial Egyptian Arabic. In Z. Hua, B. Dodd (eds.), *Phonological development and disorders in children*, pp. 204–32. Clevedon: Multilingual Matters.
- [4] Blevins, Juliette. 2004. *Evolutionary phonology: the emergence of sound patterns*. Cambridge: CUP.
- [5] Boersma, Paul & Weenink, David. 2009. *Praat: Doing Phonetics by Computer* (Version 5.1.0) [Computer program]. Retrieved July 22, 2009, from <http://www.praat.org/>.
- [6] Cohen, Jack. 1988. *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates
- [7] Khattab, G. 2007. Lebanese speech acquisition. In S. McLeod (ed.), *The international guide to speech acquisition*, pp. 300–12. Clifton Park, NY: Thomson Delmar Learning
- [8] Khattab, G., Al-Tamimi, J. 2014. Geminate timing in Lebanese Arabic: The relationship between phonetics and phonology. *Laboratory Phonology*, 5, 231 – 269.
- [9] Khattab, G., Al-Tamimi, J. 2013. The role of geminates in shaping early word patterns by Lebanese-Arabic speaking children. In M.M. Vihman & T. Keren-Portnoy (eds.) *Readings on Phonological Development: Templatic Approaches, Cross-Linguistic Data*. Cambridge: CUP.
- [10] Khattab, G., Al-Tamimi, J. 2008. Durational cues for gemination in Lebanese Arabic. *Languages and Linguistics*, 22, 39–56.
- [11] Kunnari, S., Nakai, S., Vihman, M. M. 2001. Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication*, 5, 13-24.
- [12] Rose, Y. (2012). Phon (version 1.5.2). [Computer program]. Retrieved from <http://phon.ling.mun.ca/phontrac/wiki/Downloads>.
- [13] Vihman, M.M., Kunnari, S. 2006. The sources of phonological knowledge: A cross-linguistic perspective. *Recherches linguistiques de Vincennes*, 35, 133-164.
- [14] Vihman, M. M., McCune L. 1994. When is a word a word? *Journal of Child Language*, 21, 517-542.
- [16] Vihman, M. M., Nakai, S., DePaolis, R. 2006. Getting the rhythm right. In Goldstein et al., eds. *Laboratory Phonology* 8, 341-366. De Gruyter.
- [17] Vihman, M. M., Velleman, S. 2000. Phonetics and the origins of phonology. In Burton-Roberts et al., eds., *Phonological knowledge*, 305-339. OUP.

---

<sup>i</sup> Long vowels can also precede geminate consonants in Arabic (e.g. /ʕaːdːe/ *having counted*, 3<sup>rd</sup> pers.fem.), but the children never targeted these in their production, most probably due to their complex morpho-syntactic structure.