The production and perception of peripheral geminate/singleton coronal stop contrasts in Arabic

Mohamed Yassine Frej

A dissertation submitted for the degree of Doctor of Philosophy

The MARCS Institute for Brain, Behaviour and Development

Western Sydney University, Australia

2021

Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

…….. (Signature

ACKNOWLEDGEMENTS

An-Nu'maan ibn Al-Basheer رضي الله عنه said: "*He who does not thank Allah for small favors will not thank Him for great ones; and he who does not express gratitude to people (for doing him favors) will not express gratitude towards Allah*." [Ahmad]

In the name of Allah S.W.T, the most Merciful, the most Gracious. We praise Him, Seek his help, and ask for His forgiveness. I am thankful to Allah who supplied me with the courage, the guidance, and the love to complete my journey as a PhD student.

Indeed, the fruition of this PhD journey involved lots of hardship, bitterness and inconvenience, but I am only walking away with sweet memories, life-changing lessons and everlasting friendships. The completion of this thesis would definitely not have seen light without the support, help and encouragement of several people including my supervisors, colleagues, family and friends. I believe "Thank you" is not enough as a reward, but it is infused with all the love and reverence I have for all of you.

I am very obliged to Prof Catherine Best, my academic supervisor. I really fall short stating your countless favors for me. Your guidance and lavish suggestions represent the backbone of this research. During the course of my PhD, there were many ups and downs, but your patience and support remained unflinching. One could not have asked for a better supervisor. I am also grateful to my associate supervisor – Dr Michael Tyler for all the time

you spared to review and comment on my work. Your valuable knowledge and advice and your keen attention to detail helped me a great deal improve this thesis. To Dr Christopher Carignan, I am especially indebted to you for all the skills you have taught me. This thesis would not have come to fruition without your guidance, encouragement and knowledge.

My time at the MARCS was made enjoyable in large part due to the many friends and colleagues. I feel very fortunate to have been surrounded by very kind and compassionate colleagues whose favors cannot be described in words. Through their care and help, they acted like my second family. They include, but not limited to, Muneeb Ahmad, Sarah Fenwick, Irena Lovcevic, Peta Mills, Gloria Maria Pino Escobar, Sarah Wright, Jia Ying, Jenny Yu and Valeria Peretokina. Particular thanks go to my Lab colleagues Juqiang Chen, Yanping Li and Tina Whyte-Ball for their suggestions and encouragement.

I would like to extend a huge thank you to my dear friends back home: Zineb Zerouali, Abdelali Elfakir, Amine Ammor, Youssef Belmiloud, Thami Lahbi, Omar Tounsi and many others. Despite the distance that separates us, you were always there for me during the happy and hard moments. I always cherish you as my best friends.

Most importantly, I must convey my deepest love and gratitude to my family. My mother Saadia, My father Salah Eddine, my sister Jihane and my two brothers Abdellah and Youssef. Their inexhaustible loving support and patience and their unsparing prayers were the genuine fuel of my academic pursuit. I would like to thank my uncles, aunts and other members of my extended family who are always excited to hear about my achievements and who inspire me to be more successful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Gemination is typologically common word-medially but is rare at the periphery of the word (word-initially and -finally). In line with this observation, prior research on production and perception of gemination has focused primarily on medial gemination. Much less is known about the production and perception of peripheral gemination. This PhD thesis reports on comprehensive articulatory, acoustic and perceptual investigations of geminate-singleton contrasts according to the position of the contrast in the word and in the utterance. The production component of the project investigated the articulatory and acoustic features of medial and peripheral gemination of voiced and voiceless coronal stops in Modern standard Arabic and regional Arabic vernacular dialects, as produced by speakers from two disparate and geographically distant countries, Morocco and Lebanon. The perceptual experiment investigated how standard and dialectal Arabic gemination contrasts in each word position were categorised and discriminated by three groups of non-native listeners, each differing in their native language experience with gemination at different word positions.

The first experiment used ultrasound and acoustic recordings to address the extent to which word-initial gemination in Moroccan and Lebanese dialectal Arabic is maintained, as well as the articulatory and acoustic variability of the contrast according to the position of the gemination contrast in the utterance (initial vs. medial) and between the two dialects. Kinematic analysis of ultrasound data related to dynamic lingual articulation indicated that speakers of the Moroccan and Lebanese dialects maintained word-initial gemination contrast articulatorily even in the utterance-initial position, where closure of voiceless stops is inaudible and its duration cannot be identified in the acoustic recordings. Utterance-medially, articulatory and acoustic closure duration played a robust role in contrasting geminates with singletons.

Utterance-initially, rather than using closure duration, speakers differentiated the contrast through both durational and non-durational acoustic cues that are associated with the forcefulness of articulation. This finding challenges the common view that gemination is primarily a distinction in closure duration.

The second experiment compared the production of word-medial and -final gemination in Modern Standard Arabic as produced by Moroccan and Lebanese speakers. The main finding was that final gemination was contrastive and maintained articulatorily and acoustically by both groups of speakers. Similar to word-initial gemination in utterance-medial position, word-medial gemination was realised primarily through differences in closure duration. Conversely, non-closure duration acoustic cues were particularly enhanced to signal word-final gemination. Differences between the two dialects with regard to the realisation of these acoustic cues also increased in word-final position.

The aim of the perceptual experiment was to disentangle the contribution of phonological and phonetic effects of the listeners' native languages on the categorisation and discrimination of non-lexical Moroccan gemination by three groups of non-native listeners varying in their phonological (native Lebanese group and heritage Lebanese group, for whom Moroccan is unintelligible, i.e., non-native language) and phonetic-only (native English group) experience with gemination across the three word positions. The categorisation and discrimination results showed that sensitivity to the gemination contrast was consistent with the listeners' native-language experience. Phonological experience with the contrast facilitated in more accurate discrimination than phonetic-only experience, but the benefit was largely restricted to the word positions where gemination occurs in the native language of the listener's group. More nuanced differences between the groups in discrimination accuracy were accounted for by perceptual assimilation to native phonological categories. Moreover, reduced

acoustic cues, dialectal differences, language use patterns and voicing were all important factors affecting the categorisation and discrimination of non-native gemination contrasts. These results highlight the importance of considering low-level phonetic alongside phonological effects on non-native perception of gemination.

The findings in this thesis constitute important contributions about positional and dialectal effects on the production and perception of gemination contrasts, going beyond medial gemination (which was mainly included as control) and illuminating in particular the typologically rare peripheral gemination.

CHAPTER 1: OVERVIEW

The vast majority of research on consonant gemination, i.e., a consonant length distinction, has been conducted on word-medial gemination with little attention to peripheral gemination (word-initial and -final). Because peripheral gemination lines up with the edges of one or more prosodic domains, there is compelling phonological, prosodic and perceptual evidence that it differs substantially and in multiple ways from medial gemination. Despite the importance of understanding how these factors impact on gemination, peripheral gemination is heavily understudied. Therefore, this PhD thesis takes up a comprehensive articulatory, acoustic and perceptual investigation of gemination in two distinct varieties of standard Arabic and its vernacular dialects, Moroccan and Lebanese, in order to determine the various regional similarities and differences between the geminate/singleton contrast across word utterance positions, particularly peripheral positions in relation to medial positions. Knowing the acoustic and articulatory similarities and differences between the regional varieties is crucial to interpreting the results of the perception experiments using the Moroccan acoustic materials with naïve English versus heritage and native Lebanese listeners.

Gemination or the geminate/singleton contrast, terms used interchangeably throughout this thesis, is at one level a *phonetic* phenomenon whereby a geminated, or long, consonant is produced with an elongated duration relative to its singleton, or short, counterpart. At the phonological level, gemination contrasts can also serve to distinguish between minimal-pair words in some languages, as it does in Arabic and many of its regional dialects. For example,

the Arabic word /ħamaːm/ with a short medial /m/ means “pigeons,” while the word /ħamːaːm/, with a long medial /m/, means “bathroom.” The difference in phonetic durations and phonological lengths of the word-medial consonant in the two otherwise identical words are what gives them different meanings. Thus, consonantal duration differentiates words contrastively in Arabic.

The catalyst for this PhD project was my experience with teaching Modern Standard Arabic (MSA) and Moroccan Arabic dialect (Moroccan) to second language learners. Serving as an Arabic teaching assistant at the University of California, Davis, in the United States gave me a first-hand exposure to the myriad of challenges English L1 speakers encounter when learning Arabic. I noticed that one of those challenges is with learning consonant gemination, particularly word-finally, which is crucial for differentiating certain otherwise identical Arabic word pairs such as /amal/ أمل “hope” and /amalː/ أملّ “more boring”. Prior to this, I had the opportunity to teach Moroccan dialect to Korean cultural volunteers in Morocco. Once again, the acquisition of gemination was not an easy task for Korean learners of Moroccan. The challenge this time was word-initial gemination, which is allowed only in some local vernacular dialects of Arabic such as Moroccan and Lebanese. Surprisingly, neither American nor Korean learners found word-medial gemination as challenging as word-final gemination in MSA or word-initial gemination in Moroccan. This stirred in me a thirst for knowledge about gemination in Arabic in general and its perception by L2 learners in particular. My starting question was: what makes gemination at the edges of the word, henceforth referred to as peripheral gemination, particularly difficult for L2 learners to perceive?

One plausible explanation is that gemination is non-contrastive in languages such as the native languages of the learners I taught: English and Korean. There are no pairs of words in either of those languages that contrast merely due to gemination. There is a general

agreement among theories of L2 perception, e.g., the Perceptual Assimilation Model (PAM; Best, 1995) and the Speech Learning Model (SLM; Flege, 1995), that L2 listeners find it challenging to discriminate between phonemes in an L2 that are not contrastive in the listener's L1. Therefore, this lack of a phonemic length contrast in English and Korean might make it relatively difficult for native speakers of those languages to discriminate the geminate/singleton contrast in geminating languages such as Arabic. However, this cannot explain why English and Korean speakers apparently find it less challenging to perceive gemination at the word-medial position than at peripheral positions. English native speakers *are* able to detect durational differences in vowels or in consonants in word-medial position in other geminating languages (Dmitrieva, 2012; Kraehenmann, 2001). This, in return could explain why my American and Korean students experienced less challenge with word-medial gemination in MSA and Moroccan, respectively.

Another possible, more direct explanation for why American and Korean speakers are more sensitive to consonantal duration contrasts in medial than in utterance- final and/or utterance-initial position is that consonant lengthening occurs phonetically or pragmatically in medial but not in peripheral positions in their native languages. Despite lacking a phonological contrast for consonantal duration, in these languages consonant lengthening does occur across word boundaries when the same consonant appears at the end of one word and at the beginning of the next word. For instance, in English <top pick> differs contrastively from <topic> (Hayes, 2001; Pickett & Decker, 1960). This type of gemination has been called "fake gemination" (Oh & Redford, 2012), because the geminated segment is phonetically long, but it is spread across two syllables falling on either side of a word or morpheme boundary and does not contrast with otherwise identical cases that differ only in having a shorter segment in that position. Therefore, it does not engage in a phonemic length contrast within any lexical items. The key thing to note is that the context in which fake gemination arises in English is

between words or morphemes, phonetically akin to medial gemination in languages such as Arabic. Peripheral gemination, however, does not occur in languages such as English and Korean, as even this type of purely phonetic fake gemination cannot occur in initial or final position of an utterance.

In comparison with gemination, peripheral gemination presents less durational information in the acoustic signal, especially for voiceless stops in utterance-initial position. Given that the onset of voiceless stop closure cannot be ascertained from acoustic information in absolute initial cases, detecting the differentiation of closure duration is certainly problematic. Thus, it is fairly clear why peripheral gemination poses a challenge for L2 speakers in any language where it is contrastive.

Indeed, an important question that must be asked is whether peripheral gemination is difficult even for speakers of languages that do employ peripheral gemination in lexical distinctions to perceive. if speakers of languages that use peripheral gemination themselves find it challenging, we should not expect speakers of non-geminating languages to successfully perceive the geminate/singleton contrast occurring at the periphery of words in Arabic or its dialects.

Non-native peripheral gemination contrasts may indeed be difficult for even native listeners of geminating languages. Among the reasons this may be found is that there is no consensus among Arabic scholars over the phonetic maintenance versus neutralisation of initial and final gemination in Arabic and its dialects. Unless a comprehensive acoustic and articulatory investigation of this contrast in peripheral positions is conducted, we cannot rule out the possibility that both native and non-native listeners find peripheral gemination difficult to perceive due to the contrast being neutralised in these positions. Thus, an investigation of the production of peripheral gemination in Arabic and its regional dialects was needed to

determine whether or not the contrast is phonetically maintained in these positions. This thesis undertook such an investigation with two geographically-distant regional varieties of MSA (Experiment 1) and their corresponding regional dialects. The acoustic recordings from that study also provided the stimuli for the experiment on the perception of gemination by non-native listeners with non-geminating versus geminating native languages.

A detailed review of relevant literature, as well as the development of the rationale and design of the research program of this dissertation project are provided in the following three chapters. The full thesis is organised as follows: Chapter 2 provides a general background for this research by introducing gemination as a phonetic and phonological phenomenon, and by describing its different types and distributions across languages. It also reviews relevant acoustic and articulatory research on the production of gemination. I especially show how research focuses mostly on word-medial gemination, with little attention to gemination occurring at word-periphery Chapters 3 introduces the Arabic language and its dialects in order to familiarise the reader with the phonetic and phonological properties and differences between MSA and dialectal Arabic as well as how gemination in particular is formulated in these varieties. I also explore research on non-native perception of gemination especially studies involving the current project's target listener groups. Chapter 4 describes the project and its objectives, and its design based on material covered in the literature review, which resulted in two production experiments and one cross-language perception experiment.

Two production experiments and one perception experiment comprise the core of this research project. Chapter 6 presents the first production experiment on initial gemination in Moroccan and Lebanese regional vernacular dialects. Chapter 7 presents the second production experiment on medial and final gemination in Moroccan and Lebanese varieties of Modern Standard Arabic (MSA). Chapter 7 concerns the non-native perception experiment of word-

initial, -medial and -final gemination in Moroccan dialect and MSA by Arabic-naïve Australian English speakers, native speakers of Lebanese regional dialect and their variety of MSA, and heritage speakers of Lebanese dialect born in Australia (Lebanese dialect was their L1, before they learned English). Chapter 8 outlines the general discussion of the findings in relation to the existing literature and theoretical issues raised in the introductory chapters.

CHAPTER 2: GEMINATION

This chapter provides an in-depth literary review of the phenomenon of gemination. It is composed of five main sections. In the first section, the phenomenon of gemination is defined and explained (2.1). Different types of gemination (2.2) and its positional distribution across languages (2.3) are also reviewed and discussed. The fourth section is dedicated to the production of gemination by native speakers (2.4). The main findings of the chapter are summed up in the final section (2.5).

2.1. Terminology

Scholars have referred to gemination using a range of terms, such as consonantal length (Hankamer, Lahiri & Koreman, 1989), geminacy (Kawahara, 2005), length distinction (Abramson, 1986), and consonant elongation (Ahmed, 2016). Throughout this thesis, the term *gemination* will be used interchangeably with the term *geminate/singleton contrast*. There have also been various definitions of the term gemination that have been put forward in the literature. For example, Carr defined it as "a process whereby a single, non-geminate, consonant undergoes lengthening to become a geminate consonant" (2008: 62). For Church, “Gemination is a term from phonetics for the combination of two short adjacent consonants into one long one” (2012, p. 12). These definitions are incomplete in the sense that they do not recognise the influence of both phonetics and phonology on gemination in a particular language. To clarify, a geminate, or long, consonant is not only distinct from a singleton, or short, consonant *phonetically* through various phonetic cues such as consonant closure and/or surrounding vowel duration, which can occur in languages where the differences do not yield any lexical

contrasts between words. It is also a *phonological* contrast in languages where it does engage in lexical distinctions. Therefore, a more comprehensive definition of the term gemination would be: Gemination is both a phonetic phenomenon whereby the duration of a geminate consonant is increased relative to its singleton (i.e., non-geminated) counterpart, and a phonological phenomenon whereby the length difference is phonologically contrastive, i.e., distinguishes between lexical items. Thus, a clear distinction should be between phonologically contrastive gemination and phonetic lengthening in languages in which gemination is not phonologically contrastive.

In the next section, I introduce the functional and positional typologies of gemination as well as their relative occurrence across languages.

2.2. Gemination typology

Gemination can be divided into two main types: Lexical and Morphological gemination. The difference between these two types is explained, and their distribution across languages is also discussed in the next subsections.

2.2.1. Lexical gemination

Gemination is lexically contrastive in a number of languages: Bengali (Lahiri and Hankamer, 1988), Berber (Ridouane, 2010), Bernese (Ham, 2013), Buginese (Cohn, Ham & Podesva, 1999), Estonian (Engstrand & Krull, 1994), Finnish (Engstrand & Krull, 1994), Cypriot Greek (Tserdanelis & Arvaniti, 2001), Guinaang Bontok (Aoyama & Reid, 2006), Hindi (Ohala, 2007), Hungarian (Ham, 2013), and Italian (Payne, 2005; Pickett et al., 1999). In these languages gemination occurs within word boundaries. This is also referred to as “true” gemination, because it represents a distinctive feature in the phonology of a language. That is,

the short and long versions of a consonant give different meanings to otherwise identical words. For instance, the phonological length distinction between the medial consonants in /da**r**asa/ 'he studied' versus /da**r**:sa/[1] 'he taught' is what gives them distinct meanings in MSA. Thus, Arabic allows lexical gemination.

2.2.2. Morphological gemination

There exist two types of morphological gemination: concatenation, which involves the same consonant appearing twice, on each side of the morpheme boundary, whereas assimilation involves different consonants with the same place of articulation and/or voicing on the two sides of the morpheme boundary, and the non-matching features of the first consonant assimilate to those of the second consonant. Both of these morphologically derived types of gemination can be either truly phonological and signal lexical contrasts, or can be merely/purely phonetic and does not signal lexical contrasts. See Table 2.1.

Table 2.1: Morphological types of gemination

Gemination type	**Level**	**Examples**	**Reference**
Concatenation	Phonological	/rˤas + saf / [rˤas:af] 'the head of a hawk' (Moroccan Arabic)	(Ridouane & Turco, 2019)
	Phonetic	top pick = [top:ik]	(Pickett & Decker, 1960)
Assimilation	Phonological	/al + dar/ = [ad:ar] 'the house' (MSA)	(Alfozan, 1989: 84)
	Phonetic	'unnatural' (English)[a]	(Hedia, 2019: 7)

[a] *In English, double consonants resulting from affixation have been shown to be longer in duration (fake geminate) than their single consonant in the same context (i.e., unnatural vs untold: Hedia, 2019). Phonetic assimilation has been also observed in young children's acquisition of English, e.g., in pronouncing 'biscuit' they assimilate the [s] to the following [k], and thus produce the word as [bik:it], with a phonetic lengthening (phonetic gemination) of the [k].*

[1] In the International Phonetic Alphabet (IPA) the length diacritic /:/ following a consonant or vowel is used to indicate gemination.

Concatenated gemination can be phonological in some languages such as Maltese (Galea et al., 2014), Berber (Ridouane, 2010) and Moroccan Arabic (Frej, Carignan, Proctor & Best, 2018). That is, the duration of the concatenated phoneme consonantal is contrastive, i.e., consonantal duration does play a role in contrasting lexical items. Phonetically concatenated gemination occurs in languages such as German (Bergmann, 2014) and English (Hedia, 2019) in which phonetically long segments arise when two identical consonants spanning a word boundary are concatenated in production. To exemplify from English, the concatenated /p/ in <top pick> is longer in duration than the word-medial /p/ in <topic> (Hayes, 2001; Pickett & Decker, 1960). This kind of gemination is also referred to in the literature as "fake gemination" (Oh & Redford, 2012), because the geminated segment is phonetically long, but it does not constitute a phonological contrast. Non-phonemic, or allophonic, extension of consonantal duration in English and other languages is governed by factors such as stress, voicing, phrasal position and pragmatic focus.

Similarly, Assimilated gemination in bimorphemic words emerges word-internally in prefixed words when the final consonant of a prefix and the first consonant of the stem word assimilate, creating a phonetically long consonant. This type of gemination can either be phonological, i.e., true or, phonetic, i.e., fake. The difference between true gemination and fake gemination is that consonantal duration within one word distinguishes it from an otherwise identical word in the lexicon, either in mono-morphemic words or even in the case when gemination spans a morpheme/word boundary (Oh & Redford, 2012). In some Arabic dialects, including Moroccan and Lebanese, affixation results in true gemination. When the the definite article /al/ is followed by a coronal phoneme in the stem word, the vowel is deleted and the /l/, a coronal consonant itself, assimilates completely to the stem's initial consonant, resulting in gemination word-initially (e.g. /da:r/ "a house" vs /d:a:r/ "the house"). This makes initial gemination contrastive in some Arabic regional vernacular dialects, such as Moroccan and

Lebanese dialects. Unlike Arabic, in languages like English, assimilated gemination in bimorphemic words is deemed not contrastive. In these languages, however, consonant lengthening does occur across word boundaries when the same consonant appears at the end of one word and at the beginning of the next word. For instance, in English <fun name> differs from <fun aim> in the duration of closure associated with the nasal /n/. This type of gemination has been called "fake gemination" (Oh & Redford, 2012), because the geminated segment is phonetically long, but it is spread across two syllables falling on either side of a word or morpheme boundary and does not contrast with otherwise identical cases that differ only in having a shorter segment in that position. Therefore, it does not engage in a phonemic length contrast within any lexical items.

2.3. Distribution of gemination: Positional effects

In very few languages, such as Arabic (including their vernacular dialects), gemination can occur across all word-positions, i.e., word-initially, word-medially and word-finally, as shown in Table 2.2.

Table 2.2: Positional types of gemination

Gemination type	**Example**	**Reference**
Initial gemination	/pitu/ "door" /ppitu/ "at the door" (Kelantan Malay)	(Hamzah, 2013)
Medial gemination	/ħakam/ "he ruled", /ħakkam/ "he treated" (Arabic)	(Khattab & Al-Tamimi, 2008)
Final gemination	*ren* /ren/, "clean" *renn* / ren:/ "ski competition" (Norwegian)	(Payne, Post, Garmann & Simonsen, 2017)

Word-medial gemination is of a different nature and more common cross-linguistically compared to peripheral gemination. This is explained in more detail below.

2.3.1. Word-medial position

In the vast majority of languages where gemination is contrastive, the geminate/singleton contrast arises only in medial position, e.g., in Bengali (Lahiri & Hankamer, 1988), Hindi (Ohala, 2007), Icelandic (Pind, 1995, 1999) and Estonian (Engstrand & Krull, 1994). To examine gemination cross-linguistically, Thurgood (1993) reviewed various languages sourced from the Stanford Phonology Archive and other sources. He found that geminates were typologically preferred in medial position. Likewise, Hamzah (2013) compiled a list of 40 geminating languages, and found that word-medial gemination was permissible in 39 of those languages. Gemination was restricted to medial position in only 28 of the languages. The other 12 languages allowed gemination in one or both peripheral positions in addition to medial gemination in all but one of them.

Among the explanations that have been proposed for why word-medial gemination is the most common cross-linguistically is ease of perceptibility. Many researchers posit that the environment in which gemination occurs either helps or hinders the discriminability of gemination (Kawahara, 2005). Accordingly, the environment in which gemination is easiest to perceive is the medial position of the word (Dmitrieva, 2012), where flanking vowels offer a suitable phonetic context for gemination contrasts to be distinguished (Padgett, 2003). Typological studies found that geminates are preferred word-medially especially after a short vowel (Dmitrieva, 2012). This might be due to relative ease of producing and perceiving gemination is this environment (Hamzah, 2013). According to this account, the geminate /k:/ in *takka* /tak:a/ "fireplace" is easily differentiated from its singleton counterpart /k/ in *taka* /taka/ "back" in Finnish, because the preceding and following short vowel /a/ highlight the onset and offset of the medial /k/.

2.3.2. Peripheral gemination

Compared to word-medial gemination, peripheral gemination, i.e., word-initial or word-final, is rare cross-linguistically (Ladefoged & Maddieson, 1996; Muller, 2001; Thurgood, 1993). Thus, peripheral gemination is regarded as typologically marked (Dmitrieva, 2012). Of the languages in which peripheral gemination can occur, the majority nevertheless use either word-initial or -final gemination and rarely both. The least typologically preferred type of gemination is non-vowel-adjacent peripheral gemination, i.e., gemination occurring word-initially followed by a consonant or word-finally preceded by a consonant. In a survey of 40 languages that use lexically contrastive gemination, this type of gemination was found in only four of these languages: Moroccan Arabic, Swiss German, Yapese, and Berber (Pajak, 2013). Our focus will be on the more "typical," though still rare, peripheral gemination, i.e., vowel-adjacent initial and final consonant gemination.

Initial gemination is uncommon among geminating languages. Among the 40 languages surveyed by Hamzah (2013), only 11 allow gemination word-initially. Moreover, most languages with initial gemination do have gemination occurring in the word-medial position as well (Thurgood, 1993). Muller (2001) collected a database of 29 languages where initial gemination is contrastive, likewise noting that with few exceptions (Ngada, Nhaneun, Pattani Malay, Sa'aban) languages that include initial gemination have word-medial gemination as well. This is also the case with Swiss German (Kraehenmann, 2001), Berber (Ridouane, 2010) and all Arabic vernaculars that allow initial gemination. The latter include Moroccan and Lebanese, which are the focus of this study.

Final gemination appears to be rarer than even initial gemination (Dmitrieva, 2012). The existence of medial gemination in a language does not presuppose the existence of final gemination. However, whenever final gemination is found in a language, its medial counterpart

is found too (Muller, 2001). Final gemination is contrastive in Hungarian (Polgárdi, 2005), Norwegian (Lunden, 2006), Swiss German (Kraehenmann, 2001) as well as Arabic. In Hungarian, final geminates occur mostly in words borrowed from other foreign languages such as English (e.g., sokk [sok:] 'shock'), and is especially common in monosyllablic words (Magyar, 2014). Along with Arabic, Berber is one other language that allows gemination across different positions of the word, including word-finally (Ridouane, 2007).

The following section (2.4) offers an in-depth review of the literature on geminate consonants in the fields of acoustic phonetics and articulatory phonetics. It outlines some broad issues pertaining to consonant gemination and describes the phonetic parameters, including temporal and non-temporal correlates, which are chosen for examination in the current study.

2.4. Gemination production

Gemination exists in a number of languages, including Italian (Payne, 2005), Finnish (Doty et al., 2007), Estonian (Engstrand & Krull, 1994) and Berber (Ridouane, 2003, 2007, 2010). The only way to determine whether gemination is contrastive and maintained in these languages is through investigating a) whether differences along various acoustic correlates/parameters/cues responsible for contrasting geminates with singletons are significant and systematic in production, and b) whether these cues are relevant to native listeners for perceiving the difference between geminates and singletons. Because of positional variability in the realisation of some correlates signalling the geminate/singleton contrast, I discuss utterance position and prosodic hierarchy as two contributing factors. In some contexts, certain temporal cues are not measurable acoustically. This calls for the use of various kinematic techniques to investigate the articulatory characteristics of gemination.

2.4.1. Acoustic studies

The majority of research on the production of gemination has generally explored the acoustic correlates that signal the geminate/singleton contrast. These correlates can be divided into temporal as well as non-temporal acoustic cues.

2.4.1.1. Temporal cues to gemination

The primary goal of experimental studies on the production of gemination has been to investigate temporal acoustic correlates of geminates. For the most part, the results of these studies indicate that the relative duration of the constriction or closure is the primary cue[2] for creating the geminate/singleton contrast when multiple cues are available, and it is the only cue that is sufficient on its own, especially in the medial position in many languages, such as Bengali (Lahiri & Hankamer, 1988), Berber (Ridouane, 2010), Bernese (Ham, 2013), Buginese (Cohn, Ham & Podesva, 1999), Estonian (Engstrand & Krull, 1994), Finnish (Doty, Idemaru, & Guion, 2007), and Cypriot Greek (Tserdanelis & Arvaniti, 2001). Constriction duration is used broadly for all manners of articulation in gemination contrasts. For fricatives, the primary acoustic duration cue is the frication duration. Since stops involve a complete closure in the oral track, the duration that distinguishes between geminated and non-geminated stops is the duration of the oral closure. Since this PhD project is concerned with voiced and voiceless coronal stops in Arabic, the term 'closure duration' is used exclusively to refer to stop gemination throughout the remainder of this thesis. The longer closure duration for geminates

[2] The concept of "primary" cues has been used in phonetic studies to describe the consistency of a particular phonetic correlate across speakers, languages and different phonetic and phonological conditions (Kawahara, 2015). However, the definition we adopt here for a cue being "primary" is that it is the dominant cue relative to other "secondary" cues (Lahiri & Hankamer, 1988).

than singletons can be seen, for example, in the waveform and spectrogram of word-medial singleton /d/ and geminate /d:/ in MSA (see Figure 2.1).

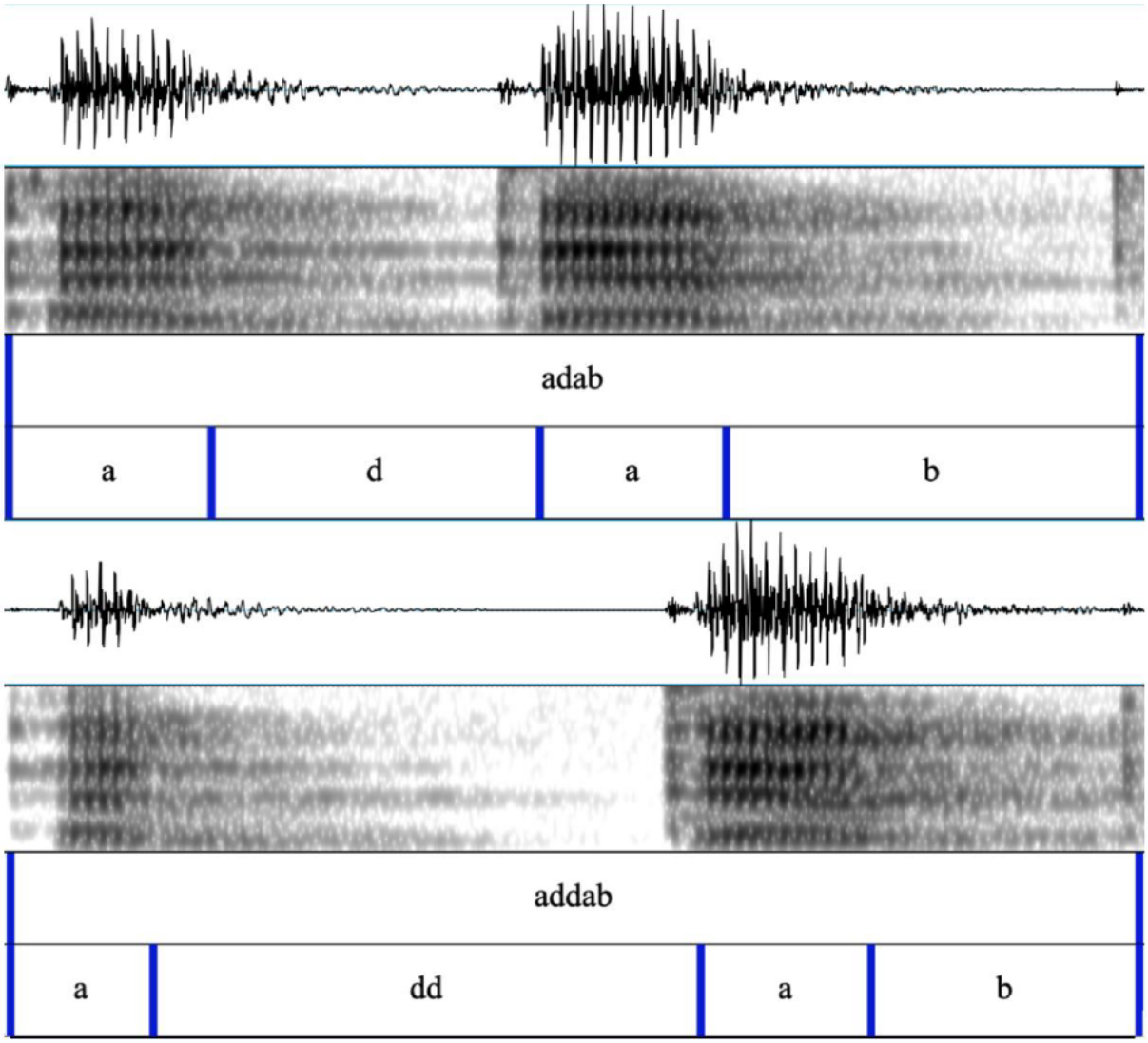

Figure 2.1: The waveforms and spectrograms of the target words for medial gemination, /adab/ and /ad:ab/, with the segmentation landmarks indicated underneath.

Closure duration is regarded as a primary cue due to its consistency and robustness relative to other cues that may also characterise the contrast. For example, Lahiri and Hankamer (1988) focused mainly on investigating durational properties of the geminate/singleton contrast in two different languages, Turkish and Bengali. The objective of the study was to determine the main durational acoustic cues responsible for the distinction and to verify whether these durational cues are found in both languages. For Bengali, they also analysed the differences

that may exist between within versus betwen morphemes geminates. For Turkish, in addition to the duration of the closure, measurements of the preceding vowel duration and Voice Onset Time (VOT) duration[3] were performed. The results revealed the predominance of closure duration as the main parameter for the geminate/singleton contrast. The duration of geminates was on average three times longer than that of singletons in Turkish. The duration of the preceding vowel did not vary much, while the VOT was significantly different for geminates and singletons, but was not significant for all speakers. This establishes closure duration as the primary cue in contrasting geminates with singletons. To further attest to the robustness of closure duration as a cue, its perceptual relevance was explored (Hankamer et al., 1989) with five native speakers of Turkish. The study investigated five acoustic cues other than closure duration (i.e., VOT, preceding vowel amplitude at offset as well as RMS values of burst, second syllable and impulse index). Closure duration was indeed found to be a significant perceptual cue for the distinction between singletons and geminates. VOT differences by themselves, on the other hand, did not support the listeners' perceptual differentiation between these two classes of consonants. The authors concluded that the additional cues beyond closure duration play an important role in perception but only when the duration of the closure is not easily distinguished.

Another factor that reveals the robustness of closure duration as a cue to gemination is the invariance of the geminate-to-singleton duration ratio[4] across different speaking rates. Four native Japanese speakers were recorded producing geminates and singletons (e.g., /kako/ 'past'

[3] VOT is the time between the release of the stop and the onset of the voicing of the vowel. For example, in English /ba/, the opening of the lips and the beginning of the vibration of the vocal cords are almost simultaneous; for English /pa/ the lips begin to open before the onset of vibration of the vocal cords (Liberman, Harris, Hoffman & Griffith, 1957).

[4] The geminate-to-singleton duration ratio is quantified as a ratio between the average geminate duration and the average singleton duration (Dmitrieva, 2012).

and /kak:o/ ‘parenthesis’) in a carrier phrase at three speaking rates (Hirata & Whiton, 2005). Durations of the closure and the preceding vowel and VOT were measured. The study found an overlap between singleton and geminate duration values between disparate speaking rates. That is, values registered for singletons in slow speaking rate overlap with those of geminates in fast speaking rate. However, this overlap does not affect the overall geminate-to-singleton closure duration ratio. The study concluded that a durational invariance is relationally formed between singleton and geminate categories. Indeed, Bouarourou and colleagues (2011) assert that the necessity to distinguish geminates from singletons in spite of the compression imposed by increases in speaking rate establishes closure duration as a robust parameter.

While closure duration is a strongly consistent feature differentiating geminates from singletons, the geminate-to-singleton closure duration ratio varies cross-linguistically. Acoustic studies have found the duration ratio to be as small as 1.38:1 (Norwegian: Fintoft, 1961) and as large as 3.40:1 (Malayalam: Local & Simpson, 1999). Ratios established for Arabic dialects lie between those extremes and vary across dialects. The ratio is larger in Moroccan than Lebanese dialects: medial geminates were found to be 2.40 times longer than singletons in Moroccan (Pajak, 2009), but only 1.82 times longer than singletons in Lebanese (Khattab, 2007). On perceptual grounds, it has been argued that a more acoustically distinct contrast based on a higher ratio is cross-linguistically preferred in comparison to the contrasts based on a lower ratio (Dmitrieva, 2012).

Moreover, the size of the geminate-to-singleton ratio varies not only from one language to another but also within the same language under the influence of word position. For example, in Moroccan Arabic, the highest duration ratio was registered for gemination occurring word-medially compared to word-final gemination (Frej, Carignan & Best, 2017). Nonetheless, a higher ratio in word-initial position compared with word-medial position was produced by

speakers of Berber, which was interpreted as a strategy employed by speakers to maintain the contrast word-initially (Ridouane, 2007). Therefore, the magnitude of the geminate-to-singleton ratio gives us an idea not only about the discriminability of the contrast but also about some language-specific strategies to maintain the contrast.

Although closure duration is considered to be the primary acoustic correlate for the geminate/singleton contrast, closure is inaudible for voiceless stops in absolute initial position: due to the absence of a preceding vowel and a lack of vocal fold vibration during the closure, there is no acoustic context to provide a perceptual cue for the onset of the closure. Yet native speakers of languages that contrast geminates with singletons word-initially, e.g., Pattani Malay, identify words with initial gemination quite well (Abramson, 1986). This finding raises the question of whether there might be additional acoustic correlates important to the perception of gemination in absolute initial position and possibly also in final position for similar reasons (Ridouane, 2010). One of these additional acoustic correlates is VOT, another durational property. Many studies recognise the importance of VOT as a cue to the geminate/singleton contrast particularly for voiceless stops in absolute initial position. Unlike closure duration, VOT is audible even in initial position because it is defined by the release of consonant closure and the start of voicing, acoustic events that are present even in the absence of a vowel preceding the consonant. VOT differences are employed primarily in languages with contrastive word-initial gemination. For example, speakers of Cypriot Greek differentiate between initial voiceless stop geminates and singletons through VOT (Arvaniti & Tserdanelis, 2000). In this language, the mean VOT duration for voiceless stop geminates is 114 ms, whereas singletons were realised with only 43 ms long VOT. However, the realisation of this durational parameter appears to be language-specific. While VOT was significantly longer for geminates than singletons in Cypriot Greek voiceless stops (Muller, 2001), the opposite was reported for Finnish (Doty et al., 2007) and Kelantan Malay (Hamzah, Fletcher & Hajek, 2011).

Moreover, a larger number of studies report non-significant differences in VOT, including but not limited to: Bengali (Lahiri & Hankamer, 1988), Bernese (Ham, 2001), Hungarian (Ham, 2001), Italian (Payne, 2005) and Japanese (Kawahara, 2012). Therefore, VOT is less cross-linguistically consistent than closure duration in distinguishing geminates from singletons.

Experimental studies on gemination have also investigated the duration of the vowels flanking the geminate/singleton consonant. The temporal characteristics of the geminate/singleton contrast for medial consonant gemination in three Indonesian languages were investigated by Cohn and colleagues (1999). In all three languages, the vowel preceding the geminates was shorter than the vowel preceding singletons. Notably, VOT did not contribute to the length distinction in these languages. In many other languages, the duration of the preceding vowel is also significantly affected. Generally, the vowel preceding a medial geminate is shorter than the one preceding a singleton, as shown for 12 of 16 languages reviewed by Hamzah (2013), which include Bengali (Lahiri & Hankamer, 1988), Buginese (Cohn et al., 1999), Hindi (Ohala, 2007), Icelandic (Pind, 1999), and Italian (Payne, 2005). For example, Ridouane (2007) reported the duration of the preceding vowel in Tashelhiyt to be 65 ms before geminates and 85 ms before singletons (i.e., vowels are 20 ms shorter in the geminate context). Geminates and the preceding vowel exhibit bahaved in a similar fashion in Moroccan, i.e., the vowel is shorter before a geminate than before a singleton (Pajak, 2009). However, such effect was not reported for Lebanese, i.e., no durational compensation (Khattab & Al-Tamimi, 2008). Because vowel shortening before geminates happens mostly in languages with word-medial gemination, this effect is taken by some scholars as evidence for the heterosyllabicity of word-medial gemination, i.e., geminates span across two syllables (Maddieson, 2018). The occurrence of final gemination in the final syllable of the word-makes final gemination structurally different from medial gemination. This is one of the motivations

for comparing the production of medial with final gemination in Modern Standard Arabic (MSA) in this thesis (Chapter 6).

Languages showing the opposite trend, i.e., vowel duration lengthening before geminates, also exist. These include Japanese (Port, Dalby & O'Dell, 1987), Persian (Hansen, 2004), Sinhala (Letterman, 1994) and Turkish (Lahiri & Hankamer, 1988). In Japanese, for example, the vowel [ɯ] was 68 ms before singletons and 86 ms before geminates. Both Hungarian (Ham, 2001) and Finnish (Lehtonen, 1970) allow long and short vowels to occur before geminated consonants. Interestingly, in both of these languages, short vowels are longer before geminated consonants than before singleton consonants. However, contrastively long vowels are still distinct from the lengthened singleton vowels that occur before geminated consonants. In Finnish, for example, long vowels are 177 ms long, while short vowels are 119 ms long, before geminate consonants. The vowel-lengthening effect before geminate consonants is argued to be due to the geminated (medial) consonant constituting a bimoraic foot in mora-timed languages (Burgel & Faehndrich, 2005) such as Hungarian and Finnish. This effect is attested for Lebanese Arabic (Khattab & Al-Tamimi, 2014), despite Arabic dialects not being considered to be mora-timed (Hamdi, Barkat-Defradas, Ferragne & Pellegrino, 2004).

Investigations of the duration of the vowel following the consonant are less reported in the literature. In the few reported cases, the following vowel generally tends be shorter after geminates than singletons, including in Japanese (Han, 1994; Idemaru & Guion, 2008), Estonian (Engstrand & Krull, 1994) and Ingrian (Markus, 2010). For example, in Japanese the mean duration of the vowel following geminates is 63 ms, while the vowel following singletons is 76 ms (Idemaru & Guion, 2008). The results for the post-consonantal vowel are particularly pertinent to gemination occurring word-initially due to the absence of a preceding vowel,

leading to reduced acoustic cues to gemination (Hamzah, 2013). However, because of the limited studies exploring this parameter, it remains unclear whether this cue is suited to describing the geminate/singleton contrast in Moroccan and Lebanese Arabic. This is investigated in our experimental chapter on initial gemination (Chapter 6).

The results from previous studies seem to suggest an inverse relationship between the preceding and the following vowel durations. When the vowel is shorter before geminates, the following vowel tends to be longer, and vice versa, i.e., a trading relationship between the two flanking vowels. For example, in Japanese, vowels were found be shorter in duration after geminates than singletons, unlike the opposite durational pattern for preceding vowels (Idemaru & Guion, 2008). This finding, however, is not universal. In Finnish, for example, there does not seem to be any such trading relationship between the flanking vowels as a result of gemination (Doty et al., 2007). Instead, the duration of the pre-consonantal vowel varies according to the phonemic length of the vowel occurring after geminates. That is, singleton vowels preceding a geminate consonant are shortened when the vowel occurring after a geminate is phonemically long, whereas in the context of post-geminate short vowels, pre-geminate singleton vowels are lengthened. It would be important to determine whether a trading relationship between the flanking vowels accompany gemination in either Moroccan or Lebanese. This is addressed in more detail following our investigation of acoustic correlates of medial gemination in MSA as produced by Moroccan and Lebanese native speakers (Chapter 6).

In this section, we have reviewed various temporal acoustic parameters at play in the realisation of the geminate/singleton contrast. Of particular importance is the closure duration that is described by most studies on gemination as the primary cue to the distinction. While all languages in which gemination is phonologically contrastive distinguish geminates from

singletons by lengthening the closure duration for geminates, the geminate-to-singleton duration ratio varies cross-linguistically and across different word positions in the same language. Closure duration is often accompanied by other durational acoustic cues, such as VOT and the duration of flanking vowels. However, languages vary greatly in terms of how these parameters are used, suggesting that the parameter settings are not universal.

In the next section, we review evidence that some non-temporal correlates could be as important as temporal correlates in contrasting geminates with singletons, contrary to the common assumption that gemination is purely a durational distinction (Lahiri & Hankamer, 1988).

2.4.1.2. Non-temporal cues

Experimental studies investigating gemination, particularly at the periphery of the word, have focused on non-temporal cues that might play a role in signalling the contrast. Commonly investigated non-temporal correlates are the amplitude of the release burst and the fundamental frequency of the following vowel. For example, Abramson (1991) examined the nature of the cues signalling the geminate/singleton contrast in word-initial position in Pattani Malay, where only initial gemination is allowed. The measurements of disyllabic words revealed significant differences in burst amplitude between words starting with a geminate and those starting with a singleton. Similarly, in Kelantan Malay, a language in which gemination occurs only word-initially as well, the release burst of voiceless stops is realised with a higher amplitude for geminates than for singletons, which was interpreted as resulting from the use of greater articulatory effort to mark gemination (Hamzah, Fletcher & Hajek, 2012). Word-medial geminated stops are also realised with greater amplitude of the release burst relative to their singleton counterparts in some languages, such as Hindi (Shrotriya et al., 1996) and Finnish (Doty et al., 2007). For example, in Finnish, the mean amplitude of the release burst for

singletons was measured as 55.5 dB versus 57 dB for geminates. Therefore, amplitude of the burst is one non-temporal parameter we chose to investigate. We were particularly interested in finding out how it signals the geminate/singleton contrast and how it varies across word positions (Chapter 6 and 7).

Another non-temporal parameter that has been investigated is the amplitude of the flanking vowels. Abramson (1987) measured the average amplitude for vowels following word-initial stops in Pattani Malay and found that those following a geminate had higher amplitude than those following a singleton. The same tendency was observed for Finnish (Doty et al., 2007). However, as is the case with most secondary cues to gemination, this pattern is not consistent across all languages. This inconsistency may be related to language differences in the prosodic domain. In some languages with only word-medial gemination, such as Japanese, amplitude is affected for the vowel preceding rather than following the consonant (Idemaru, 2005). In this thesis, the amplitude of the vowels flanking the geminate/singleton contrasts is examined for MSA, Moroccan and Lebanese (Chapter 6 and 7).

In addition to the amplitude of the vowel, f0 of the following vowel has also been investigated. Most studies have reported higher values for the f0 of the vowel following a geminate than a singleton in initial position. For example, f0 is higher for initial-position post-geminate than post-singleton vowels in both Pattani Malay (Abramson, 1999) and Kelantan Malay (Hamzah, 2013). In this thesis, calculating f0 of the vowel is not limited to the word-initial gemination, but is done across the other two positions of the word (Chapter 6 and 7), because we are interested to find out how this parameter varies in defining the geminates/singleton contrast across different positions of the word in MSA, Moroccan and Lebanese dialects.

To summarise, along with temporal acoustic cues we reviewed in the previous section, native speakers of some languages have been found to rely on some non-temporal cues as well (e.g., burst release and vowel amplitude) to signal the geminate/singleton contrast. These cues are particularly found in contexts other than the medial position of the word. In the next subsections, we explore factors that contribute to this effect.

2.4.2. Factors differentiating medial from peripheral gemination

Although consonant gemination is traditionally considered to be a distinction based on duration, it has been proposed that it may instead be based on articulatory force, particularly at the periphery of the word. We explain how the position of the contrast at the edges of some prosodic domains can strengthen the articulation of gemination. Conversely, we also discuss how certain positions of the word can lead to the contrast being neutralised.

2.4.2.1. Gemination as segmental duration or articulatory tenseness

The contrast between a geminate and a singleton has traditionally been considered to be based purely on the duration of the consonant (Ham, 1998). That is, a geminate consonant is identical to its singleton counterpart except that their closure duration is different. This, however, has been proven not to be the sole difference in numerous research studies on gemination (e.g., Abramson, 1991; Hamzah, Fletcher & Hajek, 2012; Shrotriya et al., 1996; Doty et al., 2007). As noted above, the geminate-singleton contrast is often reflected in additional, non-temporal cues, such as higher amplitude of the burst and higher fundamental frequency of flanking vowels (e.g., Kelantan Malay: Hamzah, 2013). These features are often associated with so-called “tensed" consonants. These additional non-closure-duration correlates of gemination, which have been observed particularly at the peripheries of the word, have led some researchers to question whether the articulatory nature of gemination is forceful

articulation rather than longer closure duration (Ridouane, 2007). However, most studies on gemination have been conducted primarily on gemination occurring in the word-medial position. Therefore, exploring how the different cues are manifested across all positions of a word, especially in peripheral positions in absolute utterance-initial or -final position where prosodic boundary effects can be assessed, is crucial to discovering the articulatory nature of gemination.

Forceful articulation, also referred to as “tenseness” (Ridouane, 2010) “articulatory strength” (Al-Tamimi & Khattab, 2011), or “fortition” (Hamzah, 2013), can be described as an articulation of a consonant with a stronger muscular effort or tighter constriction than its non-tensed counterpart (Matthews, 2014). Early investigations of whether certain consonants are characterised by differences in articulatory force found that voiceless stops are typically articulated with a higher intra-oral air pressure relative to their voiced cognates, which can indicative of a forceful articulation (Malécot, 1968). Thus, the feature [fortis] was used to represent consonants articulated with such increased pressure. Later investigations further described the feature [fortis] to be manifested phonetically by an articulatory timing and laryngeal power / voltage component (Kohler, 1984). Evidence for more forceful articulation of geminates compared to singletons comes from Tashlhiyt Berber, where singletons, but not geminates, are often produced without a release burst (Ridouane, 2010). As an unreleased stop is a "weaker" version of its released counterpart (Fougeron, Kuehnert, Imperio, & Vallee, 2010), singletons appear to be produced with lower articulatory force than their geminate counterparts.

Slis (1971) highlighted the relationship between the tenseness and the duration of the consonant. In an electromyographic (EMG) experiment, he measured the duration of the closing movement of the lips and the activity of the muscles during the closure of labial stop

produced with different degrees of effort, and found that a longer duration of closure was associated with an increase in muscle activation. Thus, articulatory force and duration do not exist independently of each other. Kohler (1984) discusses support for the relationship between articulatory force and duration more broadly. Not only the duration of consonantal closure but also the duration of the previous vowel, of aspiration (for voiceless tops) and of the consonant in total are associated with more forceful articulation (also see Jakobson & Waugh, 1979; Ladefoged & Maddieson, 1996; Jessen, 1999).

The observations just reviewed raise the question of whether tenseness is the result or the cause of longer duration for geminated consonants. On the one hand, it can be argued that longer duration is a consequence of a tense articulation (Ridouane, 2003), i.e., that a closure is longer because it involves a greater compression between the active and passive articulators. However, the opposite argument is plausible too: it may be that increased muscle activation is needed to maintain a longer closure.

The question of whether the phonological and articulatory nature of gemination is a forceful articulatory distinction (fortis/lenis or tense/lax), or a duration/elongation distinction remains debatable mainly because aspects constituting a forceful articulation include also an increase in duration. Very few studies have tackled the articulatory nature of gemination. In this study we propose that gemination is the result of both forceful articulation and elongation, that the balance of their effects is language specific, and that more forceful articulation plays a larger role in peripheral geminates, where durational cues are less apparent, than in medial geminates, and that this influence is enhanced at the boundaries of utterances/phrases (phrase-initial and -final position). Next, we explain how this effect can be explained through principle of contrast maximisation at the edges of the prosodic domain (Cho & Jun, 2000; Cho & Keating, 2001).

2.4.2.2. Prosodic hierarchy

An important suprasegmental factor that could affect how gemination is produced in specific contexts is the contrast maximisation principle (Cho & Jun, 2000; Cho & Keating, 2001). This principle stems from the idea that prosodic domains are ranked hierarchically, with certain domains being higher than other domains (Beckman & Pierrehumbert, 1986). One of the posited effects is that increased articulatory strengthening/enhancement is associated with each higher level in the prosodic hierarchy (syllable < word < smaller phrase < intonational phrase < utterance). That is, the prosodic position of a contrast determines how strong its articulation is. Articulatory strengthening can be described as an increase in the spatio-temporal magnitude of one or more key articulatory gestures/actions. As such, consonants in higher prosodic domains have been found to be articulated with more extreme and longer constriction/closure durations than those in lower domains, thereby maximising their contrastiveness with other consonants (Cho & Jun, 2000; Cho & Keating, 2001). One of the primary prosodic positions that has been shown to cause articulatory strengthening/enhancement is the initial position in a prosodic unit, i.e., the initial position of a syllable, word, phrase, or utterance, with the effect increasing as initial position lines up across multiple prosodic units, i.e., greatest for phrase-initial (e.g., Pierrehumbert & Talkin, 1992; Cho & Keating, 2001). For example, the coronal stop /t/ in English has a longer VOT phrase-initially compared to when it occurs in phrase-medial position (Pierrehumbert & Talkin, 1992). Similarly, investigations using Electropalatography (EPG), an instrumental technique for recording information about the tongue's contact with the hard palate during speech, have shown that consonants are produced with a larger and longer lingual contact phrase-initially than phrase-medially (e.g., Fougeron & Keating, 1997).

Two types of domain-initial strengthening may arise: syntagmatic strengthening and/or paradigmatic strengthening (Fougeron, 1999). Syntagmatic strengthening refers to articulatory/acoustic amplification of the relationship between a segment and those that are adjacent to it. For example, the difference between a consonant and the following vowel may be enhanced. Paradigmatic strengthening refers to articulatory/acoustic amplification of the differences between a segment and another contrastive segment in the sound system of a language (e.g., geminate vs. singleton). Syntagmatic enhancement of a domain-initial geminate consonant would be indicated by longer consonantal closure duration at the expense of a shortened post-consonantal vowel in utterance-initial position. Paradigmatic enhancement, on the other hand, would be indicated by longer closure duration for geminate consonants than for singletons in utterance-initial position (as compared to how the contrast would be realised in utterance-medial position). It is important to note that both paradigmatic and syntagmatic processes can co-occur in prosodic strengthening contexts.

Word-initial gemination is ideal for testing the principles of domain-initial strengthening, because unlike word-medial gemination, word-initial gemination can occur in the initial position in an utterance. This allows us to compare the same word-initial gemination contrast at two levels of the prosodic hierarchy: the domain-initial level, i.e., utterance-initially, versus the weaker domain-medial level, i.e., utterance-medially. Word-initial gemination in citation form represents the highest level of the prosodic hierarchy. This provides an emphasised, carefully pronounced and prosodically strengthened environment, whereas word-initial gemination in medial position of the utterance is prosodically weaker because this context sits at a lower level of the prosodic hierarchy.

Strengthening/lengthening is not limited to phrase-initial position. Phrase-final lengthening is well documented in the literature. It is in fact a common feature in most

languages (Blevins, 2008). Phrase-final syllables ending with a stop are reported to be lengthened compared to phrase-medial ones in English (Oller, 1973). Scholars theorise that phrase-final lengthening is a strategy used by speakers to cue the location of the boundaries of words, phrases, or sentences. In the next section, we explain how phrase-initial and -final lengthening does not ensure the maintenance of initial and final gemination.

2.4.2.3. Contrast neutralisation

A contrasting strategy to the strengthening of the contrast at the edges of phrase is that peripheral gemination is prone to being neutralised. Neutralisation[5] of initial gemination, i.e., degemination processes, could result in loss of the contrast word-initially. The closure duration is inaudible especially for voiceless stops in absolute initial position: due to the absence of a preceding vowel and the corresponding lack of vocal fold vibration, there is no acoustic context to provide a perceptual cue for the onset of the closure. Poor acoustic closure duration cues to gemination in word-initial position could, for example, lead to diachronic neutralisation of the contrast (Blevins, 2004). Findings from some studies are compatible with this possibility. In Swiss German, the geminate/singleton phonological contrast is phonetically neutralised in word-initial position (Kraehenmann, 2001). Similarly, in Russian geminates are freely degeminated word-initially (Kasatkin & Choj, 1999). That is, the geminate/singleton phonological contrast is not maintained acoustically. However, initial gemination is maintained in some languages, such as Kelantan Malay (Hamzah, 2013). One of the motivations for the current project is that it is currently unknown whether initial gemination is maintained as contrastive in either Moroccan or Lebanese. The dearth of experimental studies on initial gemination in both Moroccan and Lebanese warrants a comprehensive acoustic and

[5] The elimination of acoustic cues distinguishing between a geminate and a singleton, leading to the loss of the geminate/singleton contrast

articulatory study of this kind to investigate whether initial gemination is contrastive in these dialects.

Similar to initial contrast, word-final stops can be realised with reduced acoustic cues. This happens when voiceless stops are produced without a burst release. Since there is no vocal fold vibration to provide acoustic information about the temporal duration of the consonant and no final release to indicate the end of the voiceless closure, the closure duration becomes inaudible. As a result, word-final geminate/singleton contrasts can be subject to neutralisation. Historical data from Middle High German show that word-final gemination was contrastive in the language and that this contrast was diachronically neutralised, resulting in the absence of final gemination in the East Central Dialect of New High German (Griffen, 1990). Similarly, Swabian, the South Western dialect of German, also underwent the diachronic process of degemination word-finally. However, the loss of the contrast was compensated for by the lengthening of the preceding vowel. Word-final geminates undergo the same process of neutralisation in an unrelated language, Hungarian. This process affects glides and liquids, and although degemination is optional or occurs in some but not other regional vernaculars, the contrast is carried by the lengthening of the preceding vowel (e.g., [tol:] – [to:l] ‘pen’) (Curtis, 2003).

Importantly, the lack of closure onset information in the acoustic signal for voiceless stops in utterance-initial position makes acoustic measurement of closure duration for these segments impossible. The same holds true for final geminates phrase-finally produced without an acoustically identifiable release. Thus, it is necessary to apply articulatory (especially kinematic) techniques in order to investigate the articulatory characteristics of the geminate/singleton contrast.

2.4.3. Articulatory studies

One of the objectives behind articulatory investigations of gemination is to find out whether the geminate/singleton contrast is maintained articulatorily when the closure duration is acoustically neutralised in some languages and some contexts (e.g., voiceless stops in absolute initial position). However, in comparison with acoustic studies, articulatory investigations of gemination are not nearly as common. As a result, with the exception of a longer articulatory contact for geminates than for singletons, it is hard to presume the universality of other articulatory strategies deployed by speakers of some languages to signal the contrast. Similarly, there is a dearth of articulatory studies on gemination in Arabic. Thus, little is known about the articulatory mechanisms employed by Arabic speakers to contrast geminates with singletons across different word positions. In the next section, we review the various common kinematic techniques used to investigate the articulation of gemination. We also explore whether the geminate/singleton contrast is based on articulatory differences alongside the acoustic differences we have reviewed.

2.4.3.1. Articulatory investigations of gemination with various kinematic techniques

Quite a few studies have reported that even when the contrast is not maintained acoustically, as with word-initial voiceless stops, significant articulatory differences are found between geminates and singletons in languages such as Tashlhiyt Berber (Ridouane, 2007) and Swiss German (Kraehenmann & Lahiri, 2008). Using EPG, these studies found that a longer contact between the tongue and the palate is maintained for geminates than singletons. Using the same technique to investigate the non-durational characteristics of the geminate/singleton contrast in Italian, Payne (2006) found coronal geminates to involve a more palatalised articulation than their singleton counterparts, resulting in a laminal configuration of the tongue for geminates but an apical tongue configuration for singletons.

Dunn (1993) investigated the movements of the lips during the closure phase of labial geminates and singletons in Italian and Finnish. He found that the lips maintain contact longer for geminated consonants, but that lip movement was found to be comparable between geminates and singletons. Similarly, Lofqvist (2005, 2006, 2007) conducted kinematic experiments using electromagnetic articulography (EMA) on the movement of the lip, jaw and tongue during the articulation of the geminate/singleton contrast in Japanese. The most important findings were: 1) longer articulatory closure for labial, alveolar and velar geminated stops relative to their singleton counterparts, 2) geminates were realised with slower tongue movement velocity and slower vowel-to-vowel movement and 3) for the most part, the peak vertical position of the lower lip during the oral closure/constriction was higher for geminates than singletons, but no difference in lip velocity was observed.

Other kinematic techniques, such as X-ray and real time magnetic resonance imaging (rtMRI) have also been employed to investigate the articulation of gemination. For example, an X-ray investigation of voiced and voiceless stop gemination word-initially, -medially and -finally in Tarifit Berber revealed that the maximum value for contact, measured from the mid-sagittal profiles, is longer for geminate stops than for singletons even for voiceless stops in word-initial position (Bouarourou, Ridouane & Hirsch, 2008). The same technique (X-ray) revealed another articulatory mechanism involved in the production of geminates, namely that the maximal contact is achieved at a later phase in the closure for geminates than for singletons in word-medial position in Japanese (Takada, 1985). Using a rtMRI analysis of the midsagittal vocal tract plane, the release of geminated coronal stops in Miyako language were characterised by not only a longer linguopalatal contact, but also a larger area of contact compared to singletons (Fujimoto & Shinohara, 2015).

Articulatory techniques are most useful for measuring closure duration when this closure is not measurable acoustically. For that, as we have seen, various techniques such as EPG (Ridouane, 2007) and oral airflow (Ridouane, 2003) have been put into use. The results from these studies suggest that geminate stops are produced with a longer articulatory contact (i.e., extended over a longer period of time) compared to singletons. However, the most informative data would capture the actual temporal dynamics of articulator movements in the formation of singleton versus geminate stop closure, whereas EPG and oral airflow provide only indirect estimates of those motions. More direct techniques, such as EMA or ultrasound, would be particularly useful for this, but have been used only sparsely in research on gemination thus far, as we show for ultrasound in the next section.

2.4.3.2. Ultrasound

Studies using ultrasound to investigate the geminate/singleton contrast are scarce. Ultrasound is an ideal technique for tracking midsagittal tongue shape by obtaining real-time video sequences of the upper surface of the tongue by sending and receiving high frequency sound waves through the tongue using a transducer held under the chin. The waves reflect at the boundaries between soft tissue and air, such as the upper surface of the tongue, allowing imaging of these boundaries in a non-invasive manner. Ultrasound is easily accessible to researchers due to its relative low cost and portability compared to other articulatory methodologies, e.g., rtMRI. Moreover, since the ultrasound machine is quiet in operation and the transducer is placed completely outside of the oral tract, there is less acoustic interference with the speech signal and less physical interference with speech production compared to other commonly used kinematic imaging technologies, such as rtMRI: MRI scanners produce a great deal of mechanical noise or EMA: sensors adhered to the tongue can interfere with natural speech production.

Despite the sizable number of studies conducted on gemination, articulatory investigations of the contrast are few and far between, and far less common are studies employing ultrasound to examine the temporal dynamics of tongue movements in gemination contrasts. Moreover, the few studies on gemination using ultrasound have focused on the investigation of tongue shape during the production of geminates and singletons, rather than on the dynamic temporal characteristics of the consonant closure. Ultrasound has been used to examine tongue shapes for geminate dental /t/ and geminate affricate /ts/ along with retroflex /ʈ/ and velar /k/ in Kannada (Kochetov, Sreedevi, Kasim & Manjula, 2012). The results showed that the affricates were characterised by a more anterior and raised tongue body and a laminal constriction, relative to the lower, more posterior tongue body and apical constriction of the dental geminate. It should be noted that no comparison was made with their singleton counterparts. In a more recent ultrasound investigation of coronal stops in Eastern Oromo, it was found that singleton ejectives were produced with a more raised tongue body than their geminate counterparts (Percival, Kochetov & Kang, 2018).

Although studying the differences in tongue shape during the production of geminates and singletons can be useful, temporal analysis would be more informative for a contrast that is based primarily on durational differences. The main reason why ultrasound investigations of gemination disregard its temporal aspect is that they tend to use the most common approach for analysing ultrasound data, namely tongue contour tracing, which can be relatively time-consuming. In order to minimise the amount of work and time necessary to process an entire ultrasound video recording, tongue contours are usually traced at single specific time points, i.e., a static approach, rather than for every image frame in the recording (see Falahati, 2013, for an example of the latter). This static approach to ultrasound analysis can be useful in investigating characteristics of speech articulation at a particular point in time, but it ignores the temporal aspects of speech production. Therefore, this analytic approach would be

ineffective for studying contrasts such as the geminate/singleton contrast, for which temporal characteristics are central to the distinction.

Some studies have addressed this limitation by generating dynamic information from ultrasound videos, either by analysing pixel differences between time frames (Palo, Schaeffler & Scobbie, 2014), using optical flow analysis (Moisik et al., 2014), or conducting a principle components analysis of pixel intensity changes over time (Hueber et al., 2007). In this project, we adopted a derivation of the latter approach, Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS: Carignan, 2014). TRACTUS analysis identifies a set of principal component (PC) coefficient matrices for an entire ultrasound recording that account for the greatest amounts of variance in the ultrasound data. The PC scores of frames associated with a closure are used as predictors in training Linear Discriminant Analysis (LDA) models. Subsequently, predictions made from the LDA model produce a temporal vector representing the probability that a particular image corresponds to a stop closure. The higher the prediction scores produced by the model, the more likely that an ultrasound frame at a particular time point corresponds to a closure. In this way, TRACTUS performs temporal articulatory analysis on ultrasound images automatically without the need for image tracing, by creating a time-varying "closure" signal. This method is ideally suited to addressing our research questions about closure duration because it provides temporal information about speech kinematics, but with substantially less manual processing than tongue contour tracing. More details on TRACTUS and ultrasound data analysis are provided in the methodological sections of Chapter 6 and 7.

2.4.4. Summary

Like most phonological contrasts, the geminate/singleton contrast is realised phonetically in ways that give rise to various temporal and non-temporal acoustic differences

in the consonants as well as the flanking vowels. Most studies on gemination have found that closure duration plays an important role in contrasting geminates with singletons. Meanwhile, many studies have also observed the importance of additional cues to the contrast, such as VOT and the amplitude of the burst. It has been reported that these non-closure duration cues are enhanced especially in contexts where the closure duration can be acoustically neutralised, namely, in the word-initial and word-final positions. Because acoustic means can fall short of measuring the closure in such contexts, kinematic techniques are useful for measuring the articulatory mechanisms of gemination. We also explained that although ultrasound offers numerous advantages over other commonly employed techniques, such as EMA and rtMRI, ultrasound investigations of gemination are rare, particularly those using more automatic processes to analyse ultrasound.

In order to understand how gemination can arise at the edges of the word in Arabic, in the next chapter, MSA and two Arabic dialects (Moroccan and Lebanese), which are the target languages of this thesis, are introduced prior to reviewing research on non-native and cross-dialectal perception of gemination.

CHAPTER 3: GEMINATION IN THE TARGET LANGUAGES: MODERN STANDARD ARABIC AND MOROCCAN AND LEBANESE DIALECTS.

This PhD thesis takes up a comprehensive articulatory, acoustic and perceptual investigation of gemination in two distinct varieties of standard Arabic and its vernacular dialects, Moroccan and Lebanese, in order to determine the various regional similarities and differences between the geminate/singleton contrast across word peripheral positions in relation to medial positions. Knowing the acoustic and articulatory similarities and differences between the regional varieties is crucial to interpreting the results of the perception experiments, whose main objective is to determine whether and how knowledge of consonantal duration distinctions in one word-position helps or hinders perception of gemination occurring in other positions that are unfamiliar to the listener. This is done by exploring how three groups of non-native listeners who differ in their native language experience with consonant duration distinctions discriminate initial gemination in Moroccan and medial and final gemination in Modern Standard Arabic (MSA) as produced by Moroccan speakers in the production experiment.

Arabic-naïve native English listeners (henceforth, English group), Moroccan-naïve native Lebanese listeners who speak both Lebanese vernacular dialect and the Lebanese variety of MSA (henceforth, Lebanese group), and Moroccan-naïve Lebanese heritage listeners who learned Lebanese dialect as their home language in Australia but have not learned standard Arabic (henceforth, Heritage Lebanese group).

However, the acoustic and articulatory properties of word-initial and -final gemination have been understudied, particularly in Arabic and its vernacular dialects, as I explain in this chapter. Furthermore, some experimental studies have shown that there is a tendency for gemination at the edges of the word to be neutralised in some languages. As such, it is reasonable to examine the maintenance of gemination occurring at the periphery of words in Arabic and its regional vernacular dialects. Little is known about regional dialect differences in realising the contrast. Therefore, an in-depth investigation of production of gemination at different positions of the word by native Moroccan and Lebanese Arabic speakers was warranted prior to investigating how the contrast is discriminated by non-native listeners.

In the next sections, we provide background knowledge about Arabic and its regional dialects as well as about the phonotactics of peripheral gemination in these dialects (3.1). Subsequently, we review non-native/cross-dialectal research conducted on the perception of gemination, particularly involving the target listener groups of this PhD project (3.2).

3.1. The Arabic language and regional variations

Arabic is spoken primarily in the Middle East and North Africa, where it is the official language of 26 countries. Over 280 million people speak it as their first language. Moreover, as Arabic is the language of the Quran, the holy book of Islam, many native Muslims around the world strive to learn it for religious purposes, which makes it amongst the top five most spoken languages in the world. In Australia, Arabic has been gaining popularity among university L2 learners[6]. The Asian Studies Association of Australia has been conducting

[6] Arabic is currently offered at five universities in Australia: the Australian National University, Deakin University, Melbourne University, Sydney University and Western Sydney University.

surveys since 2001 on Languages Enrolments in Australian Higher Education, and reports that over that period Arabic has had the highest increase in rate of enrolment of any foreign language (142%).

Every Arabic L2 learner is faced with added linguistic complexity when interacting with native speakers from across the Arab world. While Modern Standard Arabic (MSA) is the shared written language of all Arab countries, and is the target language that most of them are formally learning as well as a lingua franca or shared language, there is a different regional vernacular dialect spoken in each of those countries that differs from other dialects but also from MSA in grammar, vocabulary, and pronunciation. This situation of use of a common overarching formal language plus a related local vernacular dialect is referred to as diglossia.

3.1.1. Diglossia in Arabic

The term diglossia comes from the Greek word *diglōssos* 'bilingual' (Oxford English Dictionary). Thus, diglossia was synonymous with bilingualism until 1930, when the French linguist Marçais used the term *diglossia* in his article "Diglossie Arabe" to describe the linguistic situation in the Arab world (Marçais, 1930). Ferguson (1959) defines diglossia as follows:

> "Diglossia is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any section of the community for ordinary

conversation." (Ferguson, 1959: 336).

In other words, diglossia is a situation in which a community uses two different languages or varieties of a language for different situations. One variety is used for casual everyday interactions, and a different standard version is used for more official and formal purposes. Diglossia is a relatively common cross-linguistic phenomenon. Some of the languages that are distinguished by a diglossic situation are Italian and German. However, the most notable and most commonly discussed example of diglossia is that of the Arabic language. MSA is a codified language, i.e., it is fixed in its rules and it does not change much over time, while local Arabic vernacular dialects are constantly evolving. In casual speech with other locals, Arabs exclusively use their local dialects, but in writing, in formal speeches, and in the news MSA is used. MSA is also used as a *lingua franca* when people are communicating with speakers of other dialects, but this happens only when necessary (Al-Kahtany, 1997).

Despite the standard form of Arabic, MSA, being widely accepted as the official language in all Arab countries, most Arabs are not proficient speakers of it. This is because MSA is only taught at schools. Thus, heritage speakers[7] of Arabic who did not have a formal education in Arabic can neither read nor understand MSA and are only familiar with their respective Arabic dialect in its oral form. In contrast, L2 learners of MSA are taught mainly this standard form of Arabic and have little to no exposure to dialectal Arabic. The situation of diglossia in Arabic is so extreme that certain dialects spoken in the far eastern and western regions of the Arab world are mutually unintelligible (Watson, 2002), which is the case for Moroccan and Lebanese. This reflects the amount of linguistic divergence that Arabic

[7] Migration of Arabs to other non-Arab countries has resulted in second-generation speakers of various Arabic vernacular dialects in the immigrants' adoptive countries around the world, who are referred to as "heritage speakers" of those vernacular dialects (See section 4.2 for a more detailed description of heritage language and speakers).

vernaculars exhibit from MSA, which in turn would impact the production of MSA by the native speakers of those Arabic vernaculars. Thus, Moroccan and Lebanese Arabic speakers, located ~4,000 km apart at the western versus eastern extremes of the Middle East, may produce gemination differently in MSA as well as in their native Arabic vernacular dialects. This possibility is examined in detail through a comprehensive acoustic and articulatory study on initial gemination in Moroccan and Lebanese regional Arabic dialects (Chapter 6) and another comparing word-medial and -final gemination in MSA as spoken by Moroccan and Lebanese native Arabic speakers (Chapter 7). In the next section, we review some key phonological differences between MSA and regional Arabic dialects.

3.1.2. Phonological differences between Modern Standard and dialectal Arabic

A number of studies have been dedicated to investigating the linguistic differences between MSA and the various regional Arabic dialects, or vernaculars, that are spoken in day-to-day life. These dialects manifest a clear linguistic divergence from MSA. In comparing MSA with Iraqi Arabic vernacular dialect, notable phonological, morphological, syntactic, and lexical differences between the two varieties have been identified (Altoma, 1969). These differences include the word-final /aw/ in MSA produced as /oː/ before a suffix in Iraqi vernacular dialect, and the endings /uːn/ and /iːn/ in MSA produced as /uː/ and /iː/, respectively, in Iraqi dialect. Similarly, the extent to which Northern Jordanian Arabic dialect (NJA) maintains the phonological alternations of MSA was studied by comparing the behaviour of the glottal stop in different word-positions (initial, medial and final) with that documented for MSA (Al-Tamimi & I'lawi, 2006). The paper concluded that the realisation of the glottal stop in NJA is most similar to MSA word-initially. However, in both medial and final position, the glottal stop has a stronger tendency to be neutralised in NJA than in MSA. Another language whose phonology was compared to that of MSA is Madani Arabic dialect, spoken in Palestine

(Amer, Adaileh & Rakhieh, 2011). In line with previous studies, those authors found several important differences between the two varieties. Madani differs from MSA in having long vowels /e:/ and /o:/, and realises emphatic voiced alveolar fricative /ẓ/ of MSA as emphatic voiced alveolar stop /ḍ/, voiceless dental fricative /θ/ as voiceless alveolar stop /t/, voiced dental fricative /ð/ as voiced alveolar fricative /z/ or stop /d/, voiced palatal affricate /dʒ/ as voiced palatal fricative /ʒ/, and voiced velar stop /q/ as glottal stop /ʔ/. Madani also inserts epenthetic vowels between multiple consonants, as in the clusters that occur in MSA, leading to de-clusterisation. These and other processes emerge variably in different Arabic local dialects, thus differentiating them from MSA. However, these findings have been largely observataional and descriptive. Little attention has been given to acoustic and/or articulatory comparisons between MSA and dialectal Arabic, and we are not aware of any research investigating whether and if so, how, native Arabic speakers produce gemination differently in MSA and in their native Arabic vernacular, including Moroccan and Lebanese.

Next, we will show how these two geographically distant Arabic dialects differ from each other phonetically and phonologically. Their maximal geographic, phonetic, phonological and grammatical differences are some of the reasons why they were selected to be the focus of this project.

3.1.3. The phonemic inventories of Moroccan and Lebanese Arabic dialects

Moroccan and Lebanese dialects share many phonological features with MSA. Most phonemes are native to MSA with the exception of the ones in parentheses in Table 3.1 and 3.2. These non-native phonemes, which do appear sometimes in MSA, are usually the result of word borrowing from other foreign languages (see Abdul-Karim, 1979 and Watson, 2002).

Table 3.1: Moroccan Arabic dialect consonants (Watson, 2002)

Place of Articulation / Manner		Labial		Dental-Alveolar		Palatal	Velar	Uvular	Pharyngeal	Glottal
		plain	emphatic	plain	emphatic					
Nasal			(mˤ)	n						
Plosive	**Voiceless**	(p)		t	tˤ		k	q		ʔ
	Voiced	b	(bˤ)	d			(g)			
Fricative	**Voiceless**	f	(fˤ)	s				χ	ħ	h
	Voiced	(v)		z		(ʒ)		ʁ	ʕ	
Tap					ɾˤ					
Approximant				l	(lˤ)	j	w			

Table 3.2: Lebanese Arabic dialect consonants (Abdul-Karim, 1979)

Place of Articulation / Manner		Labial	Alveolar		Palatal	Velar	Uvular	Pharyngeal	Glottal
			plain	emphatic					
Nasal			n						
Plosive	**Voiceless**	(p)	t	tˤ	(t͡ʃ)	k			ʔ
	Voiced	b	d			(g)			
Fricative	**Voiceless**	f	s			x	χ	ħ	h
	Voiced	(v)	z		(ʒ)	ɣ	ʁ	ʕ	
Tap/trill									
Approximant			l		j	w			

A close comparison of Tables 3.1 and 3.2 illustrates certain discrepancies between the two inventories[8]. In particular, Moroccan has a slightly larger consonantal system than

[8] It is important to note that the phonemic inventories of Moroccan and Lebanese presented in Tables 3.1 and 3.2 respectively are reflective of the standard spoken version of these dialects and do not necessarily represent dialectal variations within each of them.

Lebanese. For example, emphatic phonemes in Moroccan outnumber those of Lebanese: The emphatic /mˤ/, /bˤ/, and /fˤ/ occur only in Moroccan. These emphatic phonemes occur almost exclusively in word initial position. They are also pronounced with slight lip-rounding (Mitchell, 1993). The phoneme /lˤ/, however, is not very common in Moroccan. It occurs mainly in certain words borrowed from French, or when triggered by other emphatic phonemes inside the same word or morpheme (Heath, 1987). In both Moroccan and Lebanese, /p/ and /v/ occur primarily in loanwords, and they are pronounced as /b/ and /f/ respectively by some speakers. The plain /t/ in Moroccan is pronounced with a clear aspiration that makes it almost sounds like /tʃ/ in all positions (Harris, 1942), which makes it distinct from the plain, unaspirated /t/ in Lebanese. The MSA uvular stop /q/ is realised as a glottal stop /ʔ/ in Lebanese. In Moroccan, it is realised as /q/ in some sub-dialects spoken in Morocco and as /g/ in others. Likewise, MSA /dʒ/ is realised as both /ʒ/ and /z/ in Moroccan. In Lebanese, all occurrences of /dʒ/ are reduced to /ʒ/. Moreover, the MSA phonemes /θ/ and /ð/ do not occur in Moroccan and Lebanese. Instead, they are realised respectively as /t/ and /dˤ/ in Moroccan and as /t/ and /zˤ/ in Lebanese. Furthermore, the voiced (ɣ) and voiceless (x) velar fricatives in MSA are realised as the uvular counterparts (ʁ) and (χ) respectively in Moroccan. In Lebanese, the uvular fricatives occur in free variation with the velar fricatives (Mitchell, 1993). Alongside these differences in phonemic inventory, additional phonological differences also exist between Moroccan and Lebanese.

3.1.4. Phonological differences between Moroccan and Lebanese

Previous studies attribute certain crucial phonological differences between Moroccan and Lebanese to their different syllable structures (for detailed reviews see Kiparsky, 2003; Watson, 2007). In particular, studies have classified the Arabic vernaculars into different syllable structure groups: C-dialects (Moroccan: Harrell, 2004) and VC-dialects (Lebanese:

Blanc, 1953 & Cowell, 2005). VC-dialects are those in which morphologically derived CCC clusters are syllabified as CVCC and C-dialects are those in which CCC clusters are syllabified as CCC.

One important phonological difference resulting from the difference in syllable structure between the two dialects is that word-final consonant clusters occur without restrictions only in Moroccan (e.g., /hrəbt/ 'I escaped' and /qelb/ 'heart'). Lebanese generally does not permit final CC clusters (e.g., /harabit/ 'I escaped'[9] corresponds to the Moroccan /hrəbt/) but allows them in certain cases of falling sonority (e.g., /ʔalb/ 'heart'). Thus, if final gemination patterns with final clusters, this would increase the functional load of final gemination in Moroccan compared to Lebanese, which in return could enhance the final gemination contrast in MSA as produced by Moroccan speakers. This speculation warrants more investigation.

Both dialects, however, allow clusters in word-initial position. In Moroccan, words such as /kta:b/ 'a book' and /kla:b/ 'dogs' are very common, because the short vowels in word-initial position in the originating MSA words, in this example /kitaab/ and /kilaab/, are often deleted in Moroccan. In Lebanese, these words are often realised exactly as in Moroccan (i.e., *ktaab* and *klaab*), but in some sub-dialects a glottal stop and a prothetic, i.e., inserted, vowel /ʔi/ is used to break this cluster across syllables, e.g., /kta:b/ in MSA/Moroccan/some Lebanese varieties is realised as /ʔik.ta:b/ in other Lebanese varieties. The fact that initial clusters are allowed in both Moroccan and Lebanese is believed to have caused initial gemination to be phonotactically possible in both dialects, a case that does not hold true for other Arabic dialects, such as Egyptian.

[9] It is worth noting that linguists such as Hall (2013) regard vowel deletion to yield final CC clusters optionally in Lebanese, rendering /harabt/ as phonotactically possible. However, these optional cases might be indicative of speaker-specific variations within the same dialect, as noted by one of the thesis examiners.

To summarise, despite these types of phonological differences between Moroccan and Lebanese, which are attributed to the different syllable structures in both dialects, there exist common phonological features between Moroccan and Lebanese dialects, the most important of which is that geminates are permissible word-initially and -finally. Therefore, it is important to investigate whether and how the phonological differences between these two Arabic dialects are reflected in their acoustic-phonetic and articulatory realisations of gemination. This is addressed in the gemination production chapters of this thesis (Chapters 6 and 7). More on the phonotactics that make initial and final gemination permissible in Moroccan and Lebanese Arabic is clarified in the two subsequent sections.

3.1.5. Phonotactics of initial gemination in dialectal Arabic

Word-initial gemination in dialectal Arabic occurs as a result of a morphological process. In a few Arabic dialects such as Moroccan and Lebanese, the prefix ʔ*al-*, which is the definite article in MSA, is reduced to *l-*. When the /l/ of the definite article is followed by a coronal phoneme in the stem word, the /l/ assimilates completely to that phoneme, resulting in gemination word-initially (see Table 3.3 for examples). Only coronal consonants can be geminated word-initially in Moroccan and Lebanese. This is why the focus of this PhD project is on the production and perception of the coronal stops /t/ and /d/ and their respective geminated counterparts /t:/ and /d:/. Table 3.3 gives three examples for Jordanian Arabic dialect (Abu-Abbas, Zuraiq, & Abdel-Ghafer, 2011), which also apply to Moroccan and Lebanese dialects.

Table 3.3: Examples of word-initial gemination in Moroccan and Lebanese dialectal Arabic via total assimilation

Singleton	**underlying form for 'the' ___**	**Geminate**
/da:r/ 'a house'	l-da:r	/d:a:r/ 'the house'
/ti:n/ 'figs'	l-ti:n	/t:i:n/ 'the figs'
/sala:m/ 'peace'	l-sala:m	/s:ala:m/ 'the peace'

However, MSA does not allow word-initial consonant clusters or geminates. As a result, in MSA the /ʔa/ of the definite article *ʔal* is retained and only the /l/ assimilates, thus creating medial rather than initial gemination. Therefore, the examples above would be realised in MSA (see examples in Table 3.4):

Table 3.4: Examples of word-initial gemination in MSA via total assimilation

Singleton	**underlying form for 'the'____**	**Geminate**
/da:r/ 'a house'	al-da:r	/ad:a:r/ 'the house'
/ti:n/ 'figs'	al-ti:n	/at:i:n/ 'the figs'
/sala:m/ 'peace'	al-sala:m	/as:ala:m/ 'the peace'

Although word-initial gemination is phonotactically possible in Moroccan and Lebanese Arabic dialects but not in MSA, final gemination is allowed in both MSA and the regional dialects. However, it is much more frequent and productive in MSA than in these two dialects, where it is very rare and not linguistically productive. In the next section, we present the phonotactic and morphophonemic processes governing final gemination in MSA.

3.1.6. Phonotactics of final gemination in MSA

The condition under which word-final gemination occurs in MSA is through a process of phonotactically-conditioned morphological reduction, whereby a case ending is dropped in pre-pausal or utterance-final environments. *Case ending* or *case marking* refers to the inflectional suffixes 'a', 'i', 'u', 'an', 'in' and 'un', which are attached to the end of Arabic words to indicate the grammatical function of that word in a sentence. Interestingly, when the word occurs at the end of a sentence, the case ending is omitted to mark a pause. This means that two forms of pronunciation exist for the same word: pausal and non-pausal forms. Final gemination occurs only when words are pronounced in their pausal forms (see Table 3.5).

Table 3.5: Pausal and non-pausal forms of Arabic words

singleton		geminate	
pausal	non-pausal	pausal	non-pausal
/ʃa:b/ (v. "to become white-haired")	/ʃa:ba/	/ʃa:bb/ (n. "youngster")	/ʃa:bbun/
/man/ (pr. who)	/man/[a]	/mann/(v. p "to bestow a favour")	/manna/
/kul/ (v. imp "eat!")	/kul/	/kull/ (pr. "all")	/kullu/
/ʔalam/ (n. "pain")	/ʔalamun/	/ʔalamm/ (v. p "to affect")	/ʔalamma/

[a] It is important to note that some words in Arabic have an invariant case ending across different positions in the sentence. These include some pronouns (e.g., /man/ 'who', /kam/ 'how much'…) and some verb forms (e.g., the imperative).

In Classical Arabic, dropping the marking of case was restricted to pre-pausal words occurring immediately preceding a full stop or a comma. However, MSA speakers have gradually tended to pronounce words without their case ending even when the words are in utterance-initial or -medial, or in isolated form. This is done either unintentionally due to their limited knowledge of the Arabic grammar, or intentionally to imitate the pronunciation of their

local spoken dialects where case ending is not a common feature. In this regard, Belazi (1984) argued that case ending is gradually disappearing from MSA, not just when the speaker pauses at the end of an utterance, but also when there is no pause. This makes the pronunciation of case ending somewhat flexible and optional. If the pronunciation of suffixes that mark case was mandatory for words at all positions of the utterance, even pre-pausal, then word-final consonantal gemination would not occur in MSA. Thus, final gemination is not a purely phonological distinction, but rather is a result of morphophonemic and sociolinguistic processes. These complications affect the functional load of final gemination, which in return raise the empirical question of. whether Moroccan and Lebanese speakers of MSA maintain the geminate/singleton contrast in production.

3.1.7. Summary

Gemination occurs word-initially, -medially and -finally in Arabic and its vernacular dialects, specifically the Moroccan and Lebanese dialects. However, peripheral gemination is quite understudied. Because there is little knowledge in the literature on peripheral gemination in Arabic, it remains unclear whether gemination occurring at the edges of the word remains contrastive. This uncertainty over the contrastiveness and maintenance of peripheral gemination makes it necessary to investigate its production by native Arabic speakers before we can test how successful non-native listeners are at perceiving the contrast.

In the next chapter, we review the relevant literature on non-native perception of gemination with a focus on research involving the target listener groups of the perceptual component of this PhD project.

CHAPTER 4: CROSS-LANGUAGE AND CROSS-DIALECT PERCEPTION OF GEMINATION

In the previous chapter, we showed how the position of the gemination contrast at the periphery of the word might lead to phonological and phonetic neutralisation of the contrast. The position of the geminate/singleton contrast in the word could also be informative about how likely the contrast would be detected by listeners. The most perceptually discriminable position of gemination is word-medial, particularly intervocalic (i.e., flanked by two vowels). Studies have shown that even speakers of languages with no geminate-singleton contrasts can detect durational differences in consonant closure for non-native[10] gemination contrasts that occur word-medially (e.g., Huggins, 1972; Pajak, 2009; Pickett & Decker, 1960). In this position, the vowels on either side of the consonant are thought to provide a suitable context for consonantal duration differences to be most easily perceived. This context, however, is not presented in word-initial gemination, which could explain why the geminate/singleton contrast is less discriminable at this word-position. For example, even native listeners identify word-initial geminate and singleton stops just slightly above chance in the Swiss German dialect of Thurgovian (Kraehenmann, 2001). Moreover, word-initial gemination in Berber was the most challenging for native listeners to perceive compared to word-medial and word-final gemination (Ridouane & Halle, 2010). In a similar fashion, word-final gemination, particularly in utterance-final position, is characterised by reduced acoustic cues as due to the absence of post-consonantal vowel. This very likely to make the contrast less discriminable. In a study

[10] The term "non-native" is used throughout this thesis to refer to a language that the speaker/listener did not grow up speaking. Moroccan and Lebanese dialects are mutually unintelligible, i.e., different languages. Therefore, the stimuli produced by native Moroccan speakers would be considered non-native for Lebanese listeners

comparing American English, Russian and Italian listeners' categorisation of non-words containing geminate/singleton distinctions across different word positions, a disadvantage was found for word-final relative to word-initial and especially word-medial gemination contrasts; final contrast was categorised the least accurately (Dmitrieva, 2018).

Apart from the poorer perceptual cues that distinguish peripheral gemination compared to medial gemination, we argue in this thesis that listeners' familiarity with the contrast at one position or lack thereof could also predict the likelihood of their success in perceiving the contrast at the same position in a non-native language. Therefore, the three groups of listeners which are the focus of this perceptual investigation were chosen to answer one key question: does native exposure to the geminate/singleton contrast in one position of the word help discrimination and categorisation of the contrast in other positions where it has not been experienced in the native language(s)?

We review in Section (4.1) perception studies on gemination involving L1 English listeners, and we argue that this group of listeners is more likely to attend to differences in consonant duration word-medially because the acoustic cues are more robust and due to their phonetic experience with fake gemination, but would find word-initial and -final gemination rather challenging due to acoustic factors that obscure their perceptibility and their lack of phonetic experience with fake gemination in peripheral positions. In Section (4.2), we turn to the second group of listeners in our study, Lebanese heritage listeners, and we explain how they differ in terms of exposure and familiarity with positional gemination compared to L1 English listeners. We also show how studies on perception of gemination have long overlooked heritage listeners of a target language. We then tackle the effect of dialectal differences on the perception of speech. Finally, we review studies on the perception of gemination by native Arabic listeners (4.3). In the final section of this chapter, we review four major non-native

speech perception frameworks (4.4), starting with the Perceptual Assimilation Model (4.4.1), the Speech Learning Model (4.4.2), the Second Language Linguistic Perception Model (L2LP) (4.4.3) and the Native Language Magnet model (NLM) (4.4.4), followed by a comparison among them (4.4.5).

4.1. Native English listeners

Although consonantal duration is non-contrastive in English, native speakers can detect durational changes across word boundaries. It was found that when the closure interval in *topic* was lengthened to approximately 95 ms, listeners reported hearing *top pick* (Pickett & Decker, 1960), i.e., a shift from one word to two words, introducing a perceived word boundary. Subsequently, in another study, the stop /p/ was judged to be long even when its closure was lengthened by only 30ms (Huggins, 1972). Thus, the studies agree on the fact that native speakers of English do attend to the difference in duration between the singular segment and its double to differentiate minimal pairs such as *topic* and *top pick*.

The finding that L1 English speakers have the ability to detect phonetic duration changes in the word-medial position is partially supported by studies on the acquisition of gemination in Japanese by L1 English learners. English listeners were naïve to Japanese were capable of discriminating the geminate/singleton contrast only when their attention was drawn to it (Hisagi & Strange, 2011). In other studies, Japanese-naïve English listeners were reported to identify geminates as allophones of singletons in Japanese (Hayes-Harb, 2005).

L2 experience has been shown to play an important role in the perception of L2 gemination contrasts. To investigate the development of native English speakers' ability to encode Japanese consonant duration contrasts lexically, English listeners who had never learned Japanese and those who had been studying Japanese for a year performed listening and

production tests (Hayes-Harb & Masuda 2008). The target stimuli consisted of nonce words with a (C)VCV structure (vowel-consonant-vowel, with an optional initial consonant), with the medial consonant being either a singleton or a geminate. The words were assigned a meaning and matched with a picture. All the participants were asked to memorise the words and their matched pictures. The results from the listening task showed that inexperienced listeners were able to distinguish between singletons and geminates, but their performance was less accurate than native Japanese speakers. Importantly, the experienced L2-Japanese learners were more successful than the inexperienced listeners in discriminating the contrast, suggesting that experience in learning the foreign language does facilitate perception of the contrast. However, it should be noted that this study concerned gemination word-medially, which is easily discriminable not only due to the flanking vowels, but also because sch gemination occurs phonetically across word boundaries in English as explained above.

Compared to perceptual investigations of word-medial gemination, fewer studies have examined non-native perception of word-initial and -final gemination contrasts. In one of the few published studies (Pajak, 2009), a native Moroccan speaker produced phonotactically possible nonce words containing the geminates (s:) and (z:) in four different environments: medial intervocalic, medial consonant-adjacent, initial vowel-adjacent and initial consonant-adjacent (no final geminates were examined). These utterances were used as stimuli to investigate the perception of Moroccan geminate consonants by naïve English speakers. Participants discriminated the initial vowel-adjacent contrast more poorly than medial-intervocalic but better than medial consonant-adjacent, and their discrimination was poorest for the initial consonant-adjacent context. Another study compared how nonce words containing word-initial and word-final length contrasts produced by a Russian speaker were perceived by native Russian listeners, Russian-naïve listeners of the medial-geminating language Italian and Russian-naïve English listeners (Dmitrieva, 2012). All groups

discriminated word-initial geminates better than word-final geminates. Importantly, the Russian group and the Italian group performed significantly better than the English group. Therefore, despite the limited number of perceptual studies on peripheral gemination with naïve English listeners as participants, it is fair to say that this group of listeners find peripheral gemination more challenging than word-medial gemination.

Unlike naïve English listeners, heritage Lebanese dialect listeners have native language knowledge of gemination not only in medial position but also in word initial position. We explain in the next section how this knowledge might help such heritage listeners with perception of initial but possibly also final gemination.

4.2. Heritage listeners

It is important first to identify who can be considered a heritage speaker/listener and what criteria define a heritage language. The most comprehensive definition in the literature is the following:

> A language qualifies as a heritage language if it is a language spoken at home or otherwise readily available to young children, and crucially this language is not a dominant language of the larger (national) society… An individual qualifies as a heritage speaker if and only if he or she has some command of the heritage language acquired naturalistically… although it is equally expected that such competence will differ from that of native monolinguals of comparable age. (Rothman, 2009: 159)

Therefore, heritage speakers/listeners are native speakers/listeners of the dominant language of the society they live in (e.g., English for Lebanese heritage speakers in Australia) who were also exposed to their first language (Lebanese Arabic dialect for Lebanese heritage

speakers) at home at an early age, but due to mainly limited exposure to their heritage language, their proficiency in that language is less proficient compared to native speakers from the country of origin. So, despite their heritage language being the first language they have learned, it is often regarded as the weaker language (Polinsky, 2018b). What makes heritage listeners an especially interesting group to investigate is that their linguistic competence in their first language matches neither that of native listeners nor that of late second language learners. As such, it would be plausible to expect them to perform less proficiently than native listeners but better than adult second language learners. This scenario has been upheld by some research studies comparing heritage language listeners' performance to that of native and second language listeners (e.g., Chang, 2016 & Lee-Ellis, 2012).

The most common account regarding heritage listeners' performance is that they have an advantage in both producing and especially perceiving aspects of their heritage language, relative to late second language learners. In fact, some studies have reported that heritage listeners can have a native-like perception of their first language. For example, heritage listeners of Korean were tested on their perception of Korean unreleased stops, and they performed exceptionally well (Chang, 2014). Similarly, Spanish heritage listeners' perception of Spanish lexical stress in various prosodic contexts was compared to that of native Spanish listeners and English learners of Spanish (Kim, 2019). Heritage Spanish listeners not only outperformed English L2 learners, but their success rate was comparable to that of native Spanish listeners. Since Lebanese allows initial gemination, in the present study we hypothesised that heritage Lebanese listeners would outperform naïve L1 English listeners in discriminating and categorising initial gemination contrasts in Moroccan Arabic.

Studies on Arabic heritage speakers/listeners are few and far between. In a study investigating knowledge of plural formation in the participants' vernacular variety of Arabic,

speakers showed a good understanding of the plural formation rules, but they still did not reach native-like performance (Albirini & Benmamoun, 2014). As far as gemination is concerned, the very few studies involving heritage speakers focused solely on production of gemination, disregarding perception. In a study on gemination in Farsi, in which the objective was to assess whether knowledge of consonantal gemination contrasts weakens across three different generations of Farsi-English-speaking individuals living in Canada, heritage speakers produced shorter geminates than their native Farsi counterparts. Nonetheless, heritage speakers still managed to maintain nativelike durational differences between geminates and singletons in Farsi. Unlike Farsi heritage speakers, however, Italian heritage speakers in Germany produced larger geminate-singleton duration ratios than their native counterparts in Italy (Kupisch, Einfeldt, Lleó, & van de Weijer, 2016).

It is crucial to investigate the perception of gemination in Arabic by Lebanese heritage listeners because their performance can offer insight about the role of gemination in dialectal Arabic.

4.3. Native Arabic listeners

Compared to Lebanese heritage listeners, native Lebanese listeners are not only fluent in the Lebanese vernacular of Arabic but have a very good command of MSA, the variety of Arabic where final gemination is permissible and productive. The bulk of research tackling the perception of the geminate/singleton contrast has tested listeners of languages in which gemination is not phonemic. The limited number of studies addressing the ability of native listeners of geminating languages to discriminate and categorise gemination in their native language reflects an underlying assumption that native listeners of a language should be fully adept at perceiving all phonemic contrasts of their language. Evidence supporting this

assumption comes from a number of studies. For example, native Italian listeners performed near ceiling at identifying Italian consonant length categories (Tsukada & Hajek, 2019). In another study, Kuwaiti Arabic speakers demonstrated high accuracy in discriminating the Arabic geminate/singleton contrast and outperformed English L1 learners of Arabic (Rinaldi, 2014). These results are not surprising, given consonantal length is contrastive and highly functional in these languages, especially when the contrast occurs word-medially.

Research focusing on the perception of gemination at the periphery of the word is limited. One of the few prior studies on native listeners' perception of word-final gemination found that native Tashlhiyt Berber listeners performed highly accurately on their native contrasts even for word-final voiceless stop contrasts (*fit-fitt* and *hat-hatt*) (Halle & Ridouane, 2016). However, an earlier study found that while Tashlhiyt Berber listeners performed near ceiling for word-initial voiced stops and fricatives, their performance on word-initial voiceless stop contrasts was notably less accurate, between 55% and 70% correct. Moreover, even with the presentation of audio-visual speech information, their performance was not enhanced considerably (Ridouane & Halle, 2010). The authors attributed the native listeners' difficulty to the fact that closure duration of voiceless stops in utterance-initial position is not audible in the acoustic signal, despite it being realised with differences in articulatory closure durations as measured by EPG.

Discrimination of word-initial gemination has been reported to be difficult in several studies (e.g., Ladefoged & Maddieson, 1996; Pajak, 2010). However, even voiceless stop contrasts in word-initial position were discriminated very successfully by native speakers of some languages. For example, native listeners of Cypriot Greek demonstrated ceiling level accuracy in discriminating word-initial voiceless stop geminates from singletons in both utterance-initial and utterance-medial position (Muller, 2003). This difference in performance

between Cypriot Greek and Tashlhiyt Berber listeners, for example, may be attributable to the relevance of different acoustic cues to the perception of gemination word-initially in different languages, and how reliably native listeners rely on those cues to discriminate the contrast. That is, on the one hand, if closure duration is the only reliable cue to the distinction between geminates and singletons in a particular language, then discrimination of utterance-initial voiceless stop contrasts by native listeners of that language is expected to be relatively poor. On the other hand, if the geminate-singleton contrast is realised with other reliable acoustic cues besides closure duration in another language, native listeners of that language should perform relatively well even with utterance-initial contrasts, especially if they are attentive to those secondary cues. In this regard, Abramson (1991; 1999; 2003) found that in Pattani Malay, a language where initial gemination is contrastive, secondary cues including burst amplitude in combination with F0 of the following vowel can enhance the perception of utterance-initial voiceless stop geminate/singleton contrasts when closure duration itself is not conveyed acoustically. Comparable results were found in Kelantan Malay (Hamzah, 2013), whose native listeners were shown to be sensitive to burst amplitude and vowel F0 cues. These secondary cues thus contributed to successful identification and discrimination of initial gemination in utterance-initial position, regardless of the voicing type. Therefore, investigating native Lebanese listeners' perception of voiceless versus voiced stop gemination, word-initial in particular, would reveal how sensitive they are sensitive to the acoustic cues contrasting gemination in Moroccan.

Another important reason for investigating the performance of native Lebanese listeners in discriminating Moroccan stimuli is to account for the role of dialectal differences in the perception of gemination. The vast majority of research on perception of regional dialectal differences has been conducted almost exclusively on vowels, and mostly in English, French or Spanish. Perception of dialect variation in geminated consonants, and by extension

suprasegmental properties of consonants, is lacking. We have not found any prior cross-dialectal reports on perception of gemination, either in Arabic or in other geminating languages.

There is general agreement that listeners are less effective in dealing with speech in a non-native regional variety of their native language than in their own variety (see Adank et al., 2009). However, these inter-dialectal variations have sometimes been compared to those between a speaker's native language and non-native languages. This is particularly relevant for Moroccan and Lebanese, which despite being both Arabic vernaculars are mutually unintelligible. It would be appropriate to characterise the impact that such dialectal differences may have on speech processing in a non-native variety of the listeners' language, especially with regards to gemination.

Although listeners of certain dialects manage to accommodate perceptually to the variation of another regional dialect, some studies have demonstrated that listeners do not always surmount difficulties in perceiving contrasts produced by speakers of another, unfamiliar regional dialect (Dufour, Brunellière & Nguyen, 2013). Therefore, it would be crucial to study the impact of dialectal differences between Moroccan and Lebanese on the perception of gemination in Moroccan by Lebanese native listeners. Their performance will also be analysed in light of theoretical principles of non-native and second-language speech perception models. The primary models of non-native speech perception are reviewed in the next section.

4.4. Theoretical frameworks of non-native speech perception

Several frameworks have been proposed to provide an account of, and draw generalisations about, the universal mechanisms that underlie the effects of language

experience on speech perception. They do this by exploring the phonetic and phonological differences and similarities between the target language and the listener's native language, or L1. The most widely cited and used of these frameworks are the Perceptual Assimilation Model (PAM: Best, 1995; Best & Tyler, 2007), the Speech Learning Model (SLM: Flege, 1995); the Native Language Magnet model (NLM: Kuhl & Iverson, 1995) and the Second Language Linguistic Perception model (L2LP: Escudero, 2005).

4.4.1. The Perceptual Assimilation Model: PAM

The Perceptual Assimilation Model (PAM: Best, 1994, 1995) is based on the direct-realist view of speech perception, which presumes adults perceive non-native phones via detecting dynamic properties of articulatory gestures. That is, listeners assimilate non-native phones into native categories on the basis of the perceived articulatory similarities/dissimilarities to native phones and contrasts. PAM examines all possible outcomes in the categorisation and discrimination of non-native speech phones. Particularly, the PAM framework focuses on the perception of non-native contrasts (i.e., two minimally distinctive phonemes in the non-native target language), as opposed to single phones. PAM proposes that the easiest non-native distinctions to perceive are those that map onto two sufficiently distinct phonetic forms in the learner's L1. Non-native categories that are completely outside the learner's L1 inventory fail to be perceptually assimilated as speech sounds and are consequently heard as non-speech sounds, while non-native distinctions that map onto a single segment in the L1 pose the greatest difficulty and their discrimination gets more difficult as the learner's L1 fluency progresses from early childhood to adulthood. Thus, when a non-native listener attempts to categorise and discriminate a contrast, a range of patterns of perceptual assimilations are expected, dependent on perceived articulatory similarities between the non-native phones and the phonemes of the listener's L1:

(i) **Two-category (TC),** in which the contrasting non-native phonemes are assimilated into different native phonological categories. Discrimination is predicted to be excellent because the two categories are distinct from each other in their native phonological space.

(ii) **Category-Goodness Difference (CG),** in which the contrasting non-native phonemes are assimilated to the same native phonological category, but differ in that one phoneme is considered a good exemplar while the other is considered a poor exemplar. Depending on the degree of goodness difference this type of discrimination can range from moderate to good. This reduced level of discrimination is due to the weak phonological differences between the contrasting non-native phonemes. So, the listener relies on fine-grained phonetic features to discriminate the contrast.

(iii) **Single-category (SC),** in which both non-native phonemes are assimilated to the same native phonological category and are equally different from (or similar to) the native phoneme. Discrimination is predicted to be poor (difficult) because listeners are unable to detect even honetic differences between the two categories.

(iv) Uncategorised-Categorised Assimilation (UC), in which one non-native phoneme is assimilated to a native phonological category and the other falls within native phonetic space but not cleanly into a single native phonological category. Discrimination is predicted to be very good, although it may vary depending on the perceived phonological overlap between the uncategorised and categorised phone (see Faris, Best & Tyler, 2016, 2018).

(v) Uncategorised-Uncategorised assimilation (UU), in which both non-native segments fall within the listener's native phonetic space but fail to be assimilated clearly to a

single native category. Discrimination is expected to be poor to very good, depending on perceived phonological overlap, and/or the phonetic distance between the two phonemes.

(vi) Non-assimilable (NA), in which both non-native phones fall outside the native phonetic space and are heard as non-speech sounds. Discrimination is predicted to be good to very good. This depends on the acoustic distance between the two phonemes.

PAM generally predicts TC to be the easiest to discriminate, followed by both CG and UC, while SC is expected to be the most difficult to discriminate. UU and NA assimilations cannot be easily ranked because their discrimination accuracy is contingent on the perceived similarity between the contrasting phones. Recent works have proposed more accurate predictions for UC and UU based on an approach of identifying the presence of perceived phonological overlap (Faris, Best & Tyler, 2016, 2018). Accordingly, excellent discrimination is expected for non-overlapping contrasts, followed by partially overlapping, and then completely overlapping contrasts. In line with this argument, we hypothesise in this study that the differential discriminability of the geminate/singleton contrast can be explained in terms of the amount of perceived phonological overlap involved in categorising the contrast.

Two common perceptual tasks are carried out to test PAM's predictions: categorisation and AXB tasks. The AXB discrimination task is used to determine potential difficulties with non-native contrasts. On each AXB trial, participants listen to three stimuli and are asked to decide whether stimulus X is from the same category as stimulus A or B. The task does not impose heavy memory load and does not allow the participants to respond on the basis of low-level auditory-acoustic differences and similarities (Hallé et al., 2016). In conjunction with an AXB task, a categorisation task is used. The categorisation task is usually a simple forced choice procedure. The participants are asked to categorise each token presented separately, judging whether it is a long or a short consonant, and then to rate the goodness-of-fit of the

token to the chosen category. By determining assimilation patterns, the categorisation task allows the interpretation of the discrimination performance. For this reason, an AXB task along with a categorisation task are employed in order to test the discrimination and categorisation of gemination in this PhD project as well as the predictive power of PAM (Chapter 8).

Unlike PAM, which was originally designed to account for non-native speech perception in naïve listeners, other L2 frameworks, such as the Speech Learning Model (SLM: e.g., Flege, 1995) and the L2 Linguistic Perception (L2LP: Escudero, 2005) models, were advanced mainly as accounts of L2 learners' phonetic acquisition over time. In the next two sections, we review the fundamental principles of these models.

4.4.2. The Speech Learning Model: SLM

SLM was designed to account for the learnability of L2 phonemes, or segments, by L2 learners. It specifically addresses variations in the extent to which L2 learners manage or fail to perceive and produce L2 phones accurately, relative to native speakers (Flege, 2005). SLM postulates that speech segments are specified as phonetic categories. The latter are defined as "long-term memory representations" in which "language-specific aspects of speech sounds are specified" (Flege, 1995: 239). Thus, SLM is concerned mainly with the phonetic, rather than the phonological aspects of L2 speech.

SLM argues that L1 and L2 share a common acoustic-phonetic space in the L2 speaker/listener, in which both languages mutually influence each other (Bohn & Flege, 1992). That is to say, the L1 phonetic system influences how L2 segments are perceived and produced. L1, thus, acts like a sieve for L2. However, L1 phonetic categories are not static, because eventually the interaction between L1 and L2 sounds causes L1 as well as L2 categories to change over time.

One central proposal of SLM is that the deviation in production of L2 segments from that of native speakers is caused by inaccuracy in L2 perception. In other words, the accuracy with which segments are perceived in L2 dictates how those L2 segments will be produced. More importantly, success in establishing a new L2 category depends greatly on the perceived L1/L2 similarity or dissimilarity.

According to SLM, when L2 phones are perceived as dissimilar from existing L1 categories, they will be distinguished from L1 categories, and will be easy for learners to discriminate from L1 phones. Consequently, new L2 categories are likely to be established. The established L2 category and the closest L1 category are predicted to dissimilate in order to minimise perceptual confusions in the shared L1-L2 phonetic system. This causes the production of one or both L1 and L2 phonetic elements to deviate from production obtained from monolinguals of each language.

SLM proposes that when an L2 learner fails to notice the difference between an L1 and L2 category because they are too similar (phonetically equivalent), formation of a new L2 category is highly unlikely. As such, the L2 phone is equivalence classified to the most similar L1 category, and the two categories form a "composite" (Flege, 2005). In this case, SLM further predicts that this L2 segment will initially be produced in the same way as the L1 segment, whereas the latter will shift in production towards the L2 segment. Across the learning process, both L1 and L2 segments will continue to be produced identically to each other.

Flege (1995) holds that the same processes and mechanisms children use to acquire their L1 continue to be used even in adulthood. This means that adult learners themselves have the capability of establishing new categories. However, SLM draws a distinction between the two groups of learners by positing that adults are less likely to form phonetic categories for L2 phones than children, because children's L1 attunement is not fully complete and thus allows

the easier formation of new L2 categories. Next, we provide a succinct summary of the third L2 framework, the Second Language Linguistic Perception Model (L2LP: Escudero, 2005), which like SLM focuses on L2 learning effects in perception.

4.4.3. The Second Language Linguistic Perception model: L2LP

L2LP is based on the Linguistic Perception Model (LP). The aim of the L2LP is to describe, explain and predict the entire process of acquisition in L2 speech perception. L2LP proposes that the learner goes through three phases in the learning process: starting with an initial state, a learner then faces a learning task and goes through a developmental phase, and ultimately reaches a final end state. These three phases are then broken down into five methodological "ingredients". Figure 4.1 represents the three phases and the five ingredients (Escudero, 2005: 95):

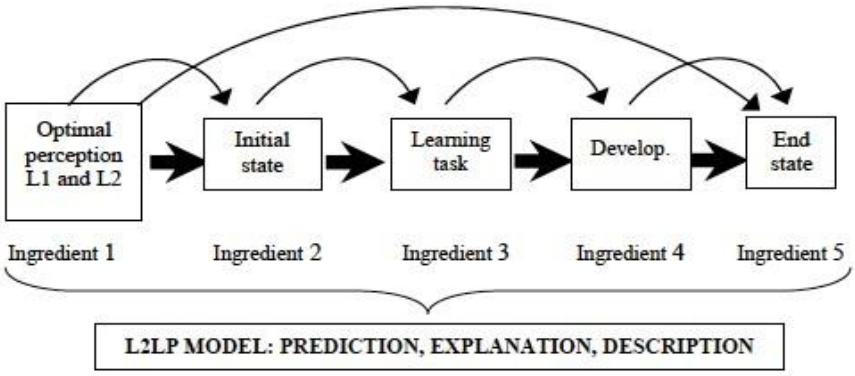

Figure 4.1: Steps and ingredients of the Second Language Linguistic Perception Model (L2LP: Escudero, 2005).

The first ingredient, optimal L1 and target L2, requires a detailed description of how auditory cues are weighed in the L1 and the target language or L2. The second ingredient, the initial state, refers to the starting phase in learning an L2 during which a learner has no knowledge of this new language. At this point, where the L1 is very influential, listeners will use the phonetic and phonological categories of their L1 to process the sounds of L2. The third

ingredient refers to the learning task that results from the differences between the perception of L2 in its initial state, when it is achieved only via the L1, and that of the target L2. Indeed, L2LP postulates that an exhaustive analysis of the inter-language differences between the optimal perception of the L1 and the target perception of the L2 provides an adequate description of the L2 learning task (Escudero, 2005: 105). The fourth ingredient, development, emphasises that in order to accomplish the loose learning of an L2, the learner needs to either create new perceptual mappings, which leads to new phonological representations, or to adapt the already existing perceptual mappings (ibid: 109). The fifth and final ingredient, the end state, indicates whether or not listeners have reached an optimal state of perception of the L2.

These five ingredients lead to three scenarios that can occur during the perception of L2 sounds. The first scenario, called the "NEW scenario", occurs if learners are confronted with a phonological category in the L2 that does not exist in the L1. For example, Spanish-speaking (beginner) learners who have only /i/ and not /ɪ/ in their language equate the two English phonemes /i/ and /ɪ/ with this single L1 category, /i/. This scenario describes the situation considered to be the most difficult for learners to overcome, since the task they must perform is to create a new phonological category for the new sound that will enable them to separate the two L2 phonemes into two distinct phonological units. The second scenario, the "SUBSET scenario", occurs when learners face an L2 phoneme that has similarities to more than one phonological category in L1. L2LP illustrates this situation by the example of native Dutch speakers learning Spanish using the three L1 categories /i/, /ɪ/ and /ɛ/ to classify the two Spanish vowels /i/ and /e/. In this situation, the Spanish /i/ can be categorised either as /i/, or as /ɪ/, and Spanish /e/ either as /ɪ/ or as /ɛ/. The "SUBSET" case scenario is considered moderately difficult for learners because the task they have to perform is to move and reduce their phonological categories. The last scenario, "SIMILAR scenario", is manifested when learners are faced with two L2 phonemes that have a phonological equivalence in their L1, but

which vary phonetically. L2LP illustrates this situation citing the example of the acquisition of French-Canadian (CF) phonemes /ɛ/ and /æ/ by English-speaking people in Canada (CE), who also have these two phonemes in their language but pronounce each with different phonetic details than the French versions. In this situation, beginner CE speakers perceive the CF /ɛ/ as /ɛ/, but perceive the /æ/ of CF as either /ɛ/ or as /æ/. The "SIMILAR" scenario is considered to be the least difficult for learners because the task they have to perform is only to slightly change their phonological boundaries to achieve optimal L2 perception.

4.4.4. The Native Language Magnet model: NLM

Next, we review the Native Language Magnet model (NLM) which, unlike the three frameworks we reviewed above, does not consider L2 learning. Similar to the other three frameworks, the Native Language Magnet (NLM) recognises the influence of L1 on perceiving speech of a foreign language (Kuhl & Iverson, 1995). The model describes the interactions between infants' auditory systems and the acquisition of language-specific categories. According to this theory, our ability to discriminate different phones is influenced by our linguistic environment from a very young age. NLM proposes that at birth, infants have the innate ability to perceive all the natural boundaries of any language that underlie the differences between phonetic categories. NLM holds that the first representation of speech in children is entirely auditory, not yet linguistic. Nevertheless, this representation includes various visual, motoric and acoustic properties of the ambient language. Speech perception is thus considered to be multimodal and influenced by the experience the baby gains as they learn the phonemic categories of their mother tongue. Moreover, because the perception of speech is organised by increasing native listening experience, perception gradually influences the production of speech in children.

NLM focuses on the shift that occurs around one year of age, when infants' capability to discriminate among non-native phones gradually declines. NLM attributes this transition to increasing neural commitment to the native language. According to NLM, the auditory space is divided into psychoacoustic boundaries corresponding to the phonemic characteristics of each language, which is referred to as a "sound map". Infants as early as six months categorise phones into neural sound maps through exposure to speech in their native language. This neural mapping of experienced native phones enables them to create abstract phonological categories specific to their native language that act as prototypes for a particular phoneme. The acquisition of a language requires a deformation of the perceptual space of the phonetic categories, leading to a reduction in perceptual distance around the prototype, called the "perceptual magnet effect". NLM explains that the perceptual reorganisation of mental representations in the baby can occur either at a late stage of speech perception (i.e., after creation of categories) or during the early development of his auditory system.

NLM proposes the concept of the "perceptual magnet effect" to describe the effect of the prototypical instances of native categories on the perception of speech. Once a phone category is created in the brain, it starts to serve as a "magnet" for non-native phones that are phonetically similar to the native categories. Instances of phones that are considered good exemplars of the native category are rather poorly discriminated compared to phones that are weak exemplars, because the closer a phone is to the prototype of a native category, the more likely it will be attracted to that prototype. Although this facilitates native language perception, it has the opposite effect on non-native speech perception. The challenge is that if a phone of the L2 is relatively close to a perceptual magnet, prototype, of the L1, it will naturally tend to be equated with the category of the L1. This has been confirmed by an experiment conducted by Kuhl and Iverson (1995), in which adults and 6-month-old babies were tasked to discriminate between synthesised vowels grouped into prototypical and non-prototypical

instances of the English vowel /i/ The results showed that adults and 6-month-old babies are not able to correctly discriminate the acoustic tokens that are very close to those of the prototypical vowel /i/ (as judged by adults), and the rate of discrimination between vowel pairs is significantly better between non-prototypical vowels.

Next, we compare among the four frameworks. We highlight some of their similarities and differences as well as what makes a particular model more suited to explain the findings of the present study.

4.4.5. Comparison among the frameworks

All four frameworks share one core, namely recognition of the central influence of the L1 on the perception of non-native speech segments. All assume that the perception of non-native segments is essentially dependent on the phonetic properties and the perceived similarities (and differences) of the phonemes in the L1 and the target language phonological systems. The main differences among these models lie in the disparate goals for which they were proposed. While SLM and NLM focus on the perception of isolated phones, L2LP and PAM predict how pairs of phones (contrasts) are categorised into native categories.

The predictive power of all these models has been tested extensively by studies on non-native perception. Although NLM correctly predicts some non-native phones as perceptually challenging, it does not clearly predict why certain other phones are easy for non-native listeners to perceive (Baese-Berk, 2018). On the other hand, the primary limitation of L2LP is that it has been mainly employed in perception of non-native vowels. The models that provide accurate predictions on perceptually difficult versus perceptually less challenging non-native consonant categories are PAM and SLM. Both PAM and SLM account for the degree of success listeners will have in perceiving non-native phones and contrasts. When a non-native

or L2 segment is too similar to an L1 category, it will be "perceptually assimilated" (PAM) to the articulatorily most similar L1 category, or equivalence classified (SLM) to the phonetically/acoustically most similar L1 category, and as a consequence, shifts in production are expected to happen. Nonetheless, PAM and SLM also differ in some points:

While SLM holds that L2 speech perception is based on the acoustic phonetic information of phone categories, PAM relies on the mapping of articulatory information. Certain key differences between SLM and PAM make the latter the most suitable theoretical framework for the present project. First of all, SLM operates mainly on the phonetic level and fails to adequately consider the influence of phonology, whereas PAM argues that non-native speech perception is based on the perception of articulatory patterning at both the phonological level and the phonetic level (Best & Tyler, 2007: 22). Another difference is that while SLM's main focus is on L2 learning, PAM focuses on naïve perception of unfamiliar perceptual distinctions. Finally, PAM focuses on contrasts, rather than only isolated phones as SLM does.

We conclude that PAM is the best suited of the four models for making predictions about the categorisation and related discrimination of non-native speech contrasts, because it only PAM addresses the importance of systematically considering the relationship between native phonetic and phonological factors in non-native speech perception.

CHAPTER 5: THE PRESENT STUDY

This research project takes up a comprehensive articulatory, acoustic and perceptual investigation of the word-initial gemination in Moroccan and Lebanese dialects and word-medial and -final gemination in Moroccan and Lebanese varieties of MSA, in order to explore the different acoustic and articulatory mechanisms employed by L1 Arabic speakers to produce these contrasts as well the perceptual relevance of these mechanisms for L1 English naïve, native and heritage listeners of Arabic.

5.1. Production experiments

In the production experiment, acoustic data were co-collected with ultrasound data from five native Moroccan and five native Lebanese bi-dialectal speakers. The speakers were recruited from the Moroccan and Lebanese communities living in the greater Sydney metropolitan area. To ensure a fluent command of both MSA and their respective Arabic dialect, only participants who were born and received education through secondary school level in Morocco and Lebanon were selected. Participants had lived in Australia for 1-4 years (Moroccan mean = 3.2 years, Lebanese mean = 3.4 years) at the time they participated in the experiment.

The objective of the production experiments was to determine how a combination of articulatory, tongue kinematics, and acoustic, durational and non-durational, parameters characterise the geminate/singleton contrast in all three positions of the word, i.e., word-initial, -medial and -final, and more importantly whether these acoustic and articulatory characteristics are indicative of a duration contrast, particularly at word periphery in Arabic and two of its

dialects. Another objective of this project is to identify whether and how two geographically distant Arabic vernacular dialects, Moroccan and Lebanese, and their local varieties of MSA may differ in production of gemination. Importantly, this PhD project tackles the debated question of whether the phonological and articulatory nature of gemination is a distinction based on articulatory force, rather than a distinction based on length/elongation.

Thus, the additional non-closure-duration correlates can be indicative of a more forceful articulation for geminates relative to singletons, which led some researchers to question whether the phonological and articulatory nature of gemination is a forceful articulation distinction, rather than a length/elongation distinction (Ridouane, 2007). This question is explored in this study as well as the possibility that gemination is both the result of a forceful articulation and elongation and that the balance between these two effects is language specific.

The production experiment was conducted using a combination of innovative kinematic and analytic methods. In order to obtain and preserve clear, high quality concurrent audio and articulatory data related to the production of the geminate/singleton contrast, high-quality acoustic data were co-collected with ultrasound data because ultrasound has a relatively smaller effect on speech output compared to other techniques, such as electromagnetic articulography (EMA), which normally requires sensors to be glued on the surface of the tongue. Co-collecting both types of data also allowed us to use the acoustic recordings from the production experiment as stimuli for the perception experiment.

5.2. Perception experiment

The perception component of this thesis consists of AXB discrimination and categorisation of initial, medial and final geminate/singleton non-lexical contrasts produced by Moroccan speakers in the production experiment. The participants in these perception tasks

were divided into three groups of non-native listeners: 1) naïve monolingual L1 English listeners with no exposure to Arabic and its dialects or any other languages with gemination contrasts, 2) Lebanese heritage listeners, who are fluent in Lebanese dialect but do not speak MSA, and 3) native Lebanese listeners, who are fluent speakers of Lebanese dialect and the Lebanese regional variety of MSA, i.e., bi-dialectal. The latter two groups are multilingual: they are also fluent in English as an early (Heritage Lebanese group) or later (native Lebanese group) L2. Because most perceptual studies investigated the sensitivity of native listeners to different phonetic cues to the geminate/singleton contrast in their native language, this PhD thesis was concerned instead with non-native perception of gemination. We did not test native Moroccan listeners because the three groups of non-native listeners were selected to test predicted differences in performance dependent on group differences in the nature and extent of their relevant phonological and/or phonetic experience with gemination.

The objective of this perception experiment was to investigate how successful these three groups of listeners are in discriminating and categorising the non-native Moroccan stimuli. Of particular interest is to investigate whether and how well the perceptual assimilation processes proposed by L2 speech perception frameworks, in particular the Perceptual Assimilation Model (PAM: Best, 1995), account for non-native perception of gemination, and how well the assimilation patterns predict the discrimination of the non-native Moroccan geminate/singleton contrast by the three groups of listeners.

We predict that the three groups' success in discriminating the medial, initial and final gemination contrasts in the Moroccan target stimuli would depend on their familiarity and experience with the contrast in each word position in their native language. In particular, we predict that phonological-experience with gemination (the two Lebanese groups) would yield better discrimination than phonetic-only experience with gemination (English group).

5.3. Rationale for the research

In terms of production, it is not yet known whether native speakers produce these contrasts using different mechanisms for word-initial versus word-final gemination nor whether they maintain the distinctiveness of the contrast at either peripheral position of the word. This research project is the first comprehensive study to investigate the distinctiveness of word-initial and word-final gemination contrasts in MSA and its regional dialects, exploring the word/utterance positional and dialectal differences in syllable structure, all of which could potentially impact the phonetic/articulatory realisation of the geminate/singleton contrast. This production investigation also provides the stimuli and the basis for the subsequent perception investigation with three groups of non-native listeners — naïve English, heritage Lebanese and native Lebanese speakers.

The three listener groups differ in terms of exposure and familiarity with gemination contrasts at different word positions in their native languages. Because gemination is not contrastive in English, English speakers are not familiar with phonological gemination. However, phonetic lengthening that arises across word boundaries in English gives the English listeners purely phonetic-level experience with gemination (called fake gemination) that is superficially comparable to word-medial gemination in geminating languages. However, English speakers lack even phonetic-level experience with initial and final gemination. Native Lebanese speakers have native phonological experience with initial and medial gemination contrasts in their vernacular dialect, and with final gemination contrasts in MSA, where they carry a relatively high functional load. The Heritage Lebanese speakers studied in this PhD project were raised in Australia and have native experience with initial and medial gemination in the Lebanese dialect that they learned at home as their first language. They had not studied MSA. With the exception of a very limited number words containing final geminated

consonants, mainly borrowed from MSA, geminated consonants do not occur word-finally in Lebanese dialect. Moreover, these rare MSA-borrowed geminates are not involved in final singleton/geminate *contrasts* in Lebanese. The Heritage Lebanese participants therefore lack native language experience with final gemination contrasts. Thus, investigating the three groups' categorisation and discrimination of Moroccan stimuli allows us to discern the relative contributions of higher-level phonological experience and lower-level phonetic experience of the native language on perception of final gemination *contrasts*. Moreover, the stimuli are non-native for all three groups of listeners because they were produced by Native Moroccan speakers, whose native Arabic dialect is unintelligible for Lebanese Arabic speakers. Therefore, testing Lebanese listeners' discrimination and categorisation of these Moroccan stimuli allows us to test hypotheses concerning the effects of native language experience on perception of a different geographically distant vernacular dialect and a different regional variety of MSA.

The broader objective of this PhD project is to strengthen our understanding of, and resolve controversy in the literature about, peripheral gemination. This production and perception investigation not only contributes to a better understanding of the relationship between production and perception of gemination in particular and non-native phonemic contrasts in general, but it will also potentially help Arabic language instructors to teach these contrasts, and L2 Arabic learners to learn them, more effectively. This objective stems from the idea that the more teachers are aware of the organisational principles underlying the production of these contrasts and how that directly affects their perception by L2 learners, the more attentive they will be to the challenges that these L2 learners face.

It is important to note that the three experimental chapters 6-8 were written and presented as stand-alone journal manuscripts. Therefore, the information presented in some

sections of these chapters is somewhat redundant with the introductory chapters and between experimental chapters.

CHAPTER 6. THE ACOUSTICS AND ARTICULATION OF INITIAL GEMINATION IN MOROCCAN AND LEBANESE DIALECTS

6.1. Abstract

This paper reports on a synchronous acoustic and articulatory investigation of word-initial gemination across two utterance positions (initial versus medial) in two disparate Arabic dialects, Lebanese and Moroccan Arabic. The results provide clear *articulatory* evidence that speakers in both dialects produce word-initial singletons versus geminates with significantly different closure durations even in utterance-initial position, where the voiceless stop contrast provides no *acoustic* evidence of closure duration per se. We also found that speakers differentiate initial geminate/singleton contrasts acoustically, however, along dimensions other than closure duration. For geminates the following vowel is longer, the stop burst amplitude is greater, and f0 at the beginning of the following vowel is higher, than they are for singletons. However, the two dialects weight the articulatory and acoustic cues differently. In Moroccan, articulatory closure duration is weighted more heavily than vowel duration, f0 and amplitude of the release burst, whereas Lebanese shows the opposite weighting. These acoustic and articulatory findings provide evidence that initial gemination contrasts are maintained in Moroccan and Lebanese. The presence of other acoustic cues even in cases where there is no acoustic evidence of closure duration suggests that lengthening of the closure in gemination may result from a more forceful/tense articulation for geminates than for singletons.

6.2. Introduction

Geminates, in any language where gemination is contrastive, are pronounced with a longer duration than singletons regardless of their position in the word, i.e., word-initially, word-medially, or word-finally. Previous studies on consonant gemination have focused primarily on the acoustic correlates of geminate/singleton contrasts. The results indicate unanimously that the relative duration of consonant closure is the only cue necessary for distinguishing the geminate/singleton contrast, especially in the most common context for gemination contrasts, which is the word-medial position (e.g., Bengali: Lahiri and Hankamer, 1988; Berber: Ridouane, 2010; Bernese: Ham, 2013; Buginese: Cohn, Ham & Podesva, 1999; Estonian: Engstrand & Krull, 1994; Finnish: Engstrand & Krull, 1994; Cypriot Greek: Tserdanelis & Arvaniti, 2001; Guinaang Bontok: Aoyama & Reid, 2006; Hindi: Ohala, 2007; Hungarian: Ham, 2013; Italian: Payne, 2005; Pickett et al., 1999). In many languages, the duration of the preceding vowel is also significantly affected by medial consonant gemination. Generally, the vowel preceding a medial geminate is shorter than the one preceding a singleton, as shown for 12 of 16 languages reviewed by Hamzah (2013), which include Bengali (Lahiri and Hankamer, 1988), Buginese (Cohn et al., 1999), Hindi (Ohala, 2007), Icelandic (Pind, 1995; 1999), and Italian (Payne, 2005). However, one major limitation of most investigations is that they have been designed to quantify phonetic cues that are intrinsic to the segmental level, without examining how suprasegmental factors could potentially affect realisation of the geminate/singleton contrast.

An important suprasegmental factor that could affect how gemination is produced in specific contexts is the contrast maximisation principle (Cho & Jun, 2000; Cho & Keating, 2001). This principle stems from the idea that prosodic domains are ranked hierarchically, with

certain domains being higher than other domains (Beckman & Pierrehumbert, 1986). One of the posited effects is that increased articulatory strengthening/enhancement is associated with each higher level in the prosodic hierarchy (syllable < word < smaller phrase < intonational phrase < utterance). That is, the prosodic position of a contrast determines how strong its articulation is. Articulatory strengthening can be described as an increase in spatio-temporal magnitude of one or more key articulatory gestures/actions. As such, consonants in higher prosodic domains have been found to be articulated with more extreme and longer constriction/closure durations than those in lower domains, thereby maximising their contrastiveness with other consonants (Cho & Jun, 2000; Cho & Keating, 2001).

One of the primary prosodic positions that has been shown to cause articulatory strengthening/enhancement is the initial position in a prosodic unit, i.e., the initial position of a syllable, word, and phrase, or utterance, with the effect increasing as initial position lines up across multiple prosodic units, i.e., greatest for utterance-initial (e.g., Pierrehumbert & Talkin, 1992; Cho & Keating, 2001). For example, coronal stop /t/ in English has a longer Voice Onset Time (VOT) phrase-initially compared to when it occurs in the phrase-medial position (Pierrehumbert & Talkin, 1992). Similarly, Electropalatography (EPG) investigations have shown that consonants are produced with a larger and longer lingual contact phrase-initially than -medially (e.g., Fougeron & Keating, 1997).

Two types of domain-initial strengthening may arise: syntagmatic and/or paradigmatic strengthening (Fougeron, 1999). Syntagmatic strengthening refers to articulatory/acoustic amplification of the relationship between a segment and its adjacent segments. For example, the difference between a consonant and the following vowel may be enhanced, i.e., the duration of the consonant is lengthened more at the expense of a more shortened post-consonantal vowel. Paradigmatic strengthening refers to articulatory/acoustic amplification of the

differences between a segment and another contrastive segment in the sound system of a language (geminate vs singleton). Syntagmatic enhancement of a domain-initial geminate consonant would be indicated by longer consonantal closure duration and higher f0 onset values in the post-consonantal vowel in utterance-initial. Paradigmatic enhancement, on the other hand, would be indicated by longer closure duration for geminate consonants as compared to singletons in utterance-initial. It is important to note that both paradigmatic and syntagmatic processes can co-occur in prosodic strengthening contexts.

Word-initial gemination is ideal for testing the principles of domain-initial strengthening, because unlike word-medial and -final gemination, word-initial gemination can occur in the initial position in an utterance. This allows us to compare the same word-initial gemination contrast at two levels of the prosodic hierarchy: the domain-initial level (the citation form of the word, which is utterance- and word-initial) versus the weaker domain-medial level (i.e., internal to a carrier phrase, which is word- but not utterance-initial). Word-initial gemination in citation form represents the highest level of the prosodic hierarchy. This provides an emphasised, carefully pronounced and prosodically strengthened environment, whereas word-initial gemination in medial position of the utterance is prosodically weaker because this context sits at a lower level of the prosodic hierarchy. In order to test these prosodic effects on the production of initial gemination and confirm that they are systematic, two regional dialects of Arabic are investigated: Moroccan and Lebanese, where initial gemination is contrastive.

Moroccan and Lebanese are the native languages of the majority of residents in Morocco and Lebanon, respectively. Although closely related, these two Arabic vernaculars exhibit considerable lexical, morphological and phonological differences from each other due their distant geographic distribution (Moroccan in the far northwest of Africa and Lebanese in the

far east of the Middle East). Previous research attributes the phonological differences between Moroccan and Lebanese to their variant syllable structures (for detailed reviews see Kiparsky, 2003; Watson, 2007). In particular, Arabic vernaculars are classified into two different syllable structure groups, namely, C-dialects (Moroccan: Harrell, 2004) and VC-dialects (Lebanese: Blanc, 1953 & Cowell, 2005). VC-dialects are those in which morphologically derived medial CCC clusters are syllabified as CV.CC (e.g., *yiktibu* 'they write'), whereas C-dialects are those in which CCC clusters are syllabified as CCC (e.g., *yktbu* 'they write'). It would be interesting to find out to what extent these phonological differences affect the realisation of initial gemination in both dialects. However, because of the lack of experimental studies on initial gemination in either dialects, the contrastiveness of initial gemination in Moroccan and Lebanese remains questionable.

Word-initial gemination is rare cross-linguistically (Ladefoged & Maddieson, 1996; Muller, 2001; Thurgood, 1993), and it has therefore been much less studied than gemination in other positions of the word. A longer closure duration for word-initial geminates compared to their singleton counterparts has led some researchers to propose that, just as with gemination word-medially, closure duration is the primary cue for word-initial gemination (Abramson, 1986; Lahiri & Hankamer, 1988). However, findings from other studies (e.g., Christodoulou, 2007) indicate additional non-closure-duration correlates play an important role in signalling the geminate/singleton contrast.

Experimental studies investigating non-temporal cues that signal the geminate/singleton contrast have found that the amplitude of the release burst distinguishes geminates from singletons in some languages. For example, In Kelantan Malay, a language where gemination occurs only word-initially, the release burst of voiceless stops is realised with a higher amplitude for geminates than for singletons, a result that was interpreted as the use of greater

articulatory effort to mark gemination (Hamzah, Fletcher & Hajek 2012). In addition, the amplitude of the following vowel can also be signal the geminate/singleton contrast. Abramson (1987) measured the average amplitude for vowels following word-initial stops in Pattani Malay and found that vowels following a geminate were greater in amplitude than those following a singleton. The same tendency was witnessed in Finnish (Doty et al., 2007). However, such as is the case with most secondary cues to gemination, this pattern is not consistent across all languages. Differences in vowel amplitude were not significant in Malayalam (Kraehenmann et al., 2000), Turkish (Lahiri & Hankamer, 1988). That inconsistency may, however, be related to prosodic domain. In some languages with only word-medial gemination, such as Japanese, the amplitude of the vowel preceding rather than the vowel following the consonant was affected (Idemaru, 2005). In addition to the amplitude of the vowel, f0 of the following vowel has also been investigated. Most studies have reported higher values for the f0 of the vowel following a geminate than a singleton in initial position. In Pattani Malay, f0 is higher for post-initial-geminate vowels (Abramson, 1998). Likewise, the mean f0 is significantly higher for vowels following initial geminates than singletons in Kelantan Malay (Hamzah, 2013).

Non-closure-duration correlates that contribute to the geminate-singleton distinction are particularly found in gemination at the periphery of the word, especially word-initially, where geminates are characterised by tenseness (Ridouane, 2007). This term has been used synonymously with other terms, such as articulatory strength (Al-Tamimi & Khattab, 2011) or fortition (Hamzah, 2013), to describe the articulatory force characterising geminates relative to singletons, in contrast with the rather commonly held concept that gemination is a distinction based on lengthening of the closure. In this paper, we use the phrase “forceful articulation” to describe this aspect of articulation. Thus, the additional non-closure-duration correlates can be indicative of a more forceful articulation for geminates relative to singletons, which led some

researchers to question whether the phonological and articulatory nature of gemination is a forceful articulation distinction, rather than a length/elongation distinction (Ridouane, 2007). This question is explored in this study as well as the possibility that gemination is both the result of a forceful articulation and elongation and that the balance between these two effects is language specific.

Neutralisation of initial gemination, i.e., degemination processes, can also occur, reflected as the loss of the contrast word-initially in a given language. Neutralisation could arise because the closure duration is inaudible for voiceless stops in absolute initial position, due to the absence of a preceding vowel and the corresponding lack of vocal fold vibration. A potential counter to domain-initial prosodic strengthening is that That is, there is no acoustic context to provide a perceptual cue for the onset of the closure. Poor acoustic closure duration cues to gemination in word-initial positions could, for example, lead to diachronic neutralisation of the contrast (Blevins, 2004). For example, the geminate/singleton contrast is not maintained acoustically in Swiss German, where the geminate/singleton contrast is neutralised in word-initial position (Kraehenmann, 2001), or in Russian, where geminates are freely degeminated word-initially (Kasatkin & Choj, 1999). Importantly, poor acoustic cues to initiation gemination make acoustic measurement of closure duration for these segments impossible. Thus, it is necessary to apply articulatory (especially kinematic) techniques in order to investigate the articulatory characteristics of the geminate/singleton contrast.

The lack of closure onset information in the acoustic signal for voiceless stops in utterance-initial position makes acoustic measurement of closure duration for these segments impossible. Voiced stops are also problematic in some cases, since the closure of initial geminated voiced stops can be partially or fully devoiced (Taylor, 1985). Thus, it is necessary to instead apply articulatory (especially kinematic) techniques to quantify closure duration

directly in stop consonants. Articulatory investigations of gemination utterance-initially have employed electropalatography (EPG) (Ridouane, 2007) and oral airflow (Ridouane, 2003) to measure closure duration. The results from these studies suggest that geminate stops are produced with a longer articulator contact (i.e., extended over a longer period of time) compared to singletons. However, the most informative data would capture the temporal dynamics of tongue movements in formation of singleton versus geminate stop closure, whereas EPG and oral airflow provide only indirect estimates of those motions. More direct techniques, such as ultrasound, which we use here, have not been used previously to measure tongue movement over time for gemination contrasts word- and utterance-initially.

Ultrasound is an ideal technique for tracking midsagittal tongue shape by obtaining real-time video sequences of the upper surface of the tongue in a non-invasive manner. Compared to other kinematic imaging technologies such as real time magnetic resonance imaging (rtMRI) or electromagnetic articulography (EMA), ultrasound produces less distortion of the speech signal because it is mechanically quiet and the transducer is placed completely outside of the oral tract with the participant comfortably seated (rtMRI: machine noise, participant in supine position in cramped quarters; EMA: sensors affixed to the tongue and other oral articulators, which can physically interfere with the production of natural speech). Ultrasound is also portable and relatively inexpensive, which further justifies the use of this technique for articulatory investigations. However, there is a dearth of studies using ultrasound to investigate the articulatory characteristics of segments. Despite the sizable corpus of research on gemination, articulatory investigations of the contrast are few and far between, and studies employing ultrasound to examine the temporal analysis of tongue movements are quite rare. The small handful of previous ultrasound gemination studies focused on tongue shape during geminate vs. singleton production, rather than on the temporal characteristics of the consonant closure. In one study on the articulation and acoustics of Kannada (Kochetov, Sreedevi, Kasim

& Manjula, 2012), the geminate affricate /ts:/ had a more anterior and raised tongue body than geminate dental /t̪:/, retroflex /ʈ:/ and velar /k:/, and had a laminal constriction whereas the dental had an apical constriction. It should be noted that there were no comparisons to their singleton counterparts in that study. A more recent ultrasound investigation of coronal stops in Eastern Oromo found that singleton ejectives were produced with a more raised tongue body than their geminate counterparts (Percival, Kochetov and Kang, 2018). Although studying the differences in tongue shape during the production of geminates and singletons could be useful in the distinction between vowels, the present study will focus on temporal analysis, which should be more informative given that the contrast is based primarily on durational differences.

The main reason why ultrasound investigations of gemination have disregarded the temporal aspect is that the most common approach to analysing ultrasound data is semi-automated tracing of the midsagittal contour of the tongue surface in individual images. In order to minimise the amount of work and time necessary to process an entire ultrasound video recording, tongue contours are usually traced at single time points, rather than for every image frame in the recording (see Falahati, 2013, for an example of the latter). This static approach to ultrasound analysis can be useful in investigating particular characteristics of speech articulation at a particular point in time, but it ignores the important temporal aspect of speech production.

Some studies have addressed this limitation by extracting dynamic information from ultrasound videos by analysing pixel differences (Palo, Schaeffler & Scobbie, 2014), using optical flow analysis (Moisik et al., 2014), or applying principle component analysis (PCA; Hueber et al., 2007). For the present study, we adopted a derivation of the PCA approach, called Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; Carignan, 2014). This method is ideally suited to addressing our research questions about

closure duration because it provides temporal information about speech kinematics, with substantially less manual processing than tongue contour tracing.

Word-initial gemination in Moroccan and Lebanese results from a morphophonological process whereby the definite article /l/ (underlyingly /al/ in Modern Standard Arabic) totally assimilates to a word-initial coronal stop, making the latter a geminate. In this case, initial gemination arises across a morpheme boundary of a bi-morphemic word in which assimilation has deleted the first morpheme's onset and nucleus, leaving only the coda consonant, which assimilates totally to the coronal onset of the stem morpheme (e.g., alduda⇒lduda⇒dduda). This makes gemination in Moroccan and Lebanese similar in some respect to the more typologically common and most often studied word-medial gemination, where gemination also arises across a boundary, but only across a word-internal syllable boundary, where it does not result from a morphophonological process.

It is important to note that only coronal consonants can occur as geminates in word-initial position. Thus, we focus only on the coronal stops /t/ and/d/ in the current study. Most of the previous studies on word-initial gemination have been conducted on languages such as Kelantan Malay (Hamzah, 2013), where it is a purely lexical distinction, i.e., within a monomorphemic word. The Kelantan Malay cases are not prosodically comparable to medial gemination because the initial geminate is the onset of a syllable, i.e., it does not arise across a syllable boundary, and it is not comparable to Moroccan or Lebanese initial gemination either as it does not cross morpheme boundaries. This is another reason why it is fair to assume that phonological differences between Moroccan and Lebanese will affect the realisation of gemination in these two dialects on the surface phonetic level.

In this study, we investigate word-initial gemination in Moroccan and Lebanese using combined articulatory and acoustic measurements, allowing us to test the principle of domain-

initial strengthening (prosodic factors) as described earlier, as well as exploring the contribution of phonological effects (syllable structure) and of dialectal differences, all of which could potentially impact the phonetic realisation of the geminate/singleton contrast.

6.3. Method

6.3.1. Participants

Five Moroccan ($Mean_{age}$: 32.8 years, range = 27-40) and five Lebanese ($Mean_{age}$: 31.4 years, range = 24-37) male speakers were recruited to participate in the experiment. Only male participants were recruited to avoid effects of gender differences on the production of gemination, as females have been reported to produce gemination in word-initial position with smaller average durations than males in Kelantan Malay (Hamzah, 2011). All the participants were born and received education through secondary school level in their respective countries. They had lived in Australia for 1-4 years ($Mean_{Moroccan}$: 3.2 years, $Mean_{Lebanese}$: 3.4 years) at the time they participated in the experiment. They reported no hearing, vision or speech impairments.

6.3.2. Materials

A list of four target items was prepared, consisting of two minimal pairs contrasting /t/ with /tt/ and /d/ with /dd/ word-initially (see Table 6.1). All the words were disyllabic real words, displaying the structure XaCaC with X being either a singleton or a geminate. There are no minimal quadrads of words in common to both Moroccan and Lebanese, therefore we used a near-minimal quadrad of words found in both dialects, which comprise two minimal-pair initial gemination contrasts for the voiced and voiceless coronal stops of the two languages.

Table 6.1: List of word-initial stimuli, their English glosses and how they were presented to the speakers (shown between brackets)

	/d/	**/t/**
Singleton	/damar/ 'destruction' [damar]	/taman/ 'price' [taman]
Geminate	/d:amar/ 'the destruction' [ddamar]	/t:aman/ 'the price' [ttaman]

Participants were asked to repeat each token 10 times in isolation (utterance-initial position) and in a carrier sentence (utterance-medial position). The carrier sentences were [galt rahaf ____] for Moroccan and its equivalent in Lebanese [alit rahaf ____] ('Rahaf said ____'). The targets were presented in random order through a PowerPoint presentation. Participants were instructed to read the stimuli without pausing after the carrier sentence. The stimuli and the carrier sentences were presented to the speakers orthographically in Roman characters rather than Arabic script. This was done because initial gemination is not permissible in Modern Standard Arabic and because it is common practice for speakers of Arabic dialects, which are not codified, to use Roman characters when, e.g., chatting on the Internet. All the participants confirmed that they used the Roman characters either solely or interchangeably with Arabic characters to write in dialect. After a modest number of practice trials (~10), all the participants were able to read the stimuli as presented in the Roman characters without any problems.

6.3.3. Procedure

The recordings were made in the Analysis of Human Articulatory Actions (AHAA) lab at the MARCS Institute, Western Sydney University. The experimental setup, presented in Figure 6.1, was designed to allow the experimenter (the first author) to run a recording session without need of an assistant.

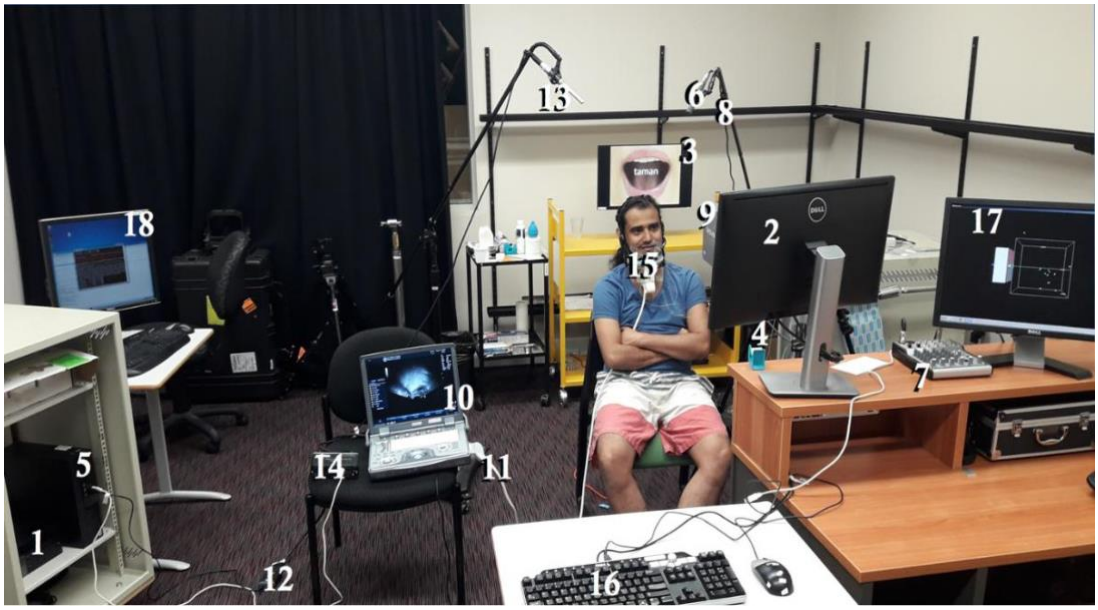

Figure 6.1: Experimental setup for ultrasound and audio recording. The numbers superimposed in white refer to the individual pieces of equipment that are referred to in the text.

Acoustic, ultrasound and electromagnetic articulography (EMA) data were co-collected and co-registered for all speakers. The EMA data were collected for additional work in separate reports on lip and jaw actions and for head-motion correction of EMA analyses. In this section, we refer to the equipment used by the respective number displayed in Figure 6.1; these numbers appear in parentheses. Three computers were used in total. One computer (1) presented the stimuli through two displays, one facing the participant (2) and one facing the experimenter (3). The second computer (4) was used to control the EMA system and record the acoustic data. The third computer (5) was used to record the ultrasound video. High quality acoustic data were recorded using a Behringer ECM 8000 omnidirectional condenser microphone (6) connected to the EG2-PCx preamplifier (7). The microphone was attached to a stand (8) and placed close to the participant and away from the EMA electromagnetic field. The audio recording was synchronised with the EMA data, which were collected using the NDI Wave system (9). No EMA sensors were glued on the tongue, in order to allow unimpeded speech articulation. Ultrasound images were generated using the GE LOGIQ e ultrasound portable

system (10) with a GE 8C-RS probe (11), operating at 8 MHz and at a depth of 8-10cm. Ultrasound video was extracted in real time from the VGA output of the GE LOGIQ e using an Epiphan VGA2USB Pro video grabber (12), which was connected to the third computer (5) via USB. FFmpeg software (FFmpeg Development Team, 2016) installed on the same computer was used to record a continuous AVI file with ultrasound video at 30 fps and synchronous acoustic data, which were collected through USB using a Sennheiser MKH 416 microphone (13) connected to a Sound Devices LLC USB Pre 2 (14). A non-metallic head-mounted ultrasound probe holder (15) was used (Derrick, Fiasson & Best, 2015; see also: Derrick, Carignan, Chen, Muayiwath & Best, 2018). The probe holder not only stabilises the ultrasound probe relative to the head within acceptable parameters for rotational and translational slippage, based on HOCUS measurements (Haskins Optically-Corrected UltraSound) (Whalen et al., 2005), but it also does not interfere with the EMA sensors because of its non-metallic components. The participants were seated in a chair with the NDI Wave EMA field generator placed to the left side of their head. The stimuli were presented through a Power Point slide show (Microsoft Office 2016) and advanced manually by the experimenter using a keyboard (16). During the experimental session, the experimenter had access to two additional displays, one for monitoring the EMA sensor placement (17), and the other for checking the occurrence of dropped ultrasound frames in the FFMPEG video recording (18).

Participants were asked to produce a “Ta, Ta, Ta” sequence at the beginning and end of recording. This was beneficial for data synchronisation because clear landmarks in the sequence could easily be identified in both the articulatory and acoustic signals; synchronisation was carried out by trimming the ultrasound and acoustic data between these landmarks.

6.4. Articulatory Analysis

For the purposes of this study, we innovated the use of ultrasound to index the temporal dynamics of initial gemination. Ultrasound tracks midsagittal tongue shape by obtaining real-time video sequences of the upper surface of the tongue. It offers advantages over other kinematic imaging technologies, such as real time magnetic resonance imaging (rtMRI), in terms of portability, safety, non-invasiveness and relatively low cost. More importantly for our experimental goals, however, ultrasound registration does not affect the acoustic output because the ultrasound probe is positioned entirely outside of the oral tract: no sensors are attached to the tongue (as with EMA) and no artificial palate is inserted into the mouth (as with EPG).

There is, however, a dearth of studies employing ultrasound to examine the temporal analysis of tongue movements, including for measurement of closure duration in gemination contrasts. A key reason behind this is that the most common approach of analysing ultrasound data relies on semi-automatic processing by tracing tongue contours, which often requires manual intervention. In order to minimise the amount of work and time necessary to process an entire ultrasound video recording, tracing tongue contours is usually done at single time points, rather than for every image frame in the recording. This method can be used to infer information about tongue position and configuration at that point in time, but it ignores the temporal aspect of speech production. Therefore, this technique would be ineffective for studying contrasts that are based primarily on duration differences (gemination). More recent techniques have allowed for time-varying measurements of speech articulation created directly from ultrasound video (Hueber et al., 2007; Moisik et al., 2014). We adopted one such method, Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; Carignan, 2014). This method is ideally suited to addressing our research questions about

closure duration because it provides temporal information about speech kinematics, with minimal manual processing compared to other methods, such as contour tracing.

TRACTUS analysis identifies a set of principal component (PC) coefficient matrices of the image (pixel) data for the video frames of the entire ultrasound recording that account for the greatest amount of variance in the ultrasound data. A principal components analysis (PCA) model was constructed for each speaker, and the PC scores of the ultrasound video frames associated with a closure are used as orthogonal predictors in subsequent Linear Discriminant Analysis models (LDA); due to the orthogonal nature of the PCs, the scores can be used as predictor variables without the risk of inflating standard error due to multicollinearity. The LDA models were trained on frames associated with tongue closure posture at 20 ms before the release. This ensures that the tongue shape is in the speaker-specific configuration associated with a full closure. By doing this, the score from the predictions based on the LDA model more closely represents probability of the tongue shape of just a closure.[11] The speaker-wise LDA models were then used to predict scores for all ultrasound frames in the speaker's entire recording, resulting in a temporal vector based on the probability that a particular image frame corresponds to a stop closure as produced by the given speaker. The higher the LDA score, the more likely that an ultrasound frame at the associated time point corresponded to a closure. This approach was applied across all frames within the designated analysis window, providing temporal information about the closure even in the absence of acoustic information (e.g., during a voiceless closure). In this way, TRACTUS yields a

[11] In previous analysis, LDA was trained on the PC scores for every frame within the boundary of a closure for every token. The issue with this method is that because the boundaries were wide, the LDA was trained on unnecessarily excessive variation in tongue shape. This generated inaccurate articulatory signal for the closure.

temporal articulatory analysis of ultrasound images automatically without the need for laborious image tracing of each video frame.

We showed in previous work on word-medial gemination in Moroccan that the articulatory duration differences between a singleton and a geminate are clearly demonstrated through the analysis generated by TRACTUS and the LDA model (Frej, Carignan & Best, 2017). However, in pilot studies on word-initial gemination, we discovered that before producing the target words in isolation, participants had the tendency to rest their tongue against or close to the hard palate. This made it difficult to capture the differences between geminates and singletons word-initially because the LDA analysis interpreted the tongue resting position as a stop closure, resulting in high-magnitude LDA scores at the beginning of the gesture (see Figure 6.2). Thus, the onset of the closure gesture could not be identified, because the increase in air pressure during the closure began without any change in midsagittal lingual kinematics. The solution we designed to circumvent this limitation was to elicit a silent /a/-like oral gesture before production of the words in isolation. This was achieved by instructing the participants to maintain a silent /a/, with jaw opening and lowering of the tongue prior to producing the stimuli and displaying a close-up picture of the mouth of a person producing the vowel /a/ in the background of the power point slides presenting the stimuli as a reminder (see Figure 6.3). This method ensured that the tongue would begin in a low position within the mouth before the closure gesture was initiated, thus capturing the entire lingual gesture for each token. Since /a/ is silent and not articulated by the participants, it does not make initial gemination comparable to intervocalic/medial gemination.

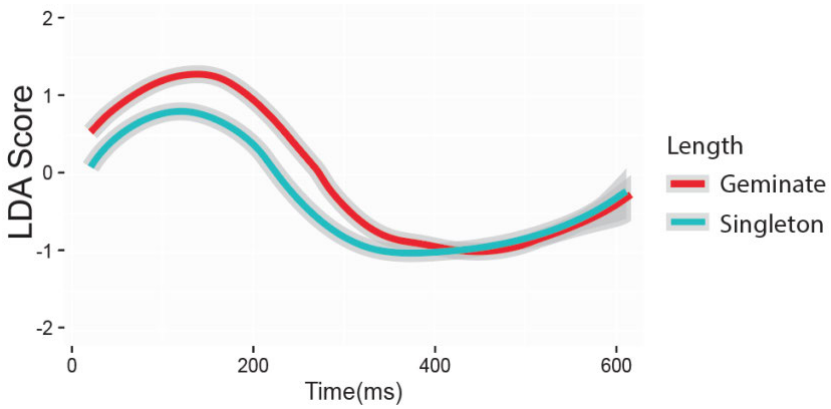

Figure 6.2: LDA scores for geminates and singletons in utterance-initial position from the pilot study (Frej, Carignan & Best, 2017), without the /a/-vowel-like silent mouth opening.

Figure 6.3: The background of the slides presenting the stimuli (left) and an example stimulus superimposed on the background (right).

In order to process ultrasound images using TRACTUS, two main Matlab functions need to be installed from http://christophercarignan.github.io/TRACTUS: "TRACTUS_prep.m" and "TRACTUS.m" along with additional functions. After having installed all the necessary functions, data processing goes through the following stages:

"TRACTUS_prep.m" has to run prior to running "TRACTUS.m". TRACTUS allows you to process ultrasound video as well images. For our data, ultrasound video was converted automatically into single images by the TRACTUS_prep function. The ultrasound video was captured at 30 frames per second. This generates a total of 1800 frames for every one minute of ultrasound video recording. After uploading all ultrasound images or video, a prompt

appears in which a threshold for automatic detection of the border of the ultrasound fan must be specified. The default given threshold is 5%. Thus means 5% of the images will analysed and the standard deviations of the pixel intensities is calculated accordingly. An example of the fan detection results ins displayed in Figure 6.4.

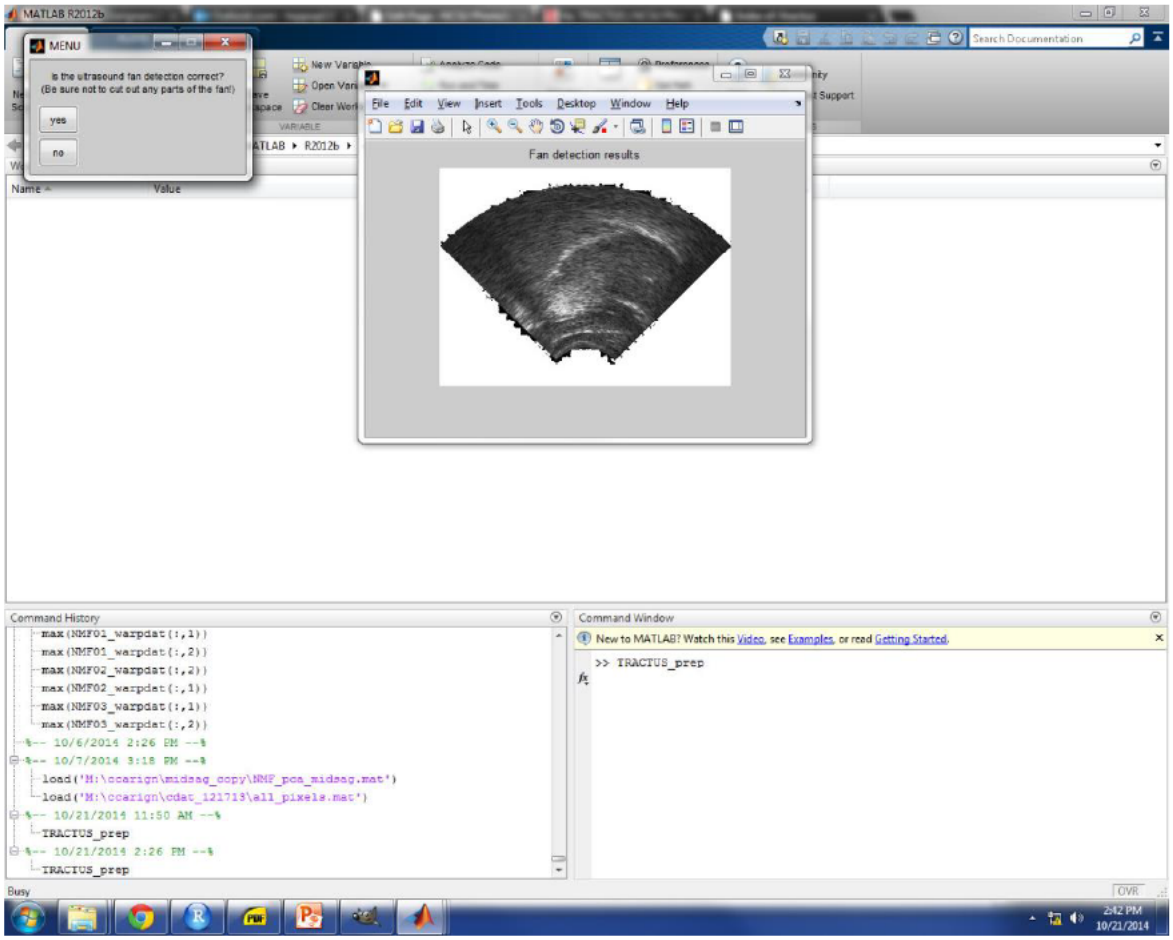

Figure 6.4: TRACTUS fan detection result.

After fan detection, a prompt appears in which the region of interest should be selected (ROI). This concerns the area that would be included in the automatic analysis. The larger the area the longer it takes for the ultrasound data to be analysed. A similar prompt to the fan detection prompt appears to help the researcher visualise the area of the tongue that should be analysed. Once the desired ROI is selected manually as shown in Figure 6.5 the preparation is

finished.

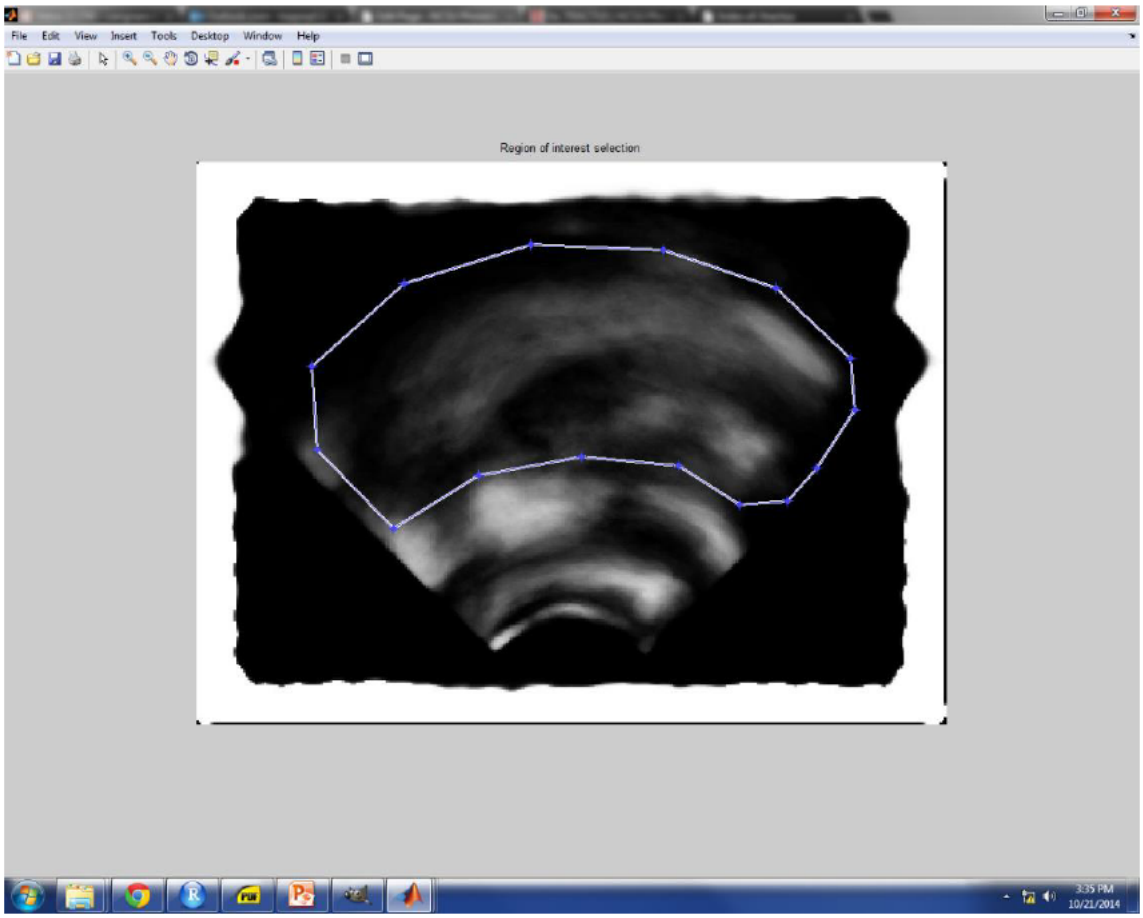

Figure 6.5: TRACTUS region of interest in ultrasound image

Once the preparation is finished, "TRACTUS.m" function is ready to be run. After running the function, a prompt appears as in Figure 6.6 where four input should be entered: 1) the number of PCs to be retained, 2) the scale of image resolution to be used in the analysis, 3) (optional) the angle of occlusal plane rotation and 4) (optional) the textfile containing information about the amount of data that the research wishes to analyse. After entering this information, TRACTUS will start analysing the data automatically.

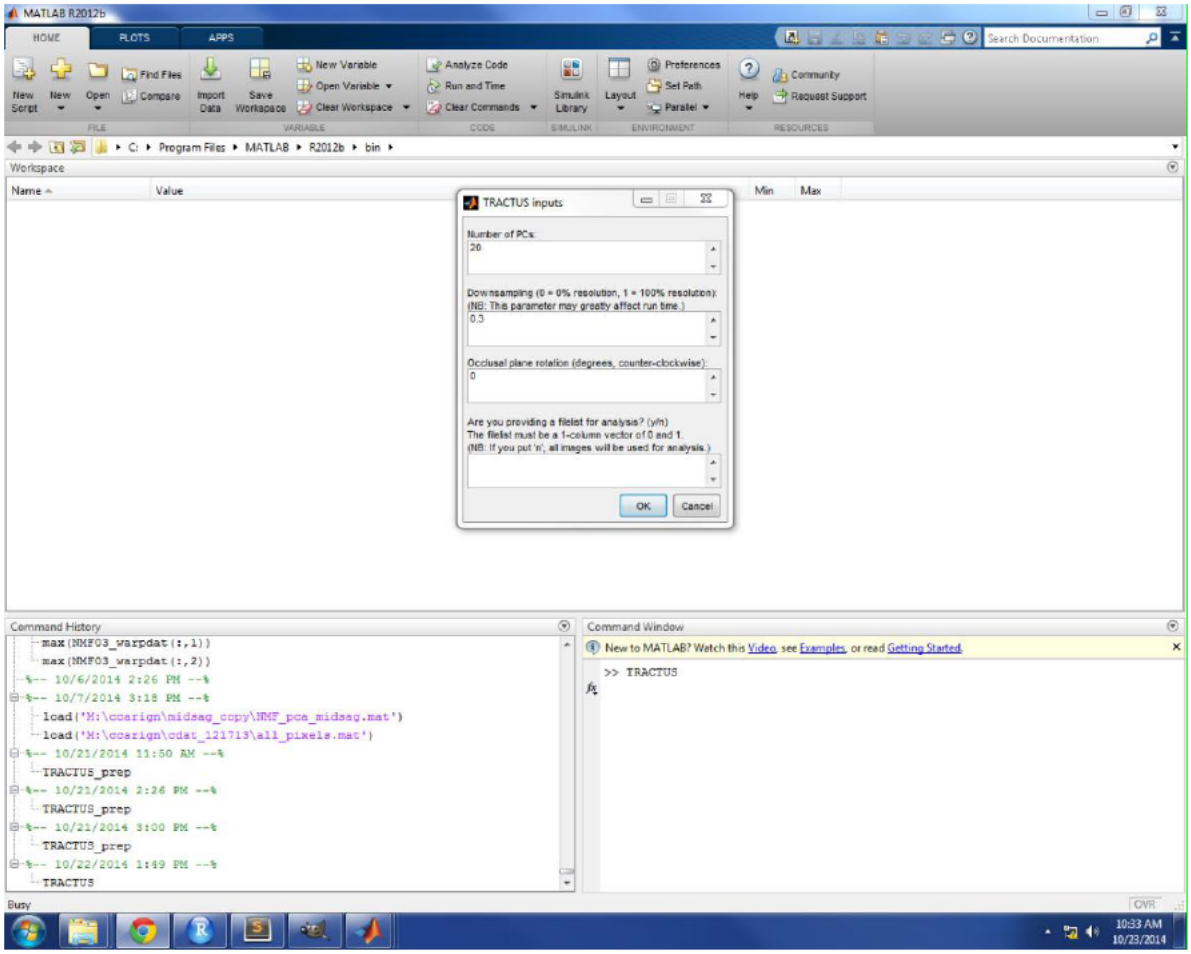

Figure 6.6: TRACTUS input

The TRACTUS analysis generates a number of files, which could be submitted for further analysis in order to derive articulatory signals. More detailed information about TRACTUS data analysis can be viewed through this link: https://phon.wordpress.ncsu.edu/lab-manual/ultrasound-and-video/working-with-data/other-methods-of-ultrasound-analysis/

6.5. Acoustic analysis

Using a custom Praat script, the following acoustic measurements were made on each token for each speaker in each utterance position:

- The closure duration of the geminate/singleton onset consonant.
- The post-consonantal vowel duration.

- The release burst RMS amplitude.
- The post-consonantal vowel onset amplitude.
- The post-consonantal vowel onset f0.

We used a separate script based on Hirst's (2011) two-pass method to accurately estimate fundamental frequency at the onset of the vowel, because the default settings for computing fundamental frequency are usually deemed unsatisfactory (Hirst, 2011). The script uses an autocorrelation method to automatically estimate f0 values. The first pass is generated based on its default parameters (5-ms time step, 40-ms Kaiser2 window, floor=75 Hz and ceiling = 600 Hz) and auto-correlation. In the second pass, the values of the *quartile*1(25%) and *quartile*3(75%) are taken because they are considered more robust estimates than the extreme values of the distribution and multiplied by 0.75 and 1.5 respectively[12].

The segmentation and annotation of the target words and phones was done manually using Praat (Boersma & Weenink, 2007), based on careful visual inspection of the broadband spectrographic and acoustic waveform and following standard segmentation procedures. The closure duration of the absolute word-initial stops in isolated words (i.e., utterance-initial) could not be measured from the acoustic signal, since the onset of the closure could not be determined especially for the voiceless stop /t/. For segments in the carrier phrase (i.e., utterance-medial position), the onset of the closure was easily segmented, as inspection revealed that the participants pronounced the target words without any pause after the carrier phrase. The closure duration of the geminate/singleton contrast was then marked from the offset of friction of the /f/ of the carrier phrase to the release burst of the stop, which was identified by a transient spike of noisy energy in the spectrogram. The vowel /a/ was segmented

[12] The script and instructions on its use can be obtained from https://github.com/JalalAl-Tamimi/Praat-f0-Accurate-Estimation

from the onset of clear periodicity and formant structure until the nasal /m/ of the next syllable, which was identified by lower frequencies of F1 compared to formants of the preceding vowel.

The segmentation and the annotation explained above is displayed in Figure 6.7 for the word-initial singleton /d/ and geminate /dd/ in carrier phrase context.

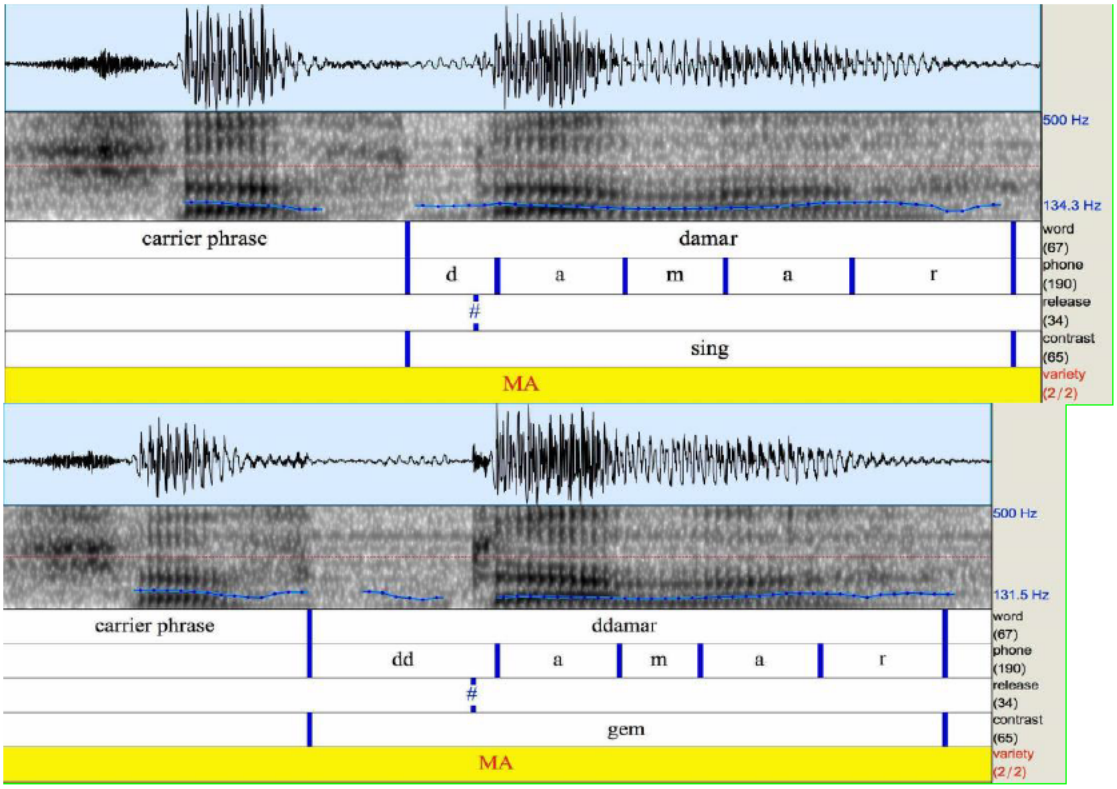

Figure 6.7: The waveforms and spectrograms of the words /damar/ and /d:amar

6.6. Results

6.6.1. Articulatory results

Figures 6.8 and 6.9 display the LDA scores of the articulatory signals for each group, as well as category means and 95% confidence intervals generated using t-based approximation ("loess" smoothing in the ggplot2 package in R). The higher the LDA score, the more likely that the tongue is producing a stop closure. For the most part, there are clear differences between the geminate articulatory signal, represented by the purple line and the singleton signal

represented by the green line. In particular, the geminate signal displays higher scores and the gesture is maintained longer than the one for the singleton signal. This is not surprising, given that articulatorily geminates have been reported to be produced with a longer articulatory contact (Ridouane, 2003).

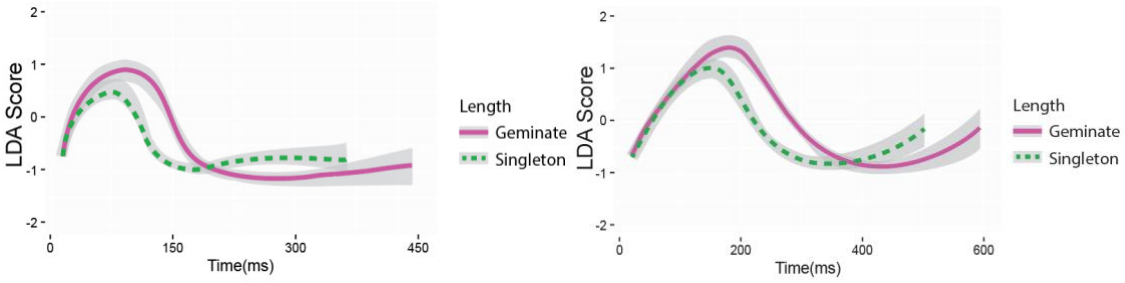

Figure 6.8: Lebanese ultrasound LDA class scores over time for geminates and singletons in utterance-initial position (left) and utterance-medial position. (right)

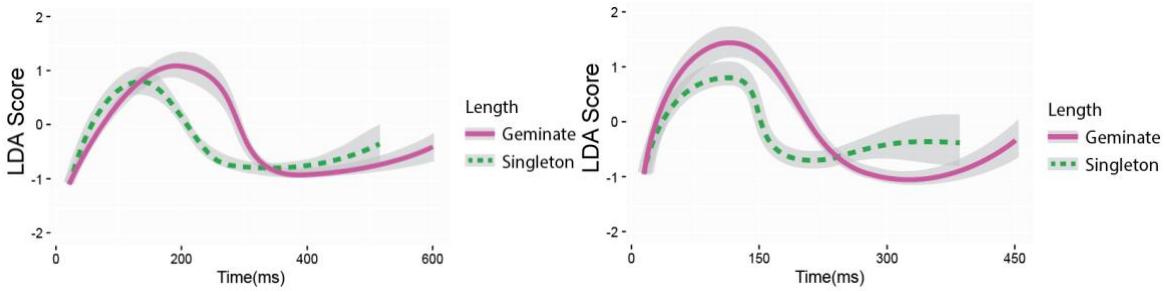

Figure 6.9: Moroccan ultrasound LDA class scores over time for geminates and singletons in utterance-initial position (left) and utterance-medial position. (right)

Given that Figures 6.8 and 6.9 are composite plots for each group, it is not possible to observe differences between different speakers. However, by looking into separate the LDA score plots for each speaker, we observed that utterance-initial figures for Lebanese speakers L2 and L5 stand out from the rest of the figures (Figures 6.10 and 6.11 respectively). Notably, their utterance-initial geminate and singleton LDA score lines overlap, which indicates a non-significant difference between a geminate and a singleton at this position. Interestingly, this is not the case for their word-initial geminates in utterance-medial position, where a difference

between the geminate and the singleton LDA score lines is clearly manifested. As for dialectal differences, it appears that the closure is sustained longer for Moroccan than Lebanese. The less pronounced difference between geminates and singleton LDA lines for the Lebanese group is believed to be partially caused by the results of the two Lebanese speakers L2 and L5. These articulatory characteristics will be further delineated in subsequent sections.

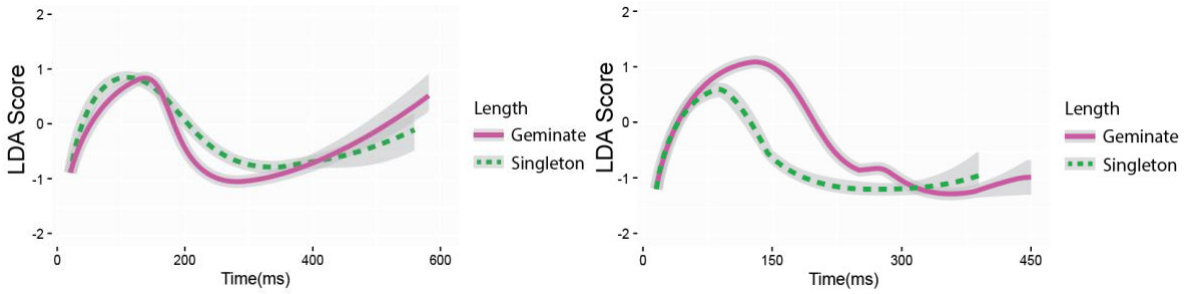

Figure 6.10: L2's LDA class scores over time for geminates and singletons in utterance-initial position (left) and utterance-medial position (right).

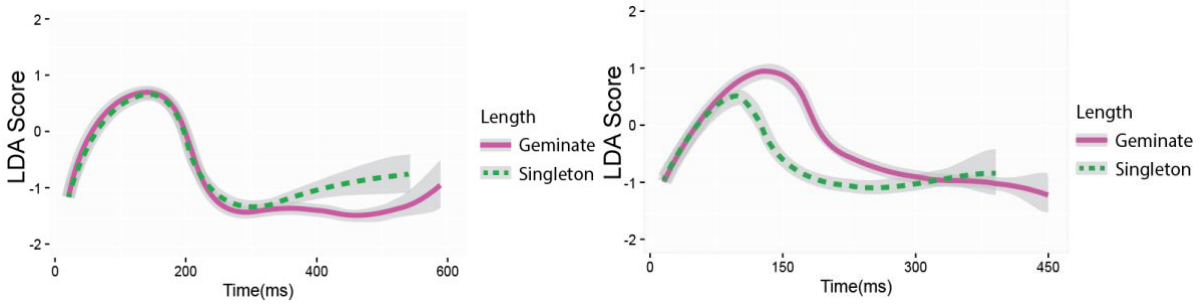

Figure 6.11: L5's LDA class scores over time for geminates and singletons in utterance-initial position (left) and utterance-medial position (right).

6.6.1.1. Calculation of articulatory closure duration

As can be seen from Figures 6.8-6.11, the articulatory signals generated from ultrasound video using TRACTUS can reliably detect changes in duration between a geminate and a singleton even in utterance-initial position for voiceless stops. In order to quantify these durational differences, we created a Matlab function to estimate duration values of the closure semi-automatically through determining gestural landmarks from the velocity signal. First, the

velocity signal was calculated from the articulatory signal. Next, the time points corresponding to the maximum velocity peak (movement into the closure) and minimum velocity peak (movement out of the closure) were identified for each token. Subsequently, the time points corresponding to samples which crossed a threshold of 20% of these peaks were calculated and used as the onset (20% of the maximum velocity) and offset (20% of the minimum velocity) of the articulatory gesture. The number image frames between these gestural landmarks was identified and stop closure duration was estimated by calculating the number of frames in the interval by the ultrasound frame rate (see Figure 6.12 for an example). In order to evaluate the validity of this closure calculation function from the articulatory data, we performed a standard correlation analysis on the results from utterance- medial data obtained from the acoustic analysis and using the articulatory closure calculation function. The results are shown in Figure 6.13.

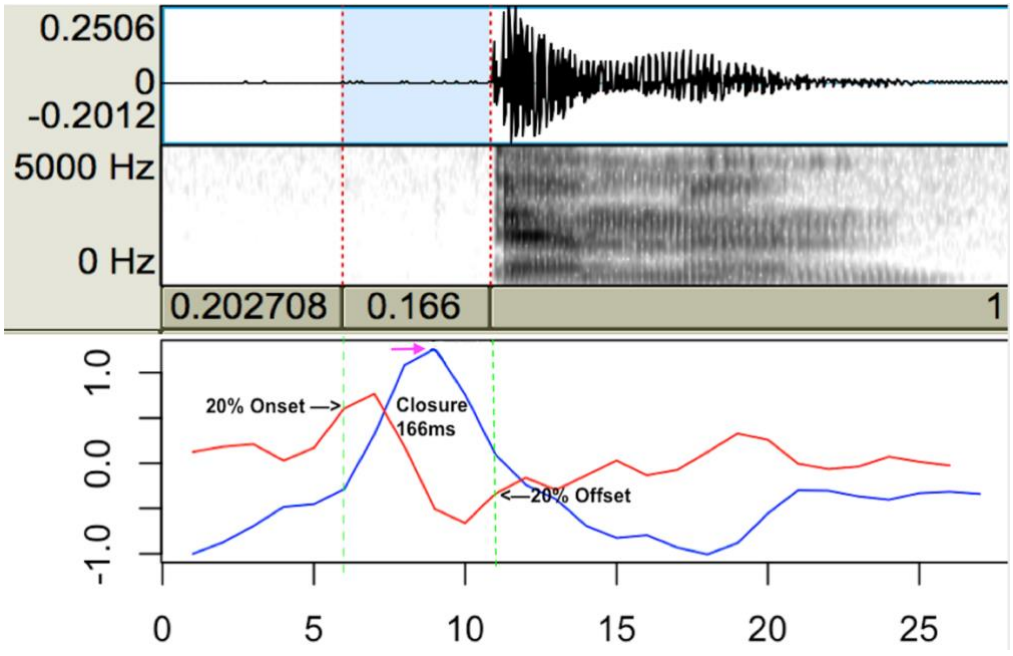

Figure 6.12: Estimation of closure duration. Articulatory signal (blue line), velocity signal (red line) and time-aligned waveform and spectrogram. The time point associated with the peak of the closure signal is automatically denoted by (pink arrow).

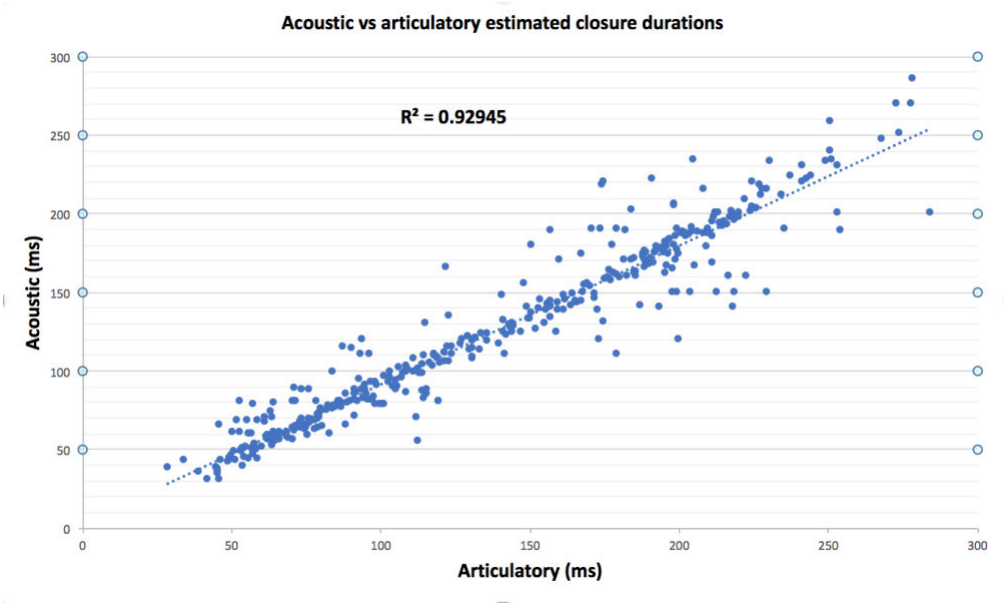

Figure 6.13: Scatter plot of acoustic closure duration vs. estimated articulatory closure duration.

There is a strong positive correlation, $r = 0.96$, between the articulatory duration and the acoustic duration of the closure in the utterance-medial position, thus validating the closure duration estimation technique we created. With this understanding, we will use articulatory closure duration as an accurate proxy for acoustic closure duration, even in cases where acoustic closure duration cannot be obtained, e.g., utterance-initial position.

Next, we set up a LME analysis in order to compare between the values generated by the acoustic duration and the articulatory estimated duration values. This was done to find out if estimated articulatory values can substitute missing acoustic values (utterance-initial voiceless closures). Our analysis showed that articulatory and acoustic closure durations do not have the same range of values despite the strong positive correlation between the estimations of duration generated by both methods. Acoustic closure duration is 22 ms longer than Articulatory closure duration for geminates and 14 ms for singletons. Using the estimates

obtained by the LME model, we corrected the values generated by the articulatory analysis before using them to replace the missing values in the acoustic analysis.

6.6.1.2. Articulatory closure duration

The articulatory closure durations of voiceless stops in utterance-initial position for both dialects are presented in Table 6.2. We used *t*-tests to determine whether there are significant differences between the means of geminates and singletons. The alpha level was adjusted for the 10 models using Bonferroni correction, in order to account for an increase in Type I error; accordingly, differences are considered to be significant at *p*-values < 0.005.

Table 6.2: Mean closure duration estimation values (ms), *p*-values and t-values for geminate and singleton voiceless stops in utterance-initial position for each speaker of both dialects.

Speakers	Geminate	Singleton	*t*-value	*p*-value
Moroccan				
M1	271	197	4.93	.005
M2	281	198	4.95	.005
M3	248	139	4.74	.005
M4	276	229	3.27	.005
M5	311	228	4.60	.005
Lebanese				
L1	261	165	5.12	.005
L2	181	174	1.13	.082 n.s
L3	260	159	4.96	.005
L4	269	171	4.02	.005
L5	171	168	0.92	.956 n.s

The closure duration results are in agreement with the results displayed by the LDA plots in showing that geminates are produced with longer articulatory closure duration even when this closure is not acoustically evident. Importantly, the statistical analysis shows that L2 and L5 failed to produce significant durational differences between a geminate and singleton, which is consistent with the LDA score plots for these two speakers.

These results suggest that this technique offers a computationally efficient and robust method for quantifying duration differences for contrasts whose temporal changes are difficult or impossible to discern from the acoustic signal alone.

6.6.2. Acoustic results

Each of the acoustic dependent variables was analysed using linear mixed effects (LME) models (Baayen, Davidson & Bates, 2008) to test the effect of the following fixed factors: group (Lebanese, Moroccan), utterance position (initial, medial), length (geminate, singleton) and voicing (voiced, voiceless). Multiple models were generated and evaluated using ANOVAs in order to come up with the best model fit. With the exception of the closure duration and the vowel duration analyses where random slopes for speakers improved the model fit, treating speaker as a random intercept resulted in better overall model fit for the other analyses. All of the dependent variables were centered and scaled for each speaker. To look into speaker specific values, a separate model was run by coding speaker as a fixed effect. Random slopes for position, length and item did not improve the model fit for any of the analyses. Tukey HSD Post-hoc tests were used to break down interactions. We generated *p*-values using the library lmerTest (Kuznetsova et al., 2017). Any *p*-value of .05 or below was deemed significant.

6.6.2.1. Acoustic closure duration

As expected, there was a main effect of consonant length, $F(1,46) = 15.6$, $p < .001$, indicating that geminates are produced with longer durations than singletons regardless of their voicing type and position in the utterance. There was also a significant effect of utterance, $F(1,48) = 32.76$, $p < .001$, indicating that consonants in utterance-initial position are produced with longer durations than those in utterance-medial position. Specifically, the results show

that initial geminates in the utterance-initial position are characterised by a longer closure duration than singletons 288 ms vs. 205 ms and 236 ms vs. 175 ms for Moroccan and Lebanese, respectively. In other words, geminates are 1.40 times longer than singletons in Moroccan and 1.35 times longer in Lebanese. Although these durations are longer than in utterance-medial position (i.e., 197 ms vs. 85 ms in Moroccan and 184 ms vs. 81 ms in Lebanese), the geminate-to-singleton ratio is larger utterance-medially than -initially, i.e., 2.31 for Moroccan and 2.27 Lebanese, as indicated by a significant interaction between length and position, $F(1,44) = 57.6$, $p < .001$. The Moroccan group's duration values ($M = 246.5$ ms) are significantly longer than those of the Lebanese ($M = 205.5$ ms), $F(1,44) = 15.8$, $p < .001$. The LME analysis also showed that the effect of position on length is significant, $F(1,44) = 32.7$, $p < .001$. These results reveal that the contrast between geminates and singletons is more enhanced in utterance-medial position than in utterance-initial position. Next, we decided to run a separate LME with speaker as a fixed factor in order to investigate speaker-specific values. Interestingly, there was a main effect of speaker, $F(9,61) = 37.6$, $p = .0129$, indicating that some speakers produced longer values than others. Importantly, there was a significant interaction between speaker and length $F(9,63) = 13.6$, $p = .0201$, and significant interaction between speaker length and position $F(9,62) = 19.6$, $p = .0111$. Tukey comparisons for each speaker and for each position showed that all speakers produced significant duration differences between geminates and singletons across both positions with the exception of Lebanese speakers L2 and L5 who failed to do so in the utterance-initial position only. This can also be seen in Figure 6.14.

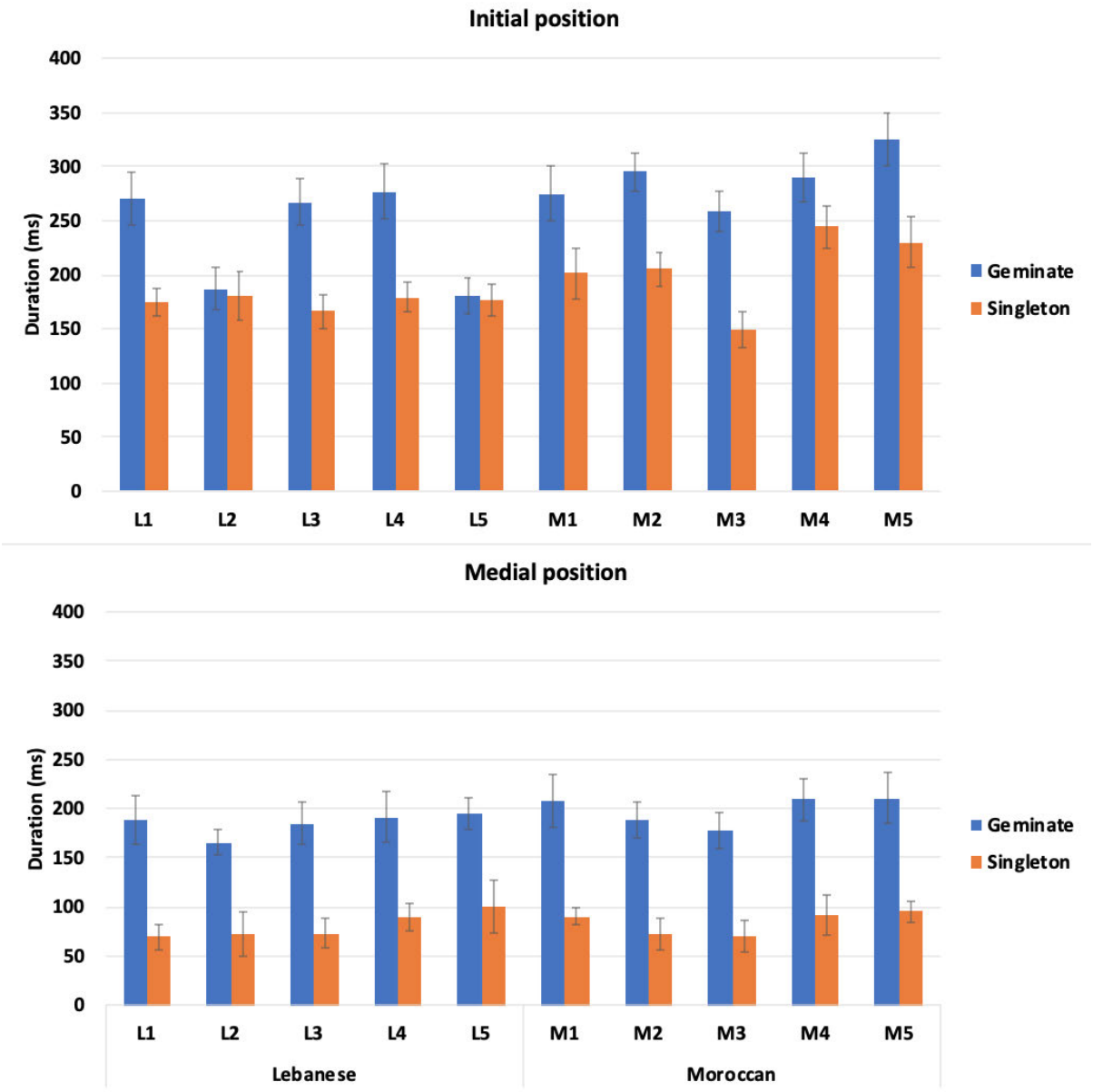

Figure 6.14: Mean closure duration values for each of the geminate and singleton targets displayed by group and by speaker (L1-5: Lebanese; M1-5: Moroccan) in utterance-initial (upper panel) and utterance-medial (lower panel) positions and standard error of the mean.

During the segmentation of the tokens, close inspection of the spectrograms revealed that the two Lebanese speakers who demonstrated a relative inability to contrast geminates with singletons in utterance-initial position in terms of closure duration (L2 and L5) produced an epenthetic vocalic element preceded by a glottal stop prior the stop closure of word-initial geminates; no epenthetic vocalic element was observed for any of the other speakers, and no

singleton closures were preceded by this epenthetic vocalic element for the two Lebanese speakers L2 and L5 either. This phenomenon is common in other Arabic dialects that do not allow word-initial consonant clusters, such as Egyptian Arabic, but it has not previously been reported in speakers of Lebanese (which does allow word-initial gemination), likely owing to the lack of research on this dialect. However, in a study on initial gemination in Maltese, a language related to Arabic, 96% of the target words were produced with an epenthetic [i-like] vocalic element alone or preceded by a glottal stop (Galea & Ussishkin, 2018). This high frequency explains why some studies argue against the existence of initial gemination utterance-initially in Maltese (Hoberman & Aronoff, 2003). This vocalic insertion makes initial gemination structurally similar to word-medial gemination. In our study, we investigated this phenomenon further by counting the occurrences of this vocalic insertion. L2 and L5 respectively produced 20% and 18% of the geminated tokens with a vocalic segment preceded by a glottal stop. The mean duration of the inserted vowels in these cases was 72ms. The closure duration of geminates following these cases of epenthetic vowels for these two speakers was on average 31 ms longer than their other geminate closures produced utterance-initially without an inserted vocalic element. This means that only geminate closures preceded by the vocalic element had noticeably longer durations than their singleton counterparts. This suggests that L2 and L5 were using this inserted vowel as a leverage to contrast geminates with singletons but were doing so inconsistently. The remainder of these two speakers' tokens did show differences in closure duration. Figure 6.15 presents the spectrogram and the waveform of two tokens produced by L2 and L5 with an inserted vocalic segment.

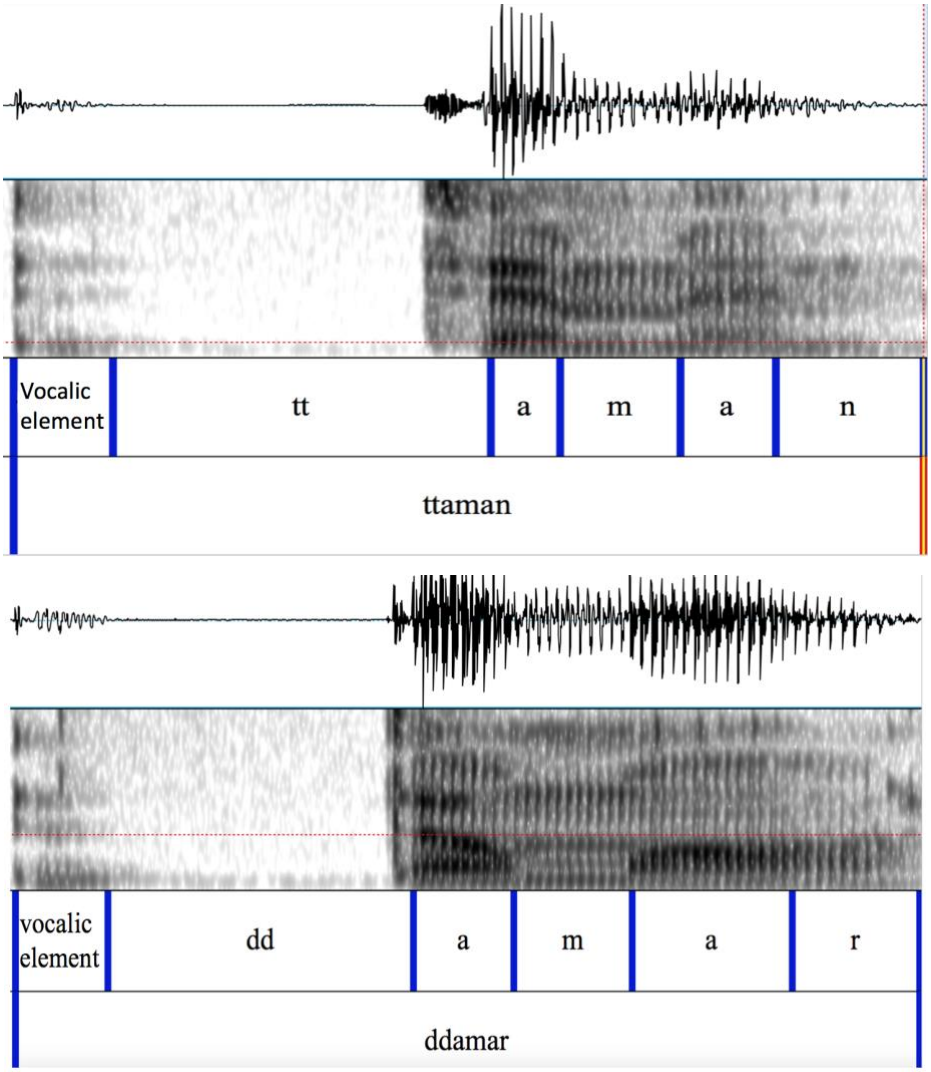

Figure 6.15: Spectrogram and waveform of tokens produced with a pre-geminate vocalic element. The upper panel is /t:aman/, produced by L2, and the lower panel is /d:amar/, produced by L5.

6.6.2.2. Vowel duration

The LME analysis showed that there was a significant effect of length of the vowel following the consonant, $F(1,46) = 43.7$, $p < .001$, both groups produced longer vowel durations after a geminate than a singleton in both utterance positions (see Figure 6.16). Moreover, the LME models revealed a significant effect of group, $F(1,44) = 57.6$, $p < .001$, indicating that Lebanese produced longer duration values for the post-consonantal vowel than Moroccan. There was also an interaction between group and position, $F(1,44) = 12.9$, p

= .0103. Tukey HSD post-hoc comparisons revealed that, in Lebanese, the vowel duration was significantly longer in initial position than in medial position ($p < .001$). However, position did not significantly affect the length of the vowel in Moroccan.

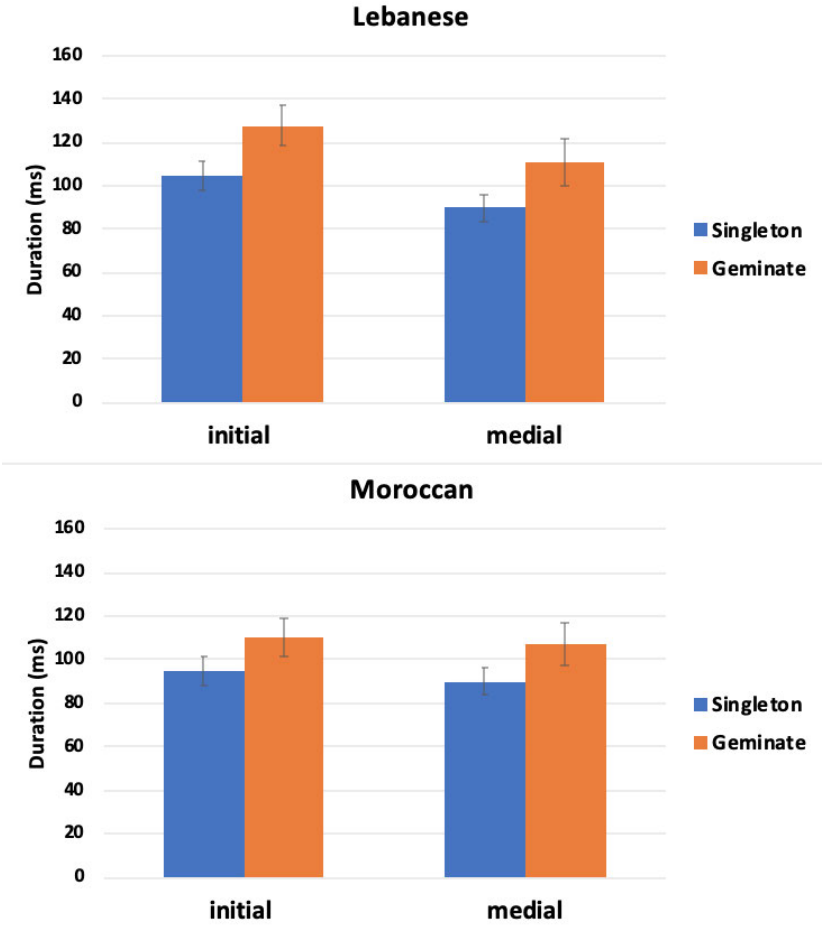

Figure 6.16: Mean vowel duration (ms) and standard error of the mean for each of the geminate and singleton tokens produced in Lebanese (upper panel) and Moroccan (lower panel) in utterance-initial and -medial positions.

6.6.2.3. Release burst amplitude

The LME analysis of the burst amplitude revealed that geminates were generally produced with higher burst amplitude than singletons, $F(1,46) = 10.9$ $p = .0411$ (see Figure 6.17). The LME models showed a significant effect of voicing, $F(1,42) = 23.2$ $p < .001$, such that voiced stops had a higher burst amplitude than voiceless stops. There was also a significant effect of group, $F(1,44) = 14.5$, $p = .001$), where Lebanese speakers produced higher values for

burst amplitude than Moroccan speakers. The interaction between length and position was significant, $F(1,44) = 42.6$, $p < .001$, After running multiple comparisons using post-hoc Tukey HSD, the amplitude of the burst was shown to be higher for geminates than singletons only in the initial position ($p < .001$). In the medial position, amplitude did not vary as a function of the singleton/geminate contrast. This can be seen in Figure 6.13.

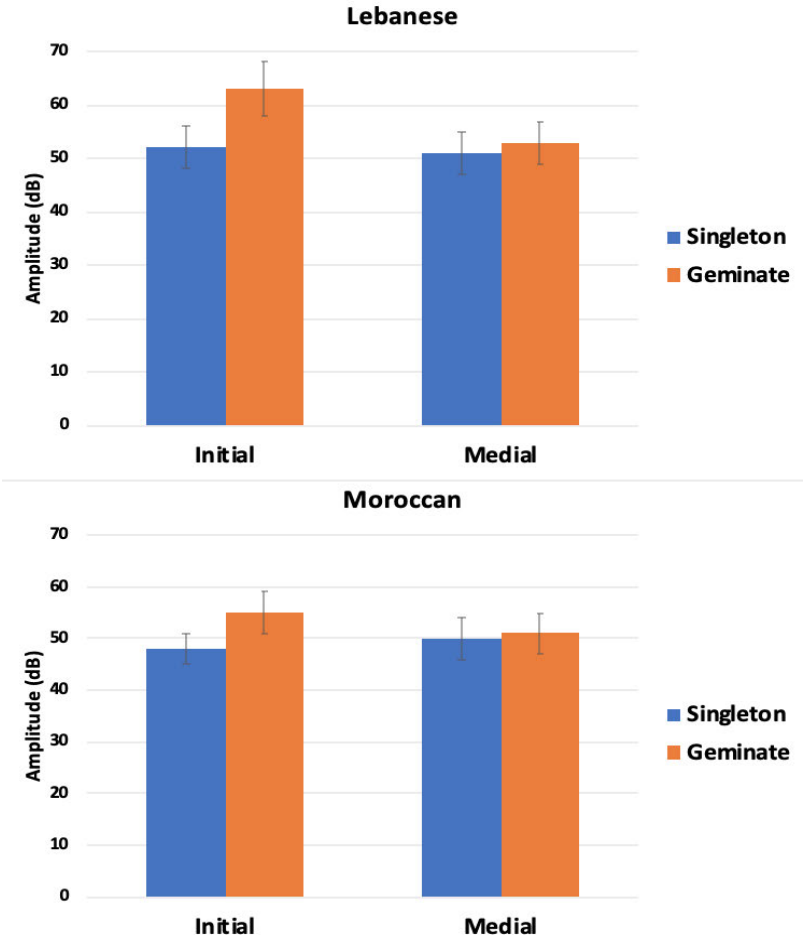

Figure 6.17: Mean amplitude of the stop burst and standard error of the mean for each of the geminate and singleton tokens produced in Lebanese in utterance-initial and -medial position.

6.6.2.4. Vowel onset amplitude

As was mentioned in the introduction, the effect of the production of the geminate/singleton contrast can be witnessed on the flanking vowels. Previous studies have shown that the vowel amplitude on either side of the consonant can contribute to the distinction

between geminates and singletons in word-medial position. In the context of word-initial gemination, which is more relevant to this study, Hamzah (2013) measured and averaged the vowel amplitude over 10% of the vowel following the consonant in Kelantan Malay. His overall results showed that post-geminate vowels are produced with 3dB higher amplitude than their singleton counterparts. The results of the mean vowel onset amplitude are presented in Figure 6.18.

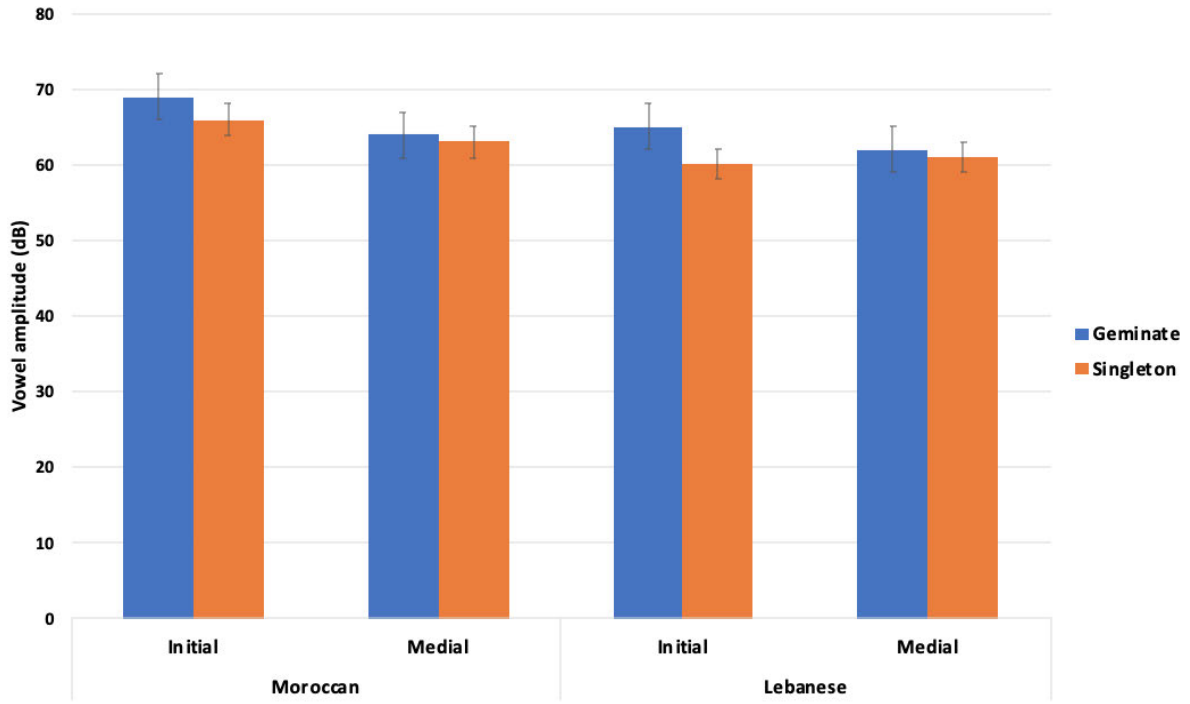

Figure 6.18: Mean vowel amplitude (dB) and standard error of the mean for geminate and singleton in both utterance positions and for both dialects

The LME model revealed a significant main effect of consonant length on vowel amplitude, $F(1,46) = 11.9$, $p = .0433$, i.e., for the most part, the amplitude of vowel onset was conditioned by the length of the preceding consonant. In particular, higher amplitudes were registered for post-geminate than post-singleton vowel onsets. There was a significant effect of position, $F(1,48) = 36.2$, $p < .001$, indicating that the amplitude at the onset of vowels was

higher in the utterance-initial than utterance-medial position. While there was no significant main effect of consonant voicing on vowel amplitude, the interaction between voicing and group was significant, $F(1,42) = 42.6$, $p < .001$. Tukey HSD post-hoc comparisons revealed that the amplitude is higher for vowels following voiceless stops in Lebanese, but not in Moroccan. Moreover, the group factor yielded a significant main effect, $F(1,44) = 79.7$, $p < .001$, i.e, Moroccans produced higher vowel onset amplitude values than Lebanese, as shown in Figure 6.14. Importantly, the interaction between position and length was significant, $F(1,44) = 34.6$, $p < .001$. After running multiple comparisons, it was revealed that the difference between vowel amplitude for geminate vs. singleton is greater for utterance-initial than utterance-medial position. Also, the difference in vowel amplitude following a geminate vs. a singleton in the utterance-medial position was not significant. Although Moroccan registered higher values than Lebanese, the mean contrast between a geminate and a singleton is larger for Lebanese (5 dB) in utterance-initial position. In Moroccan, the mean difference is barely noticeable (3 dB), as suggested by just-noticeable-difference (JND) for the vowel amplitude (3 dB) (Toole & Olive, 1998).

6.6.2.5. Vowel Onset f0

As reported in the literature, in certain languages f0 of the vowel following a geminate is enhanced compared to the vowel following a singleton. The f0 of the following vowel may, therefore, be strong non-durational cue to the geminate/singleton contrast. The results of the mean f0 values for our data based on the LME statistical analysis are reported in Figure 6.19.

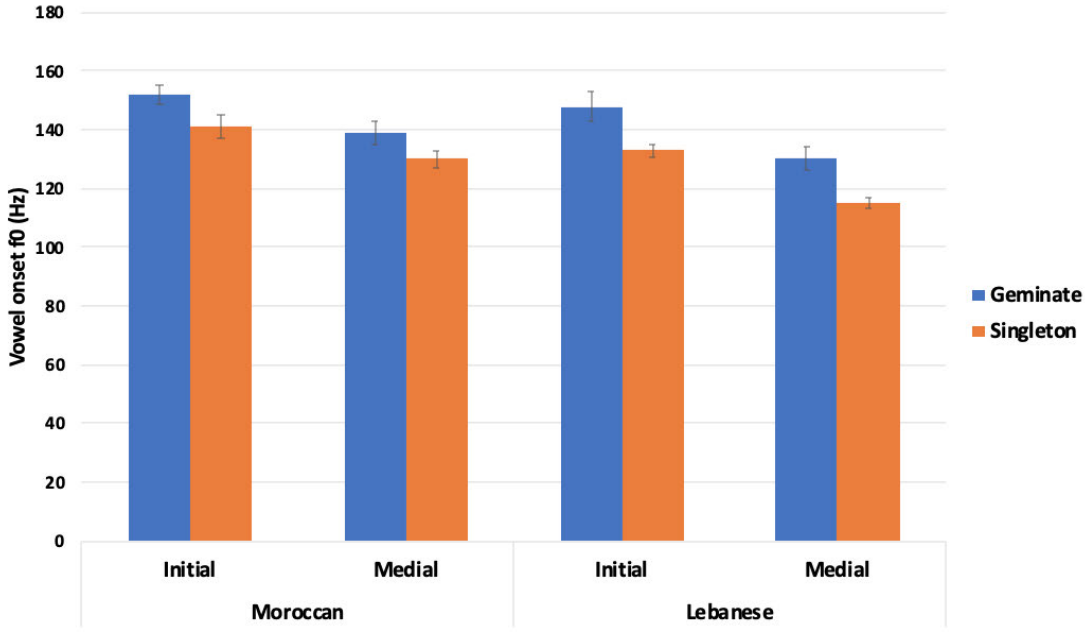

Figure 6.19: mean vowel onset f0 (Hz) and for geminate and singleton in both utterance positions and for both dialects. Error bars display standard error of the mean.

There was a significant main effect of length, $F(1,46) = 26.9$, $p < .001$), revealing that f0 values are significantly higher at the vowel onset following a geminate than following a singleton. There was also a significant effect of group, $F(1,44) = 34.7$, $p < .001$, i.e., Moroccan produced larger values than Lebanese. Moreover, a significant interaction between position and length was significant, $F(1,44) = 33.7$, $p < .001$, i.e., the f0 contrast between post-geminate and post-singleton vowels is more enhanced in initial than in medial position. Importantly, there was a significant interaction between group and length, $F(1,46) = 21.3$, $p < .001$, indicating that the contrast between geminates and singletons was larger for the Lebanese group compared to the Moroccan group.

6.7. Discussion

Given that Moroccan and Lebanese are both dialects of Arabic, it was expected that there would be some common acoustic and articulatory characteristics in the way their speakers

produce the geminate/singleton contrasts. However, given that the two dialects are spoken in two geographically distinct areas of the Arab world, it was also expected that Moroccan and Lebanese speakers would exhibit some acoustic and articulatory differences in the realisation of the contrasts.

The articulatory analysis generated by TRACTUS from ultrasound data revealed that, for the most part, word-initial geminates in these two dialects are produced with longer articulatory closure than singletons, which is in agreement with most articulatory investigations on gemination. The longer closure for geminates relative to singletons was maintained across both stop consonant voicing types and more importantly, in both utterance positions. This finding was confirmed statistically for all speakers of Moroccan. However, only three speakers of Lebanese differentiated between the two consonant types in both positions, while the other two did not do so consistently. In particular, L2 and L5 did not produce a length contrast in utterance-initial position despite their maintenance of the contrast in the utterance-medial position. Perhaps not surprisingly, around 20% of the geminated tokens in the utterance-initial position produced by these two speakers (and only these two speakers) were accompanied by a pre-geminate schwa-like vocalic element insertion. This inserted schwa appeared to be preceded by a glottal stop, which is consistent with the fact that vowel-initial morphemes are phonotactically illegal in Arabic and its dialects. This vocalic element insertion, along with failing to produce significant differences between a geminate and a singleton only in the utterance-initial position, might suggest that initial gemination in absolute-initial position is undergoing a process of neutralisation in Lebanese Arabic.

For the tokens preceded by an epenthetic vowel, the duration of the stop closure of geminate tokens was on average around 30 ms longer than for the tokens produced without the pre-closure inserted vowel. This mechanism could be interpreted as an effort exerted by those

two Lebanese speakers to maintain the geminate/singleton contrast utterance-initially. Therefore, this might suggest that they used this vocalic element as leverage to help them launch their absolute initial geminates. This makes their pre-closure vocalic insertion instances structurally similar to medial gemination. It is important to note that these two Lebanese speakers did not use this strategy in their utterance-medial productions, i.e., there was no vowel inserted between the /f/ of the preceding word in the carrier phrase and the target consonant. This might suggest that word-initial gemination in utterance-medial position may have been articulatorily easier for them to produce.

Initial gemination is more productive in Moroccan than in Lebanese. A study on the frequency of initial clusters in a number of Arabic dialects including Moroccan and Lebanese found that initial consonant clusters are three times more frequent in Moroccan than in Lebanese (Hamdi, Barkat-Defradas, Ferragne & Pellegrino, 2004). Therefore, the inability of the two Lebanese speakers to systematically contrast geminates with singleton in absolute initial position could be related to the lower productivity word-initial gemination in Lebanese.

The acoustic analysis revealed three additional significant acoustic cues: one durational cue (the duration of the post-consonantal vowel) and two non-durational cues (the amplitude of the stop release burst and f0 of the post-consonantal vowel). The realisation of these cues seemed to be dialect-specific and also varied according to the position in the utterance. In both dialects and across both positions, vowels following a geminate were lengthened more than their counterparts after singletons. This is consistent with the results found in some languages with initial gemination, such as Japanese. In terms of dialectal differences, Lebanese vowel durations were not only longer, but the contrast between post-geminate and post-singleton vowels was larger than was found for Moroccan. The fact that Lebanese speakers produced longer vowel durations might be explained by their shorter closure durations compared to the

Moroccan speakers. So, there seems to be a coarticulatory effect between the consonant duration and the following vowel duration. A stronger contrast between vowel durations for Lebanese suggests that Lebanese speakers rely more on the post-geminate vowel to implement the geminate/singleton contrast word-initially. Interestingly, this effect was witnessed even with the non-temporal cues we tested, from the burst amplitude and the f0 of the post-consonantal vowel results as well, i.e., the difference between geminates and singletons with regard to these acoustic measures was larger for Lebanese than Moroccan. Similarly, the contrast between the geminate and singleton values from the results of the vowel onset amplitude is slightly more enhanced for Lebanese than Moroccan. Interestingly, both groups used release burst amplitude contrastively in the utterance-initial position, but not in the utterance-medial position. Moreover, higher f0 values were found in utterance-initial position compared to utterance-medial position.

It important to note that some of the differences between Moroccan and Lebanese in realising the voiced and voiceless stimuli could be due to stress pattern differences between both dialects. One of the examiners of this thesis noted that, at least in Lebanese Arabic, the words with the voiced stops are of an iambic structure, whereas their voiceless counterparts are in a trochaic structure. This shift in stress pattern could potentially affect the phonetic realisation of the post-consonantal vowel.

Strengthening of acoustic cues beyond the closure duration is supported by theories such as Enhancement Theory (Stevens, Keyser & Kawasaki, 1986), which posits that the primary cue is usually supplemented by other secondary phonetic cues, whose job is to enhance the salience of the primary cue. Our results showed that closure duration is a robust cue in distinguishing geminates from singletons regardless of utterance position. However, we noticed that other non-closure duration cues are particularly enhanced in utterance-initial

position where lengthening of the closure is articulatorily difficult to achieve. Thus, phonetic cues such as release burst amplitude and f0 of the post-consonantal vowel, appear to enhance closure duration, especially when the latter is less robust acoustically.

These acoustic results collectively might challenge the traditional account that geminates are based primarily on a distinction of closure duration. The fact that geminates were distinguished from singletons by greater stop release burst amplitude, higher f0 of post-consonantal vowel and longer durations of the post-consonantal vowel — even in the case of voiceless stops in absolute initial position, where closure is not acoustically evident — might indicate that word-initial gemination is characterised by a forceful articulation. That is, the primary articulatory goal of gemination in this position might not be specifically to lengthen geminates relative to singletons. Rather, lengthening may be a byproduct of a more forceful/tense articulation for geminates than for singletons. Converging evidence for this possibility comes from the results for speakers L2 and L5, who used release burst amplitude, vowel duration and f0 to distinguish geminates from singletons in utterance-initial position, despite not having produced significant closure duration increases in their articulatory data for geminates across their productions. Nevertheless, this interpretation of geminates as marked by a forceful articulation is not in conflict with lengthening of closure duration for geminates, but rather the two can go hand in hand – articulatory force could be one mechanism for achieving/maintaining longer closures.

6.8. Conclusion

This study conducted an acoustic and articulatory investigation of word-initial gemination. The results challenge traditional assumptions that consonant length contrasts are primarily due to differences in the acoustic duration of the closure. Ongoing work will aim at

additional cues that might contribute to the geminate/singleton contract in order to further test whether gemination is the result of a lengthening process or a forceful articulation.

CHAPTER 7. THE ACOUSTICS AND ARTICULATION OF MEDIAL AND FINAL GEMINATION IN MOROCCAN AND LEBANESE ARABIC

7.1. Abstract

An acoustic and ultrasound investigation was conducted on voiced and voiceless stop gemination in Modern Standard Arabic as produced by Moroccan and Lebanese speakers. The roles of various temporal and non-temporal cues were assessed in signaling the word-medial geminate/singleton contrast as compared to the typologically rarer word-final gemination, which some Arabic scholars have suggested is neutralised in Modern Standard Arabic. Ultrasound video of the tongue was analysed via the Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; Carignan, 2014), which allowed us to capture and quantify temporal differences between geminates and singletons even for tokens whose closure duration could not be measured from the acoustic signal. The results showed that while acoustic and articulatory closure duration distinguish between geminates and singletons in both dialects and across both positions, it is more robust in word-medial position. Final gemination in domain-final position was associated with strengthening for all measures that are associated with forceful articulation, except closure duration, and with increased variability in the use of these cues across the two dialects. These findings provide evidence that gemination in MSA is contrastive in both word positions, but that non-closure duration properties are augmented in word-final position.

7.2. Introduction

This study examines dialectal differences in the realisation of word-medial versus -final coronal stop gemination (/t, d/ versus /t:, d:/) in Modern Standard Arabic (MSA), using both articulatory and acoustic data, from speakers of two geographically distant Arabic varieties, Moroccan and Lebanese. Most importantly, the study investigates the acoustic and articulatory correlates of the geminate-singleton contrast to determine whether both temporal cues (such as closure duration) and non-temporal cues (such as burst amplitude) convey a length contrast in the rare and understudied final gemination, similarly or differently to gemination contrasts in the more common word- or utterance-medial position.

Gemination, or the geminate/singleton contrast (used interchangeably), is a phonetic and phonological phenomenon whereby a geminate, or long consonant, is articulated with a longer duration (among other features) relative to a singleton, or a short consonant. Gemination can occur in different positions of the word. The most common type of gemination is the one occurring word-medially (Thurgood, 1993). As such, most knowledge about the phenomenon comes from experimental studies on medial gemination. Compared to the more commonly investigated medial contrast, final gemination has greater potential to provide insight about positional and prosodic influences on the acoustic and articulatory properties of gemination. Final gemination, unlike medial gemination, can line up with the final boundary of various prosodic domains, so it allows examination of aspects of geminate articulation that may be affected by phrase-final strengthening. Moreover, the weaker acoustic cues to closure duration that are associated with peripheral positions of the contrast commonly lead to neutralisation of the contrast word-finally (e.g., Hungarian: Curtis, 2003). As such, focusing on final gemination allows us to tease apart the acoustic and articulatory properties that may be enhanced due to prosodic boundary effects and those that are neutralised word-finally.

In previous research related to the current study, we investigated prosodic effects on the realisation of word-initial gemination in Moroccan and Lebanese vernacular dialects (Frej, Carignan, Proctor & Best, 2018). We observed that acoustic cues that characterise a forceful articulation of word-initial gemination are particularly enhanced in absolute utterance-initial position (the leading edge of the prosodic domain) as compared to utterance-medial position (analogous to medial gemination), with more dialectal variation in the utterance-initial position. Based on these results, we hypothesise that word-final gemination may open up an opportunity for non-temporal acoustic cues to come into play even more robustly than in medial position. Moreover, they may vary more across two disparate varieties of a language because the balance between segmental and prosodic goals (i.e., between lengthening closure for gemination and phrase-final prosodic processes) can differ between disparate varieties of a language.

Gemination is contrastive in MSA. That is, the long and short versions of the same consonant function as separate, contrastive phonemes. For instance, the length difference between the phoneme /r/ in /da**r**asa/ 'to study' versus /da**r:**asa/ 'to teach' is what gives these words distinct lexical meanings. In MSA, gemination is permitted in two positions of the word, word-medially and -finally. A common grammatical process that often results in word-medial gemination is the generation of the causative form of most verbs. This turns the consonant of the second syllable of the non-causative verb stem into a geminate, as in the example above ('teach' ~ to cause someone to study) and in /a**k**ala/ 'to eat' versus /a**k:**ala/ 'to feed' (to support or cause another to eat).

The condition under which word-final gemination arises in MSA is also grammatical but is the result of a process of morpho-phonological reduction, whereby a case ending is dropped in pre-pausal (e.g., utterance-final) environments. Case ending or case marking refers to the inflectional suffixes ('a', 'i', 'u', 'an',' 'in' and 'un') that are attached to the end of

Arabic words to indicate the grammatical function of the word in a sentence. When the word occurs utterance-finally before a stop, or phrase-finally before a comma, the case ending is omitted to mark a pause. Thus, there are two forms of pronunciation: (1) pausal forms that lack the full inflectional suffix and (2) the fully suffixed non-pausal forms. Final gemination occurs only when words are pronounced in their pausal forms (See Table 3.5).

In Classical Arabic, dropping the marking of case was restricted to words occurring in utterance-final position, i.e., at the end of sentence, which involves some amount of pause in the speech stream. However, MSA speakers have gradually tended to pronounce words without their case ending even when words are in utterance-initial, -medial, or in isolated form. This is done either unintentionally due to their limited knowledge of the Arabic grammar, or intentionally to imitate the pronunciation of spoken dialects where case ending is not a common feature. In this regard, Belazi (1984) argues that case ending is gradually disappearing from MSA, not just when the speaker pauses at the end of an utterance, but also when there is no pause. This makes the pronunciation of case ending somewhat flexible and optional. If the pronunciation of suffixes that mark case ending were mandatory for words at all positions of the utterance, even pre-pausal, then word-final consonantal gemination would not occur in MSA. Thus, final gemination is not a purely phonological distinction, but is a result of morphophonemic and sociolinguistic processes. Moreover, unlike medial gemination, final gemination in Arabic can occur only in the phrase-final position. This is not the case for languages where final gemination is contrastive at the purely lexical level (e.g., Norwegian: Payne, Post, Garmann & Simonsen, 2017) as opposed to morphologically derived gemination, which is understudied.

Most acoustic studies on gemination have found closure duration to be a reliable cue for distinguishing geminates and singletons in Bengali (Lahiri & Hankamer, 1988), Hindi

(Ohala, 2007), Icelandic (Pind, 1995, 1999) and Estonian (Engstrand & Krull, 1994) inter alia. Additional non-closure durational cues which contribute to the geminate-singleton contrast are cues associated with tenseness (Ridouane, 2007). Tenseness, also referred to as articulatory strength (Al-Tamimi & Khattab, 2011) or fortition (Hamzah, 2013), is an articulatory feature whereby a tensed sound is articulated with a stronger muscular effort or constriction, compared to its non-tensed counterpart (Matthews, 2014). Tenseness is contrastive in some languages, such as Korean, which distinguishes among three different versions of the same consonant based on the level of tenseness of the glottal folds. Tenseness is often conveyed through a variety of acoustic and articulatory cues, such as higher amplitude of the release burst and f0 of the preceding and following vowels. However, one of the main features that is often reported to accompany tensed consonants is length (i.e., the duration of the constriction). This makes it debatable whether tenseness results from, or causes, a longer closure. Length is another feature that is often talked about when describing gemination, because it is commonly claimed in the literature that the contrast between a geminate and a singleton is based purely on the duration of the consonant (Ham, 1998). That is, a geminate consonant is argued to be identical to its singleton counterpart except that their length is different. This claim, however, has been refuted by numerous research studies on gemination. Studies investigating cues other than the duration of the closure have found that the geminate-singleton contrast is often carried by non-temporal cues, such as higher amplitude of the burst and higher f0 of the flanking vowels (e.g., Kelantan Malay: Hamzah, 2013). Importantly, because these non-closure duration features are associated with tensed consonants and are particularly enhanced in the context of geminates at the word-periphery, this has led some researchers to suggest that the phonological and articulatory nature of gemination may be a forceful articulation distinction, rather than a length distinction (Ridouane, 2007). In this study, we hypothesise that medial gemination would be

distinguished primarily through differences in closure duration. However, non-closure duration cues would be more enhanced for final geminates than singletons.

Word-final gemination is a typologically marked type of gemination, since only a very small number of languages allow gemination to occur in the final position of the word. These languages include Berber (Ridouane, 2007), Hungarian (Polgárdi, 2005), Swiss German (Kraehenmann, 2001), and Norwegian (Lunden, 2006) as well as Arabic, which is the focus of the current study. In Hungarian, final geminates occur mostly in words borrowed from other foreign languages such as English (e.g., sokk [sok:] 'shock'), and is especially productive in monosyllable words (Magyar, 2014). In Swiss German, the geminate-singleton contrast is maintained primarily through differences in duration of the closure: geminate closures are almost twice as long as their singleton counterparts (Kraehenmann, 2001). Berber is another language that allows gemination across different positions of the word, including word-finally (e.g., /ifis/ 'jackal' /ifis:/ 'he was quiet'; Ridouane, 2010).

One explanation for this typological dispreference for word-final gemination could be due to the contrast being less discriminable in this position (Dmitrieva, 2012). Studies have shown that even speakers of languages with no geminate-singleton contrast can attend to durational differences of the consonant closure when the contrast occurs word-medially and intervocalically (e.g., Dmitrieva, 2012). In this position, the vowels on either side of the consonant are thought to provide a suitable acoustic context for consonantal length differences to be easily perceived. This is not the case for word-final gemination: due to the absence of a following vowel and the corresponding lack of vocal fold vibration, the acoustic cues of the consonant closure duration are less discriminable. Some have reasoned that this, in turn, may have led to diachronic neutralisation of the word-final geminate/singleton contrast in some languages, like the Afro-Asiatic language Afar (Blevins, 2004). Moreover, word-final stops are

often produced without a release, making the end of the closure inaudible especially for voiceless stops, for which there is no vocal fold vibration during the closure to provide acoustic information about temporal duration of the consonant. On the other hand, there are reports of the maintenance of gemination word-finally in a number of languages, such as Berber (Ridouane, 2007) and Swiss German (Krehanmann, 2001), mainly because the stops are released and therefore acoustically audible in these languages (Blevins, 2004). Thus, reduced acoustic cues to gemination does not always lead to neutralisation of gemination in the final position of the word.

The focus of the few studies on word-final gemination in Arabic has been on its phonological representation. An optimality-theoretic analysis of the phonological status of final gemination in Jordanian Arabic (Abu-Abbas, Zuraiq & Abdel-Ghafer, 2011) concluded that final underlying geminates are degeminated (i.e., the contrast with singletons is neutralised) only when they are preceded by a long vowel in a monosyllabic word (e.g., /ʕa:mm/ 'general' verus /ʕa:m/ 'year'). The authors proposed that neutralisation stems from a restriction in the language regarding how many morae a monosyllabic word can contain (NB: geminates comprise two morae, whereas singletons comprise one). It is important to note that there is controversy over the distinctiveness of final gemination in Arabic. Ghalib (1984) holds that geminates occurring word-finally are non-distinctive in MSA, alleging that words in MSA bearing a final geminated consonant are pronounced exactly as their non-geminated counterparts. He further explains that Arabs do not distinguish between the Arabic words /dam/ and /dam:/[13] in their pausal forms. They both signify 'blood'. Conversely, El Saaran (1951) argued that the contrast is maintained word-finally, providing as evidence examples of geminate/singleton word pairs that attest to the distinctiveness of final gemination in MSA

[13] The Arabic word for 'blood' is ambiguously transcribed with or without geminating the final /m/.

(e.g., ħa:d/ ‘deviated’ and / ħa:d:/ ‘sharp’). All of these phonological studies, nonetheless, remain impressionistic in nature, as they are not backed up by any empirical data. Therefore, detailed experimental examination of the articulatory and acoustic properties of the final geminate/singleton contrast in MSA is required in order to investigate whether the gemination contrast is maintained or neutralised in final position.

To our knowledge, there are only two published experimental studies investigating word-final gemination, and they examined dialectal Arabic rather than MSA. Al-Tamimi, Abu-Abbas & Tarawnah (2010) addressed whether the word-final gemination contrast in Jordanian Arabic is distinctive. They measured the preceding vowel duration and stop closure duration for the oral stops /d, d:/ and the preceding vowel duration and duration of nasal murmur for the nasal stops /n, n:/ and /m, m:/. They found that the closure was 1.5 times longer in geminate than in singleton consonants, and conversely that vowels were 1.4 times longer preceding singleton than geminate consonants. They also found that the durations of the closure and the preceding vowel together offer sufficient acoustic support for native listeners to differentiate between geminates and singletons word-finally. Therefore, the authors concluded that word-final gemination is contrastive in Jordanian Arabic dialect. More recently, Al-Deaibes and Rosen (2019) found comparable results in an investigation of final gemination in Rural Jordanian Arabic. The maintenance of final gemination in these two Arabic dialects does not guarantee the maintenance of the final contrast in MSA as the differences between Arabic vernaculars and MSA are well attested in the literature (see: Altoma, 1969; Al-Tamimi & I’lawi, 2006; Amer, Adaileh & Rakhieh, 2011).

However, with the exception of one pilot experiment related to the present study (Frej, Carignan & Best, 2017) there is no existing articulatory or acoustic evidence concerning whether final gemination is contrastive or not in MSA. In that pilot experiment with only two

Moroccan MSA speakers, the results were inconclusive. While both speakers employed closure duration to differentiate geminates from singletons word-medially, they seemed to adopt different strategies in word-final position. One speaker contrasted geminates with singletons word-finally through closure duration, whereas the other did so by differentiating release burst duration for both voiced and voiceless stops. Given the scarcity of research on languages with final gemination and on Moroccan in particular, it was unclear whether the second speaker's release burst-duration technique for contrasting geminates with singletons word-finally was idiosyncratic or might be shown by other speakers. Therefore, it was necessary to replicate the experiment with more participants to see if other speakers use this somewhat atypical pattern of release burst-duration contrast for final gemination.

Phrase-final lengthening is well documented in the literature, and it is a common feature in most languages (Blevins, 2008). Phrase-final syllables ending with a stop were reported to be lengthened compared to phrase-medial ones in English (Oller, 1973). Phrase-final strengthening is typically considered to be a strategy used by speakers to cue the boundaries of words, phrases, or sentences. Word-final gemination is well suited for testing the principles of domain-final strengthening, because unlike word-medial gemination, word-final gemination can occur in the final position in an utterance.

In contrast to the strengthening at the edges of a prosodic domain, final gemination is prone to being neutralised in other final-geminating languages (Hungarian: Curtis, 2003; Russian: Kasatkin & Choj, 1999), though as noted it is not clear whether it is neutralised in MSA. Neutralisation of final gemination, i.e., degemination processes, could result in loss of the contrast word-finally. Given that final gemination phrase-finally can be produced without an acoustically audible burst, the closure duration per se may be inaudible especially for voiceless stops in absolute final position. As previously mentioned, there is no acoustic context

to provide a perceptual cue for the offset of the closure if it is produced without an audible release burst. Poor acoustic cues to closure duration in word-final positions could lead to diachronic neutralisation of the contrast (Blevins, 2004). Importantly, these poor acoustic cues make acoustic measurement of closure duration for these segments impossible. Thus, it is necessary to apply articulatory (especially kinematic) techniques in order to investigate the articulatory characteristics of the geminate/singleton contrast.

Methods of studying speech articulation are most useful to measure the closure duration when this closure is not measurable acoustically. Various techniques such as EPG (Ridouane, 2007) and oral airflow (Ridouane, 2003) have been used in these cases. The results from these studies suggest that geminate stops are produced with a longer articulatory contact (i.e., extended over a longer period of time) compared to singletons. However, the most informative data would capture the actual temporal dynamics of articulator movements in the formation of singleton versus geminate stop closure, whereas EPG and oral airflow provide only indirect estimates of those motions. More direct techniques, such as EMA or ultrasound, would be particularly useful for this, but have been used only sparsely in research thus far.

We sought to determine whether temporal and non-temporal cues in the acoustic signal convey a gemination contrast between the coronals /t, d/ and /t:, d:/ in word- and utterance-final position (relative to medial positions), and whether articulatory analysis using ultrasound further supports and delineates the characteristics of closure duration differences. This study also aimed to resolve the controversy over the contrastiveness versus neutralisation of final gemination in Arabic, as well as to examine whether gemination results from forceful articulation and/or from closure elongation (and whether the balance between these two sources of contribution to gemination is dialect-specific). Moreover, our comparison of word-final and word-medial gemination will determine whether speakers use similar or different acoustic and

articulatory features in producing gemination contrasts, which will provide insights on general principles of how gemination is achieved and how it is affected by prosodic position in a word or utterance.

We hypothesise that both acoustic and articulatory closure duration are strong cues in signalling the geminate/singleton contrast. We predict that geminates will be produced with a longer closure duration than singletons. Moreover, we expect that differences in other temporal parameters, such as preceding vowel duration, release burst duration and non-temporal parameters (i.e., preceding vowel offset f0 and amplitude of the burst and preceding vowel) will be indicative of a length contrast between a geminate and a singleton in the word-medial position and, importantly, in the word-final position where gemination is prone to being neutralised. We also hypothesise that 1) there will be dialectal differences between Moroccan and Lebanese affecting the realisation of these parameters and 2) certain cues, other than closure duration, will be particularly enhanced in the final position of the word compared to word-medially.

7.3. Method

7.3.1. Participants

[14]Five male speakers from Morocco ($Mean_{age}$ = 32.8 years, range = 27-40) and five from Lebanon ($Mean_{age}$ = 31.4 years, range = 24-37) were recruited to participate in the experiment. Only male participants were recruited to avoid introducing possible effects of gender on the production of gemination, as females have been reported to produce gemination in some word positions with smaller average durations than males (Kelantan Malay: Hamzah,

[14] It is important to note that the speakers for this experiment are the same speakers recruited for the experiment on initial gemination, and the data collection took place on the same settings (Chapter 6).

2010). All participants were born and educated through secondary school level in their respective countries. They had lived in Australia for 1-4 years ($Mean_{Moroccan}$ = 3.2 years, $Mean_{Lebanese}$ = 3.4 years) at the time they participated in the experiment. They reported no hearing, vision or speech impairments.

7.3.2. Target stimuli

For word-final gemination, a list of four target items with two minimal pair gemination contrasts of voiced and voiceless coronal stops was used (Table 7.1). The words were carefully selected with the aim of forming an organised paradigm with a constant frame in terms of lingual articulation except for phonological length in the word-final gemination contrast. As such, we held constant everything else that could affect the phonetic details of the target consonant. The words were grouped according to voicing profile of the final consonants: voiced coronal stops (/d/-/d:/) and voiceless coronal stops (/t/-/t:/). All target items are disyllabic real words in MSA, displaying the structure ʔaCaX(:), with a glottal stop /ʔ/ followed by the short vowel /a/ in both the first and second syllable and X being either a singleton or a geminate. The medial consonant C is the nasal stop /m/ for the target words /ʔamad:/ versus /ʔamad/, but it is the oral stop /b/ for /ʔabat:/ versus /ʔabat/ due to a lack of a real word minimal pairs in MSA with the same medial consonant in the same context for both voiced and voiceless final stops. This gives a near-minimal quadrad in which the medial consonant difference should not affect lingual articulation, as both are bilabial stops that differ only in terms of nasalisation (i.e., neither would impose coarticulatory effects on tongue tip movement in the target word-final coronal stops).

Table 7.1: List of word-final voiced versus voiceless coronal stop gemination stimuli in MSA with their phonetic transcriptions and their English glosses.

	/d/		/t/	
	IPA	**English Gloss**	**IPA**	**English Gloss**
Singleton	/ʔamad/	(n. 'limit')	/ʔabat/	(v.i. 'she refused')
Geminate	/ʔamad:/	(v.t. 'to provide')	/ʔabat:/	(v.t. 'to cut off')

For word-medial gemination, a similar minimal quadrad of four disyllabic MSA real words were chosen that display the structure ʔaXab, with X being either a singleton or a geminate (Table 7.2).

Table 7.2: list of word-medial stimuli and their English glosses.

	/d/		/t/	
	IPA	**English Gloss**	**IPA**	**English Gloss**
Singleton	/ʔadab/	(n. 'literature')	/ʔatab/	(v.t. 'to weaken')
Geminate	/ʔad:ab/	(v.t. 'to discipline')	/ʔat:ab/	(v.t. 'to dress up a girl')

Participants were asked to produce each of the eight target items 10 times in a carrier sentence, and 10 times in isolation. The carrier sentences were [galt rahaf ____] for Moroccan and its equivalent in Lebanese [alit rahaf ____] ('Rahaf said ____'). The carrier sentences with the target word inserted were presented in random order in a PowerPoint presentation. Participants were instructed to read the stimuli without pausing after the carrier sentence. The stimuli and the carrier sentences were presented to the speakers orthographically in Arabic script in order to instigate an MSA style of pronunciation. The speakers were also repeatedly reminded to pronounce the stimuli in MSA, rather than in their vernacular Arabic.

7.3.3. Procedure

Acoustic data were co-collected with ultrasound data using the GE LOGIQ-e ultrasound system with a GE 8C-RS probe, operating at 8 MHz and depth ranging from 8-10 cm. Ultrasound video was extracted from the VGA output of the GE LOGIQ-e in real time using an Epiphan VGA2USB Pro video grabber. FFmpeg software (FFmpeg Development Team, 2016) was used to record a continuous .AVI file with ultrasound video at 30 fps along with synchronised acoustic data, which were collected through USB using a Sennheiser MKH 416 microphone connected to a Sound Devices LLC USB Pre 2 audio interface. A non-metallic head-mounted ultrasound probe holder, designed by Derrick, Best, and Fiasson (2015), was used to stabilise the ultrasound probe relative to the head within acceptable parameters for rotational and translational slippage (Derrick, Carignan, Chen, Muayiwath & Best, 2018), consistent with HOCUS measurements [Haskins Optically-Corrected UltraSound] (Whalen, Iskarous, & Tiede, 2005), which are considered the standard for ultrasound speech research.

Participants were asked to produce the sound sequence “ta, ta, ta” both at the beginning and at the end of recording. This was beneficial for data synchronisation because clear landmarks in the sequence could easily be identified in both the articulatory and acoustic signals. Synchronisation was carried out by extracting the ultrasound and acoustic data between these landmarks.

7.3.4. Articulatory analysis

Ultrasound images were analysed semi-automatically using the TRACTUS Matlab functions. TRACTUS analysis identifies a set of principal component (PC) coefficient matrices for the entire ultrasound recording, which explain the greatest amount of variance in the ultrasound data. The PC scores of frames associated with a closure are used as predictors in

subsequent Linear Discriminant Analysis models (LDA). Thus, the LDA produces a temporal vector based on the probability that a particular image corresponds to a stop closure. The higher the LDA scores displayed by the model, the more likely that an ultrasound frame at a particular time point corresponds to a closure. In this way, TRACTUS performs temporal articulatory analysis on ultrasound images automatically without the need for image tracing.

7.3.5. Acoustic analysis

To perform the acoustic analysis, the data were imported into Praat Boersma & Weenink, 2007), where segmentation was carried out manually using a Praat TextGrid (see Figure 7.1 for an illustration of the segmentations for tokens of the word-medial singleton /d/ and geminate /d:/ and Figure 7.2 for the word-final singleton /d/ and the geminate /d:/ words). The vowel /a/ preceding the target consonant was segmented from the onset of clearly repeating periodicity and formant structure (easily identified due to the initial glottal stop in all target words) until the stop closure. The closure interval was marked from the end of regular periodicity of the vowel, accompanied by noticeable drop of amplitude, up until the burst. The burst of the stop release was identified by a clear spike of noisy energy in the spectrogram and measured until the start of the vowel in medial position and until the cessation of friction in the word-final position. This was possible because Moroccan and Lebanese speakers produced stops with release and aspiration even in word-final position with the exception of few tokens, which we address in subsequent sections.

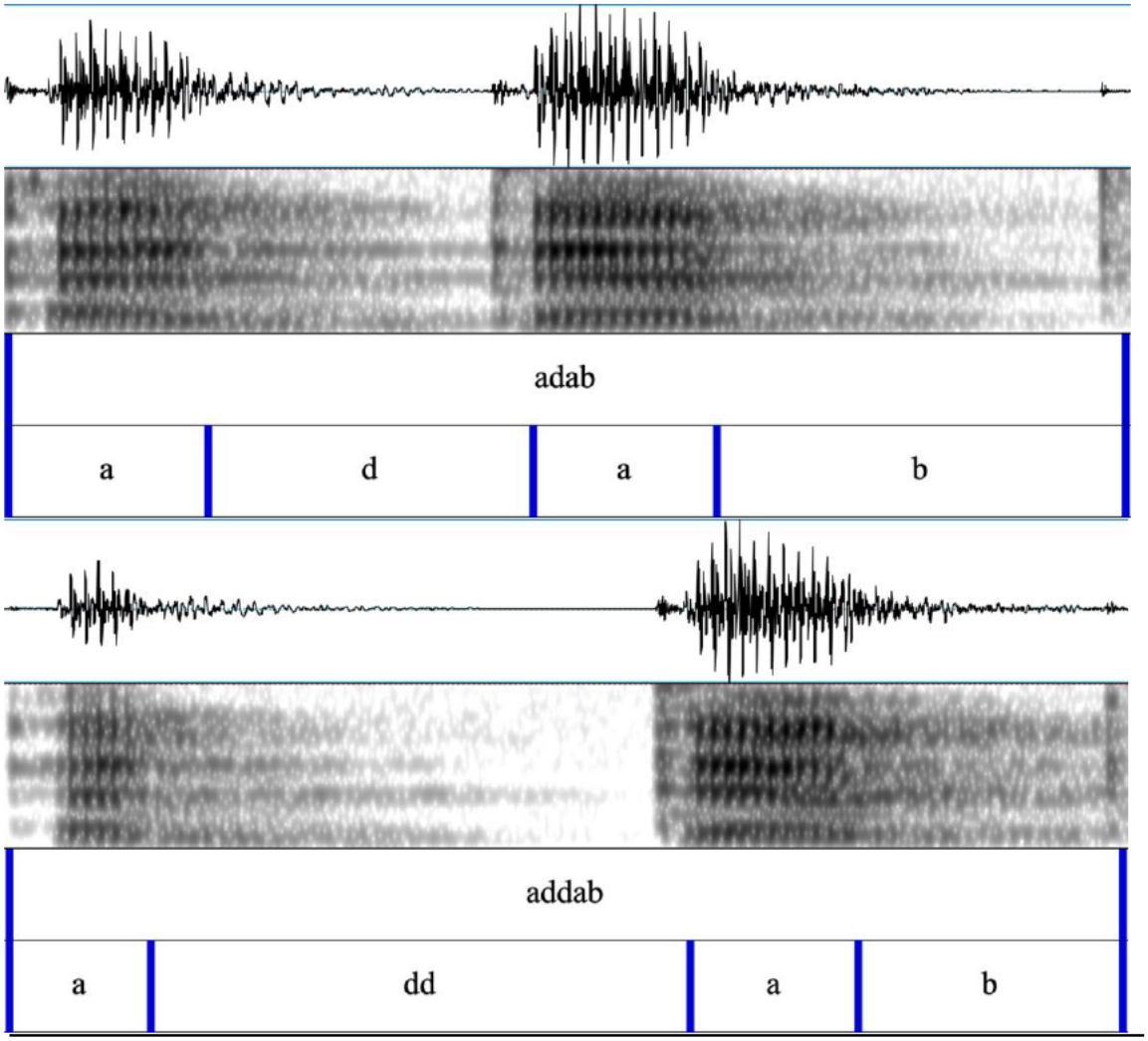

Figure 7.1: The waveforms and spectrograms of the target words for medial gemination, /adab/ and /ad:ab/, with the segmentation landmarks indicated underneath.

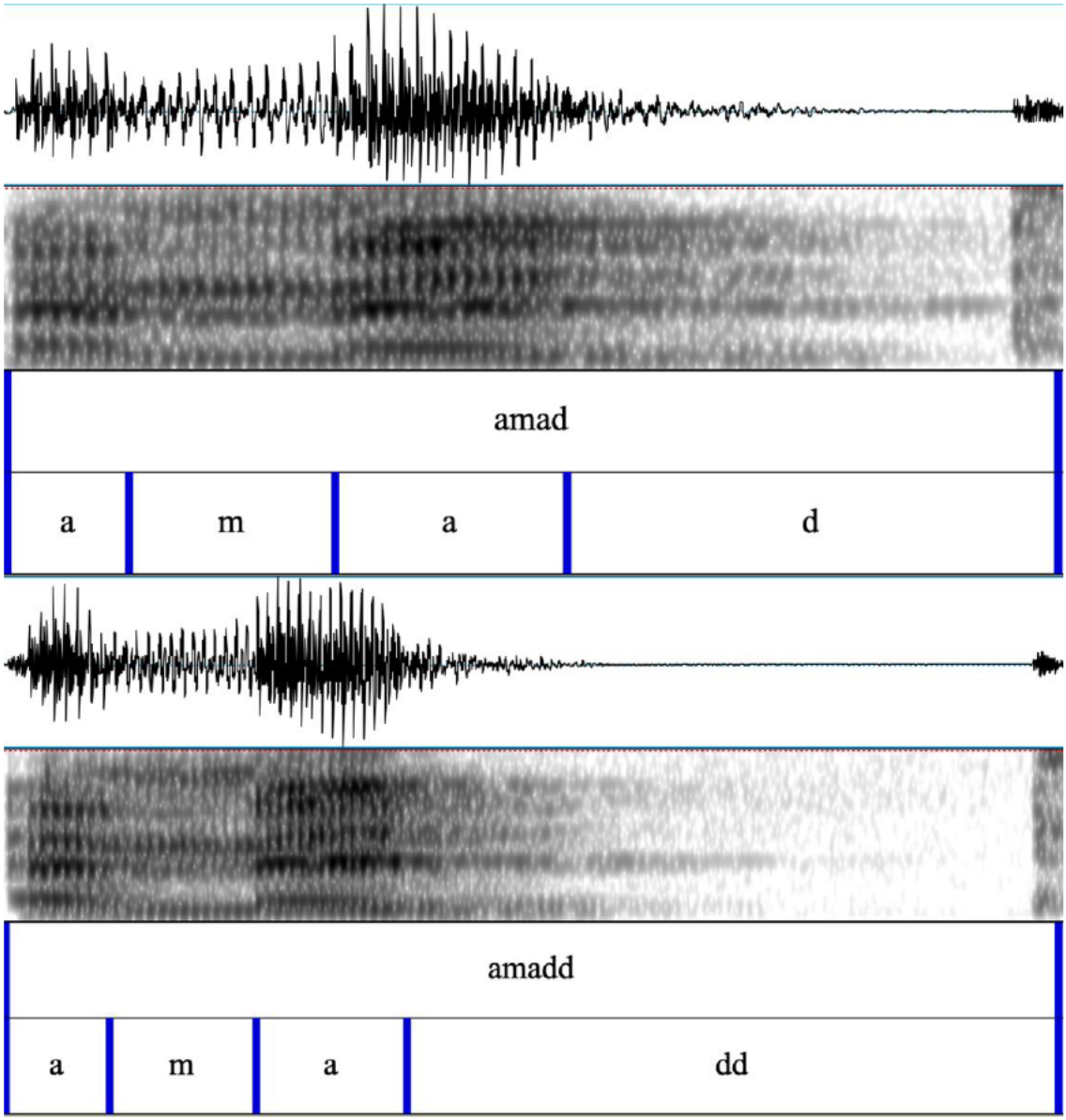

Figure 7.2: The waveforms and spectrograms of the target words for final gemination contrasts /amad/ and /amad:/, with the segmentation landmarks indicated underneath.

Once the segmentation was complete, the following durational measurements were logged using a custom Praat script: closure duration of the target singleton and geminate consonants, release burst duration as well as duration of the preceding vowel. Three additional non-durational measurements were tested: the amplitude of the release burst and preceding vowel as well as fundamental frequency (f0) of the preceding vowel offset. Similar to word-initial data (Chapter 6), we used a separate script based on Hirst's (2011) two-pass method to accurately estimate fundamental frequency at the offset of the vowel.

7.4. Results

7.4.1. Articulatory results

7.4.1.1. TRACTUS analysis

Figures 7.3 and 7.4 display the LDA scores of the articulatory signals over time for each group, as well as category means and 95% confidence intervals generated using *t*-based approximation ("loess" smoothing in the ggplot2 package in R). Because the model has been trained to identify a stop closure, the higher the LDA score assigned at any point in time, the more likely it is that the tongue is maintaining a stop closure. As can be seen, the articulatory signals generated by TRACTUS show clear differences between the geminate articulatory signal (represented by the purple line) and the singleton signal (represented by the green line). Specifically, the geminate signal displays a higher peak and wider interval for the closure than the singleton signal, indicating that the closure is more canonical and maintained for a longer period in geminates. This is consistent with studies that found geminates are produced with a longer articulatory contact than singletons (Ridouane, 2003). The scores also appear to be higher for geminates than for singletons throughout most of the consonantal interval, suggesting that singletons may also be slightly lenited (weaker closure gesture) as compared with their geminate counterparts. There also appear to be some differences between Moroccan and Lebanese articulation of geminates. In particular, the differences between geminate and singleton signals is more apparent for Moroccan than Lebanese; and closures are possibly also maintained slightly longer in Moroccan than Lebanese productions.

These observations of the articulatory characteristics reflected in the TRACTUS signals were statistically evaluated and further delineated as described in the next section.

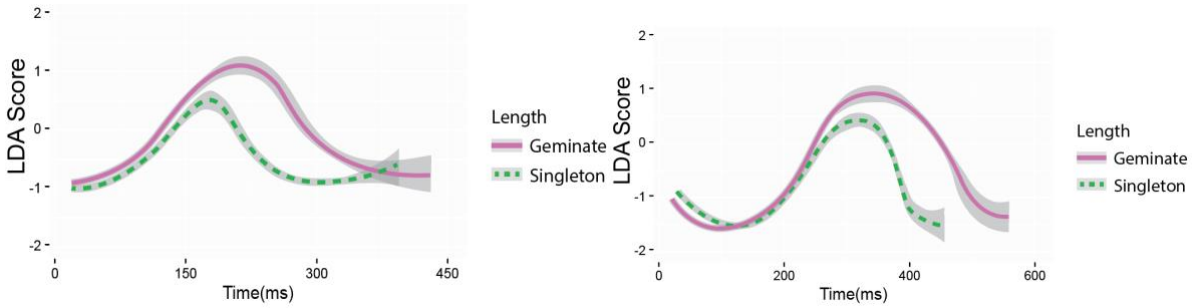

Figure 7.3: Moroccan ultrasound LDA class scores over time for geminates and singletons in word-medial position (left) and word-final position (right). Higher scores along the y-axis indicate greater likelihood of tongue configuration associated with a stop closure.

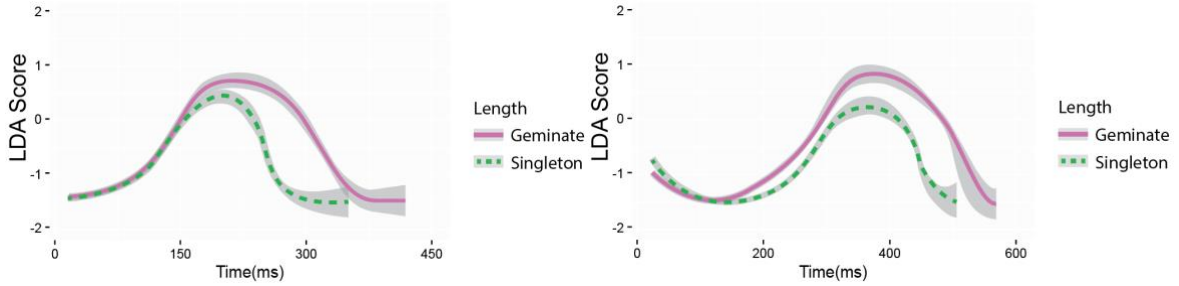

Figure 7.4: Lebanese ultrasound LDA class scores over time for geminates and singletons in word-medial position (left) and word-final position (right). Higher scores along the y-axis indicate greater likelihood of tongue configuration associated with a stop closure.

7.4.1.2. Articulatory closure duration analysis

In order to calculate the articulatory closure duration for all tokens, but importantly for the burstless tokens, which could not have been measured from the acoustic signal, we used a custom Matlab function to estimate duration values of the closure semi-automatically through determining gestural landmarks from the velocity signal. First, the velocity signal was calculated from the articulatory signal generated from the ultrasound video. Next, the time points corresponding to the maximum velocity peak (movement into the closure) and minimum velocity peak (movement out of the closure) were identified for each token. Following this we calculated gesture onset as 20% of the maximum velocity during the closure phase and gesture

offset as 20% of the minimum velocity during the release phase of the articulatory closure. This is common practice (e.g, Mooshammer & Fuchs, 2002) and is quite suitable for stops (Tomaschek, Arnold, Bröker & Baayen, 2018). 20% thresholds yield more reliable articulatory gesture landmark estimation than do lower or higher thresholds (Bombien, Mooshammer & Hoole, 2013) and are more reliable than use of local minima of tangential velocity (zero-crossings) because there is sometimes more than one zero-crossing around the target phase (see Mooshammer, Hoole & Geumann, 2006). From the calculated gestural landmarks we estimated stop closure duration by multiplying the number of frames between the closure onset and offset by the ultrasound frame rate (see Figure 7.5 for an example). This closure duration estimation method was previously validated in (Frej, Carignan, Proctor & Best, 2018).

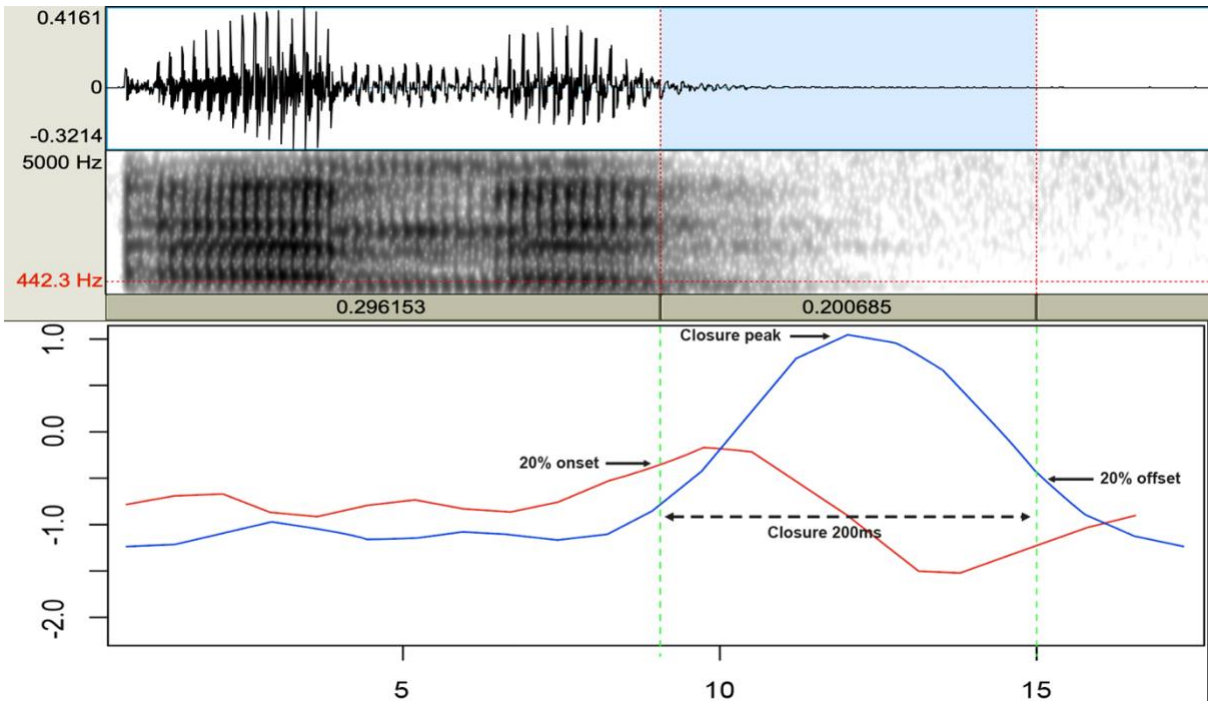

Figure 7.5: Estimation of closure duration. Articulatory signal (blue line), velocity signal (red line) and time-aligned waveform and spectrogram.

The articulatory closure durations for both positions (medial and final) and both dialects are presented in Table 7.3. We submitted the results the estimated articulatory closure durations

to an LME analysis to determine whether there were significant differences between the means of geminates and singletons.

Table 7.3: Mean articulatory closure duration estimation values (ms), *p*-values (LME analysis) and for geminates and singletons.

	Geminate	Singleton
Moroccan		
Medial	210	98
Final	302	186
Lebanese		
Medial	188	102
Final	266	166

As expected, the LME model analysis revealed a significant effect of length, $F(1, 320) = 44.8$, $p < .001$, suggesting that geminates are articulated with a longer articulatory closure duration (M = 269 ms) than singletons (M = 138 ms) regardless of position or dialect. There was also a significant main effect of position, $F(1, 312) = 86.2$, $p < .001$, indicating that the articulatory closure duration in final position (M = 230 ms) was longer than the duration in medial position (M = 149 ms). Moreover, the effect of group was significant, $F(1, 320) = 36.4$, $p < .001$, i.e., Moroccan speakers produced higher values for the articulatory closure duration (M = 199 ms) than Lebanese speakers (M = 180 ms). In terms of interaction, the LME analysis showed a significant interaction between length and position, $F(1, 310) = 99.3$, $p < .001$). A Tukey HSD post-hoc test revealed that this interaction was due to the larger contrast in word-medial position (M = 1.99/1 G/S) compared with word-final position (M = 1.61/1 G/S), $p < .001$. There was also a significant interaction between length and group $F(1, 296) = 20.11$, $p < .001$, revealing that the contrast between geminate and singleton in terms of articulatory closure duration is larger for the Moroccan group compared with the Lebanese group.

These estimated articulatory closure duration results are in agreement with the results displayed by the LDA plots in showing that geminates are produced with longer articulatory

closure duration even when this closure is not acoustically evident, i.e., for tokens produced word-finally without a burst. These results suggest that this technique offers a computationally efficient and robust method for quantifying duration differences for contrasts whose temporal changes are difficult or impossible to discern from the acoustic signal alone.

7.4.2. Acoustic results

7.4.2.1. Statistical analysis

Each of the acoustic dependent variables was analysed using linear mixed effects (LME) models (Baayen, Davidson & Bates, 2008) to test the effect of the following fixed factors: group (Lebanese, Moroccan), position (medial, final), length (geminate, singleton) and voicing (voiced, voiceless). To examine the interactions between factors that best describe the data, differences between various models were generated and evaluated using ANOVAs. All of the dependent variables were centered and scaled for each speaker. Treating Speaker as a random slope resulted in a better model fit than treating it as a random intercept for all the analyses. Random slopes for item, position and length did not improve the model fit. Tukey HSD Post-hoc tests were used to break down interactions and to run multiple comparisons. We generated *p*-values using library lmerTest (Kuznetsova et al., 2017). Any *p*-value of .05 or below was deemed significant.

7.4.2.2. Acoustic closure duration

Manual inspection of the spectrograms revealed that in word-final position only about 15 percent of voiced tokens and about 10 percent of voiceless tokens in Lebanese, and 12 percent of voiced tokens in Moroccan, lacked a release burst. All voiceless tokens in Moroccan were produced with clear release bursts. Importantly, all tokens lacking a release burst were

singletons; all geminate stops, regardless of voicing type and dialect, were also produced with noticeable release burst noise. The lack of release burst information in the acoustic signal, especially for voiceless stops in word-final position makes acoustic measurement of closure duration for these segments impossible. Thus, after applying the correction method we described in Chapter 6, the missing values for the burstless tokens were replaced by their equivalent articulatory closure durations as estimated from the LDA analysis of ultrasound data.

The LME analysis revealed a significant main effect of length, $F(1,320) = 23.7$, $p < .001$, indicating that across all other factors geminates had longer acoustic durations than singletons (M = 260 ms vs 146 ms). It can be seen from Figure 7.6 that closure duration values were greater for geminates than singletons for all Moroccan and Lebanese speakers. There was also a significant main effect of position, $F(1,312) = 32.76$, $p < .001$, i.e., closures in word-final position were produced with longer values than those in word-medial position (M = 264 ms vs. 171 ms for Moroccan, and M = 227 ms vs 152 ms for Lebanese).

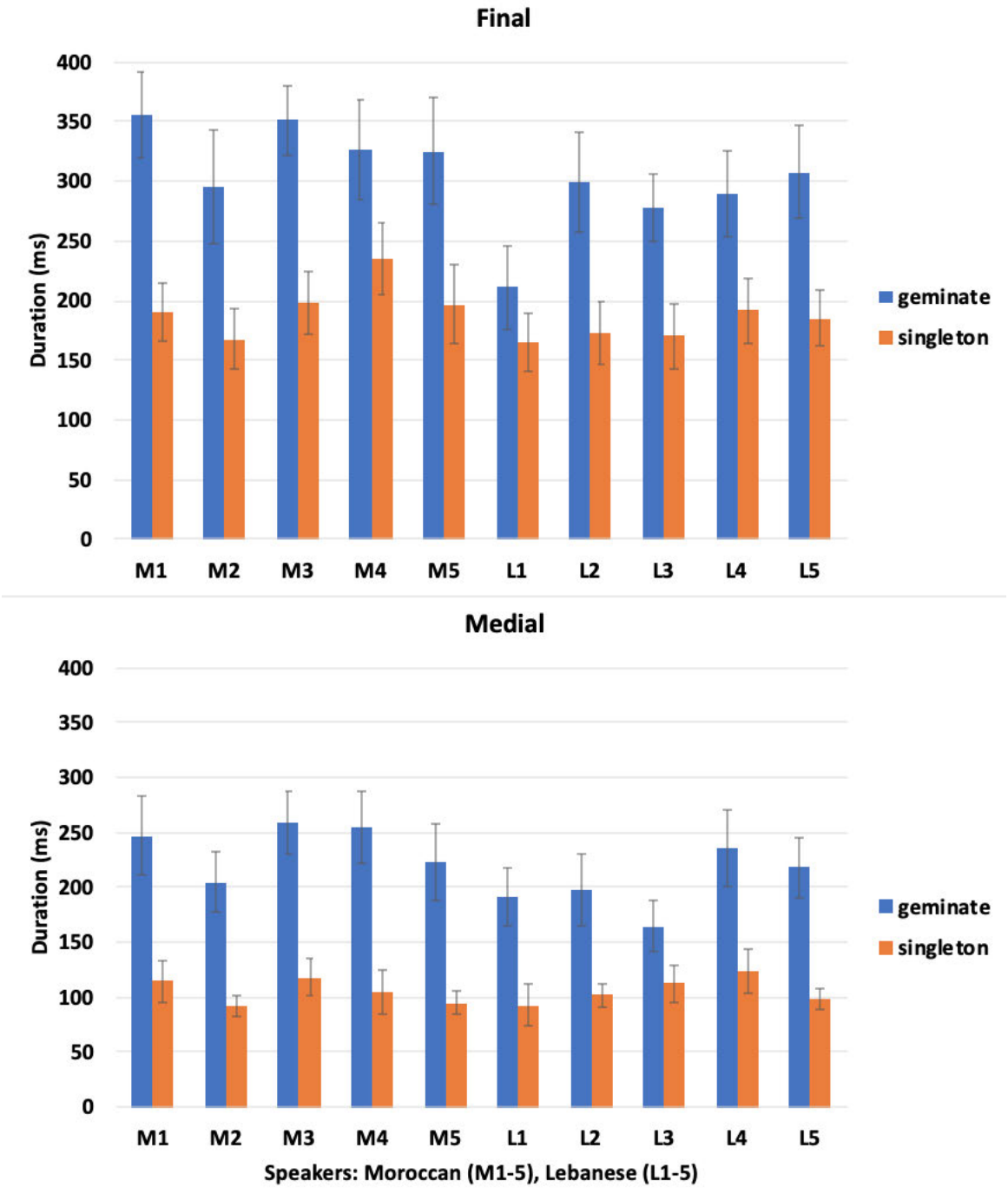

Figure 7.6: Mean closure duration values for each of the geminate and singleton targets displayed by speaker for each group (M1-5: Moroccan; L1-5: Lebanese) in word-final (upper panel) and word-medial (lower panel) positions. Error bars display standard error of the mean.

The LME model revealed a significant main effect of group, $F(1, 320) = 16.8$, $p < .001$, with Moroccan speakers producing longer closures than Lebanese speakers (M = 218 ms vs

189ms). Moreover, voiceless stops have longer closures than their voiced counterparts, as indicated by the significant main effect of voicing, $F(1, 290) = 32.7$, $p = .0112$. In terms of interactions, the LME model showed significant interactions between length and position, $F(1, 310) = 19.7$, $p < .001$. We broke down this interaction using Tukey HSD post-hoc comparisons, which revealed that contrast between geminates and singletons was proportionally larger in word-medial than -final position. Although closure durations were longer, on average, in word-final than in word-medial position, pairwise *t*-tests showed that the geminate-to-singleton ratio was nonetheless larger word-medially than -finally (M = 2.25/1 vs 1.67/1 for Moroccan , $p < .001$; M = 1.85 vs 1.56, $p < .001$) for Lebanese). There was also a significant interaction between length and group, $F(1, 296) = 16.7$, $p < .001$. After running multiple comparisons, it was found out that the Moroccan group produced longer closures than the Lebanese group mainly for geminates (M = 284 ms for Moroccan geminates vs M = 237 ms for Lebanese geminates, $p < .001$). However, closure durations for singletons did not differ significantly between the two groups (M = 151 ms for Moroccan singletons vs. M = 147 ms for Lebanese singletons, $p = .09$).

7.4.2.3. Release burst duration

The role of the release burst duration in contrasting geminates with singletons is not well investigated in the literature. It was shown in previous studies that the presence or absence of gemination affects the duration of the release burst in some languages, such as Tashlhiyt Berber (Ridouane, 2010), but the effect was mostly attested for voiced stops (with the exception of Cypriot Greek, whose inventory lacks voiced stops; Tserdanelis & Arvaniti, 2001).

Table 7.4: Mean release burst duration (ms) and *p*-values (LME analysis) for geminates and singletons in both word positions and for both dialects

Group/position	voiced			Voiceless		
	Geminate	Singleton	*p*	Geminate	Singleton	*p*
Moroccan						
Medial	33	34	n.s	56	56	n.s
Final	39	49	<0.01	69	67	n.s
Lebanese						
Medial	35	37	n.s	55	54	n.s
Final	38	45	<0.01	65	64	n.s

It is important to note that this analysis does not include values from the "burstless" tokens in the word-final position, because their burst duration could not be measured acoustically. Because they constitute only about ~10% of the data, the effect should be minimal and LME models are ideal for handling missing values. Surprisingly, the LME models did not reveal a significant main effect of consonant length on burst duration; it did not vary between geminates (M = 48 ms) and singletons (M = 45 ms), $F(1, 312) = 3.2$, $p = .075$. However, there was a significant main effect of voicing, $F(1, 284) = 33.4$ $p < .001$, i.e., voiceless stops were produced with a significantly longer release burst (M = 54 ms) than voiced stops (M = 34 ms). There was also a significant interaction between voicing and position, $F(1, 308) = 10.6$, $p = .009$. Tukey HSD post-hoc comparisons were applied to break down this interaction by position and voicing type. Results revealed that while gemination did not have any effect on burst duration in voiceless stops (M = 61 ms for voiceless geminates and M = 60 ms for voiceless singletons), voiced geminates were produced with longer release duration than singletons, but this effect was reliable only in word-final position (see Table 7.4). Interestingly, this holds true for both Moroccan and Lebanese. The significant main effect of group, $F(1,312) = 9.7$, $p = .0446$, revealed that Moroccans produced longer release bursts (M = 51 ms) than Lebanese (M = 49 ms). However, this difference (2 ms) is below the Just-Noticeable-difference (JND) for segment duration (Klatt & Cooper, 1975).

7.4.2.4. Release burst amplitude

As mentioned in the introduction, some research has reported that the geminate-singleton contrast in languages such as Berber (Ridouane, 2010) and Finnish (Doty et al., 2007) is reflected in a higher amplitude of release bursts in geminates than in singletons. Likewise, the LME analysis of the burst amplitude showed a significant effect of consonant length, $F(1, 320) = 12.9$, $p < .01$), from which it was revealed that geminates had higher burst amplitudes (M = 56 dB) than singletons (M = 51 dB) (see Figure 7.7). The LME models showed a significant effect of voicing, $F(1, 290) = 43.2$ $p < .001$, i.e., voiced stops had higher burst amplitude (M = 58 dB) than voiceless stops (M = 50 dB). This is consistent with what was found in (Hamzah, 2013) and (Al-Tamimi & Khattab, 2018). Additionally, there was a significant effect of group, $F(1, 320) = 10.5$, $p < .01$, i.e., Lebanese speakers produced higher values for burst amplitude (M = 56 dB) than Moroccan speakers (M = 51 dB). The interaction between length and position was significant ($F(1, 310) = 11.6$, $p < .01$). After running multiple comparisons, it was revealed that for word-final position only, the burst amplitude was higher for geminates (M = 62 dB) than singletons (M = 54 dB), $p < .01$. In word-medial position, there was a tendency for the release burst amplitude to be higher in the geminate (M = 51 dB) than the singleton environment as well (M = 49 dB), but the difference was not statistically significant, $p = .311$. Also, the contrast between geminate and singleton is larger in the word-final position compared to the word-medial position.

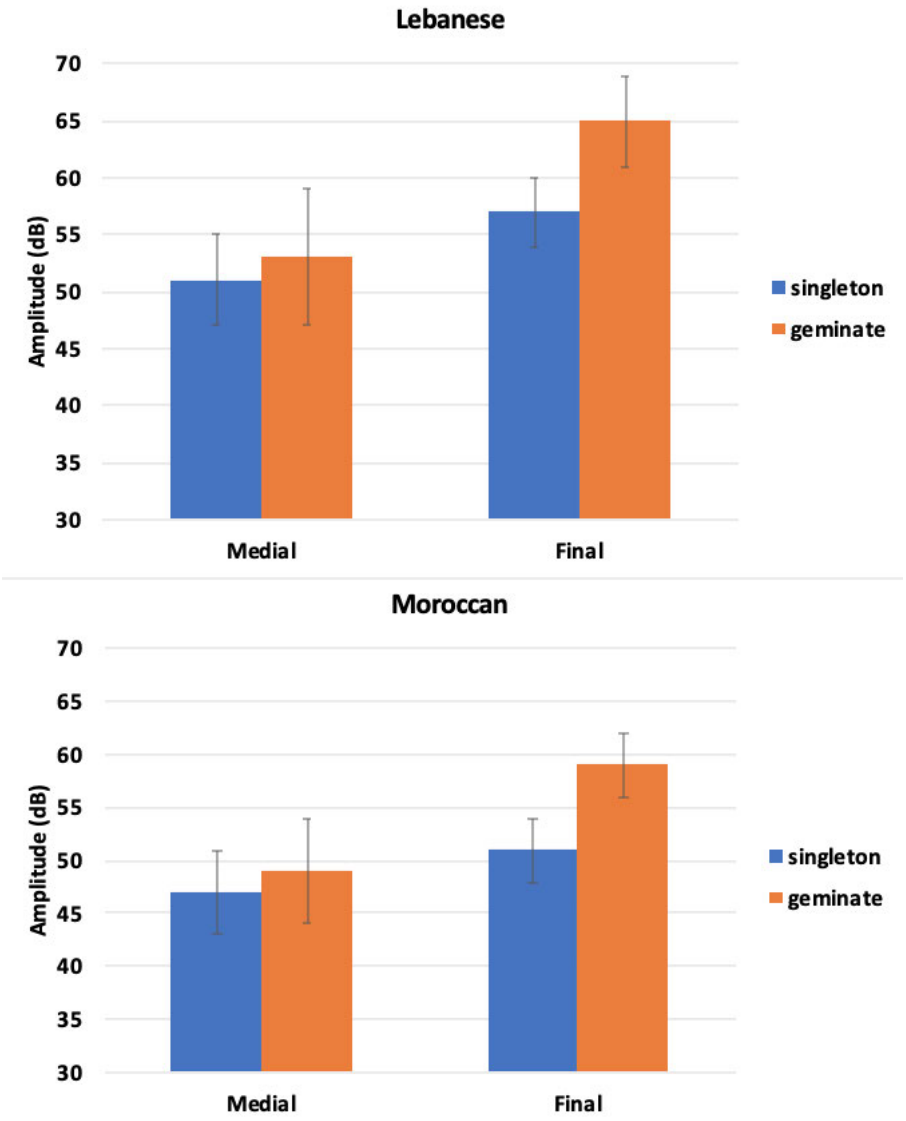

Figure 7.7: Mean amplitude (dB) of the stop bursts for the geminates and singletons produced in word-medial and -final positions by Lebanese (upper panel) and Moroccan (lower panel). Error bars show standard error of the mean.

7.4.2.5. Preceding vowel duration

The results of preceding vowel duration values were also submitted to Linear Mixed Effect modelling analysis. Although there was a tendency for vowels to be shorter before geminates than before singletons (M = 91 ms for singletons vs. M = 83 ms for geminates), the main effect of consonant length on preceding vowel duration approached but did not pass the threshold for significance, $F(1,320) = 5.3$, $p = .056$. The LME model showed a significant main effect of group, $F(1,320) = 33.7$, $p < .001$, i.e., the Lebanese group produced significantly

longer preceding vowel durations. Furthermore, the main effect of position proved to be significant, $F(1, 312) = 85.9$, $p < .001$), i.e., durations of the preceding vowel were longer word-finally than -medially. Another significant main effect was observed for voicing, $F(1, 290) = 94.7$, $p < .001$) indicating that vowels preceding voiced stops were longer ($M = 93$ ms) in duration than vowels preceding voiceless stops ($M = 81$ ms).

There was a significant interaction between length and group, $F(1, 296) = 33.7$, $p = .009$. Breaking down this interaction showed that the difference between vowel values for geminates and singletons was greater for the Moroccan group than for the Lebanese group. Next, we decided to run another LME model in which we specified speaker as a fixed effect in order to look at speaker-specific effects in each group. The results showed a significant main effect of speaker, $F(9, 310) = 25.8$, $p < .001$, as well as a significant interaction between speaker and length, $F(9, 310) = 90.6$, $p < .001$. Tukey HSD post-hoc comparisons for each speaker and position revealed that vowel duration was employed more strongly by Moroccan than Lebanese speakers to contrast geminates with singletons, particularly in word-final position, where four Moroccan speakers and one Lebanese speaker produced significantly shorter vowel durations before geminates than before singletons. In word-medial position, four Moroccan and three Lebanese speakers produced significantly shorter vowels before geminates relative to singletons. This is displayed in Figure 7.8.

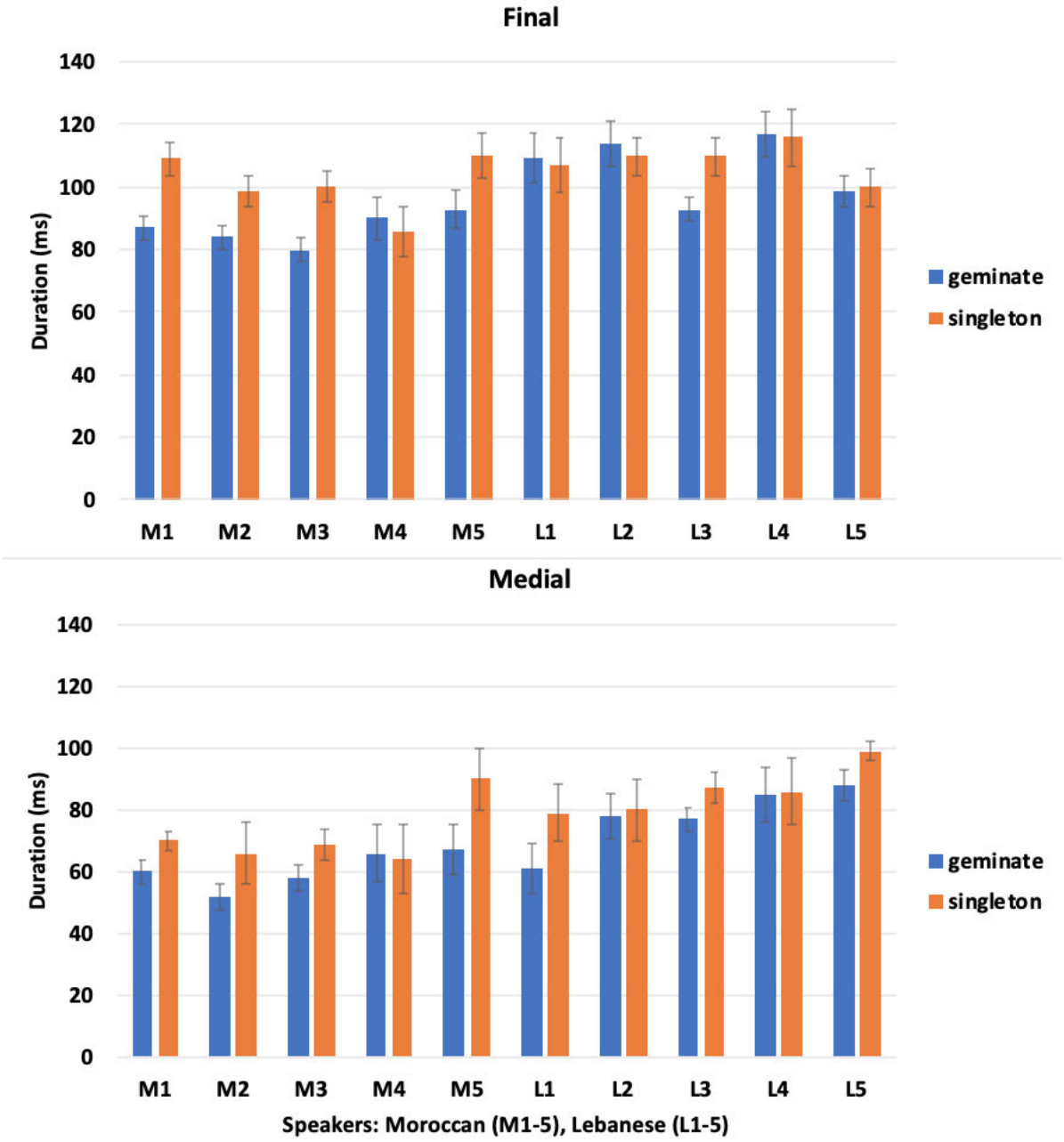

Figure 7.8: Mean preceding vowel duration values for each of the geminate and singleton targets displayed by group and by speaker (L1-5: Lebanese; M1-5: Moroccan) in word-medial (upper panel) and word-final (lower panel) positions and standard error of the mean.

7.4.2.6. Amplitude of the preceding vowel

Previous studies have also found that the vowel amplitude on either side of the consonant can be influenced by the distinction between geminates and singletons in word-

medial position (Finnish: Doty et al., 2007; Japanese: Idemaru, 2005). For our data, vowel amplitude was taken and averaged over the preceding vowel. The results of the mean vowel amplitude are presented in Figure 7.9:

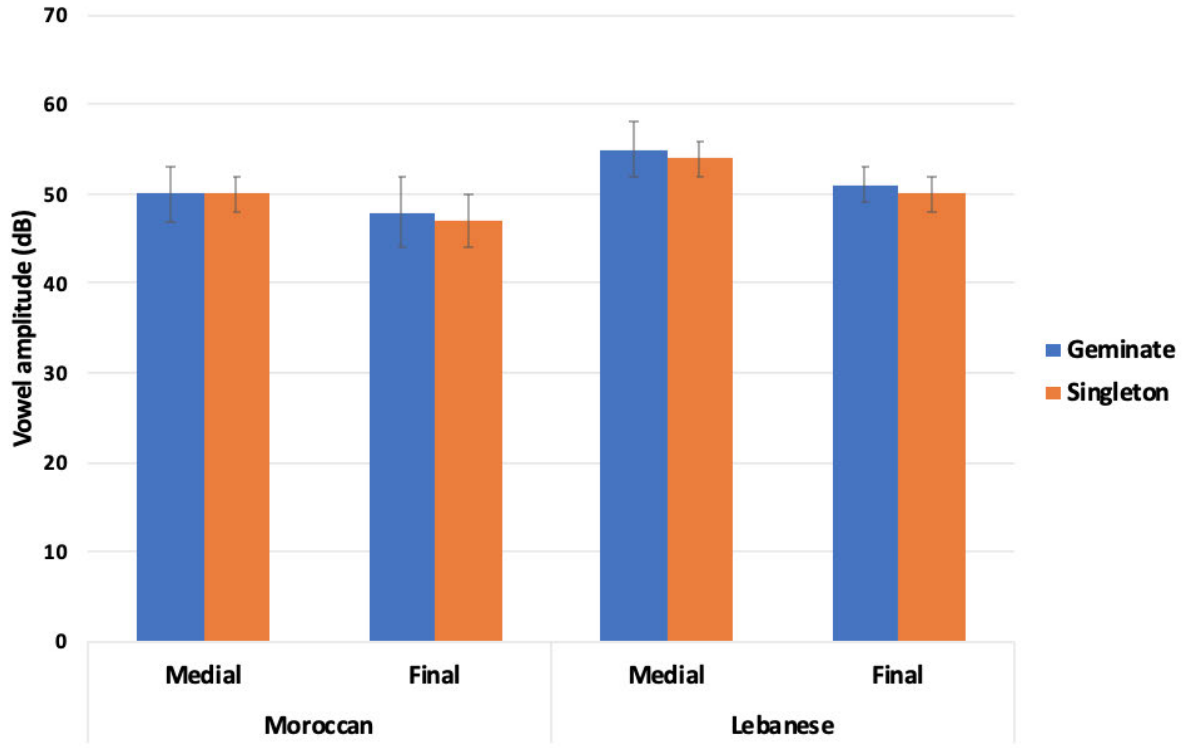

Figure 7.9: Mean vowel amplitude (dB) and standard error of the mean for geminate and singleton in both word positions and for both dialects

The results of preceding vowel amplitude values were submitted to LME modelling analysis. Although the overall amplitude of the preceding vowel appears to be slightly stronger in the pre-geminate context (M = 51 dB) compared to the pre-singleton context (M = 50 dB), this difference is not statistically significant ($F(1, 320) = 1.25$ $p = .564$),. However, there was a significant main effect of group, $F(1, 320) = 68.73$, $p < .001$, indicating that Lebanese speakers produced the geminate/singleton preceding vowel offset with slightly higher amplitude (M = 52 dB) than the Moroccans (M = 48 dB). The main effect of voicing was also significant ($F(1, 290) = 73.4$ $p < .001$), i.e., voiced stops had higher burst amplitude (M = 56 dB) than voiceless stops (M = 44 dB). There was also a significant effect of position, $F(1, 312)$

= 102.56, $p < .01$, i.e., higher values of the preceding vowel amplitude were registered for the word-medial position (M = 52 dB) than for the word-final position (M = 49 dB).

7.4.2.7. Vowel Offset f0

Alongside amplitude, studies have also investigated the contribution of fundamental frequency of the surrounding vowels in signaling the geminate/singleton contrast. In some languages f0 of the preceding vowel offset is found to be higher in the geminate compared with the singleton context. This has been confirmed for Bengali (Ghosh, 2015), Japanese (Kawahara, 2015) and Lebanese (Al-Tamimi & Khattab, 2011) (only for fricatives in medial position in Lebanese). This suggests that f0 of the preceding vowel offset is potentially a strong non-durational cue to the geminate/singleton contrast. The results of the mean f0 values for our data based on the LME statistical analysis is reported in Table 7.10.

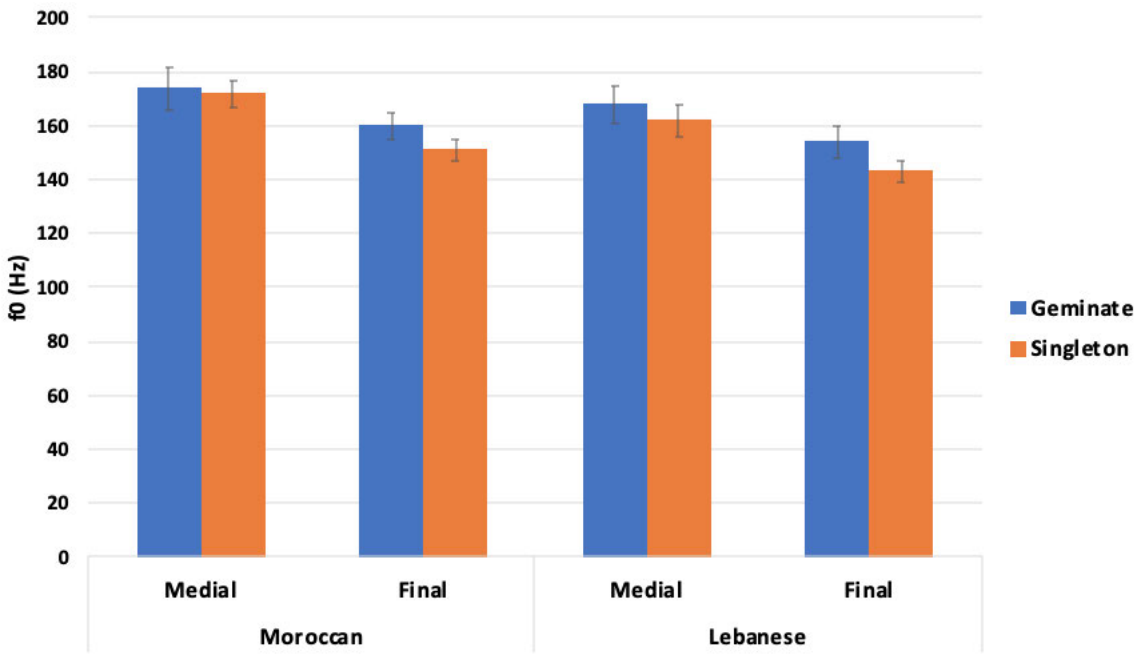

Figure 7.10: Mean vowel offset f0 (Hz) and standard error of the mean for geminates and singletons in both word positions and for both dialects

The LME model showed a significant main effect of length, $F(1, 320) = 10.4$, $p = .0097$ revealing that the mean f0 values were significantly higher at the vowel offset preceding a geminate than preceding a singleton (M = 164 Hz for geminates vs. M = 157 Hz for singletons). There was also a significant main effect of position, $F(1, 312) = 26.3$, $p < .001$), i.e., higher f0 values were registered in the medial position (M = 169 Hz) compared with word-final position (M = 152 Hz). Importantly, the interaction between length, position and group was significant, $F(1, 342) = 66.5$, $p < .001$). Breaking down this interaction by position revealed that there was a significantly higher f0 of the vowel offset for geminates (M = 157 Hz) than singletons (M = 147 Hz) only in word-final context (see Table 7.5). In the word-medial context, the effect of length on f0 approached significance ($p = 0.07$) only for Lebanese, but the difference between geminates and singletons was not significant for either group. There was a significant main effect of voicing, $F(1, 290) = 16.4$, $p < .01$), i.e., vowel offset fundamental frequency is higher for voiceless (M = 169 Hz) than voiced stops (M = 152 Hz). Furthermore, despite the f0 contrast between pre-geminate and pre-singleton vowel offset being more enhanced in final than in medial position, higher overall fundamental frequency values were produced word-medially (M = 169 Hz) than word-finally (M = 152 Hz), as revealed by a significant main effect of position, $F(1,312) = 99.4$, $p < .001$. In addition, the effect of group was significant, $F(1, 320) = 71.4$, $p <.001$, i.e., Moroccan speakers produced higher values for the f0 (M = 164 Hz) than Lebanese speakers (M = 156 Hz), along with a significant interaction between group and position, $F(1, 322) = 11.4$, $p = .0478$, i.e., the f0 contrast between a geminate and singleton is greater for Lebanese than Moroccan in final position (11 Hz vs 9 Hz), which suggests that this cue comes into play more strongly in Lebanese than in Moroccan.

7.5. Discussion

The current study was undertaken to investigate the articulatory and acoustic characteristics of word-medial and -final gemination in Modern Standard Arabic as spoken by Moroccan and Lebanese Arabic speakers. The acoustic investigation has provided a comprehensive overview of the different cues that might contribute to the geminate/singleton contrast. In line with findings from previous research on gemination mainly in word-medial position (Bengali: Lahiri and Hankamer, 1988; Berber: Ridouane, 2010; Bernese: Ham, 2013; Buginese: Cohn, Ham & Podesva, 1999; Estonian: Engstrand & Krull, 1994; Finnish: Engstrand & Krull, 1994; Cypriot Greek: Tserdanelis & Arvaniti, 2001; Guinaang Bontok: Aoyama & Reid, 2006; Hindi: Ohala, 2007; Hungarian: Ham, 2013; Italian: Payne, 2005; Pickett et al., 1999), it was hypothesised that closure duration would be strongly associated with the distinction between geminates and singletons. However, how this closure duration would be realised by speakers of two disparate regional dialects as well as the extent to which it would be maintained in word-final position were previously unanswered. Importantly, we investigated the theoretical issue of whether gemination reflects an articulatory goal of lengthening the closure or a more forceful articulation. These were the key questions that we investigated here.

The results for the closure duration (both acoustic and articulatory) clearly showed that it plays a robust role in distinguishing between geminates and singletons. This pattern was consistent not only in word-medial position but also in word-final position, the context in which gemination was previously thought by some researchers to be neutralised (Ghalib, 1984). The fact that both Moroccan and Lebanese speakers produced longer closure durations for geminates relative to their singleton cognates in this position is strong evidence for the contrastiveness of word-final gemination in MSA. Nonetheless, the lowest geminate-to-

singleton duration ratios (g/s) were witnessed in this position (*M* = 1.67 for Moroccan and 1.56 for Lebanese) compared to the duration ratios found in word-medial position (M = 2.25 for Moroccan and 1.85 for Lebanese). A smaller g/s duration ratio in final position not only suggests that the contrast may be perceptibly difficult to discern, but it might also imply that the contrast could eventually become subject to neutralisation in production in this position of the word, which is not uncommon among languages where word-final gemination was historically maintained and became neutralised over time. It is generally assumed that higher geminate-to-singleton duration ratios are found in the medial position of the word than in final position (Dmitrieva, 2012). Our results are in agreement with this assumption.

Despite being realised with a smaller geminate-to-singleton contrast compared to its word-medial counterpart, closure durations are substantially and significantly longer for geminates and singletons (even in the word-final position) in both regional varieties of Arabic, with a mean duration of 330 ms (geminates) versus 198 ms (singletons) for Moroccan speakers and of 277 ms (geminates) and 177ms (singletons) for Lebanese speakers. Both geminates and singletons are systematically lengthened by speakers of each variety relative to their word-medial cognates: 238 ms (geminates) versus 105 ms (singletons) for Moroccan and 218ms (geminates) versus 99 ms (singletons) for Lebanese speakers. This positional difference is likely to reflect domain-final lengthening, whereby the duration of segments occurring in utterance-final syllables show an increase in duration relative to the same segments in utterance-medial position. Previous acoustic analyses of stops produced in utterance-final syllables have reported that their closures are doubled (Hebrew: Berkovits, 1991) to tripled in length (French: Bell-Berti, Gelfer, Boyle & Chevrie-Muller, 1991), relative to the same stops in utterance-medial locations.

In terms of dialectal differences, Lebanese speakers almost exclusively produced shorter closure duration values than Moroccan speakers, with a mean group difference of 47 ms for geminates and 11 ms for singletons. Interestingly, Moroccan speakers not only produced longer closure durations, but their geminate-to-singleton duration ratios were also larger than those produced by Lebanese speakers. We speculate from these results that closure duration may play a slightly stronger role in contrasting geminates with singletons in Moroccan compared to Lebanese. Based on the smaller geminate-to-singleton ratio by the Lebanese group compared to the Moroccan group, we tentatively speculate that final gemination is more likely to be neutralised in Lebanese than in Moroccan.

With respect to the results for the duration of the vowel preceding the consonant, The results from our data showed that while the majority of our Moroccan speakers produced shorter durations of the vowel before geminates in the word-medial and even -final positions, this pattern was not significant for the Lebanese participants. All except one Moroccan speaker (but only one Lebanese speaker) produced shorter vowels before geminates than singletons. Thus, shortening of the vowel before geminates is a stronger cue to the geminate/singleton contrast in Moroccan than Lebanese, particularly in word-final position. This could possibly add to speculation about regional difference in susceptibility to eventual neutralisation of final contrast (more so for Lebanese than Moroccan), but this warrants further investigation. Shortening of vowel duration before geminates was a feature defining the geminate/singleton contrast in several languages: Bengali (Lahiri and Hankamer, 1988), Buginese (Cohn et al., 1999), Hindi (Ohala, 2007), Icelandic (Pind, 1995; 1999) and Italian (Payne, 2005). Although less common, the inverse pattern (lengthening of pre-geminate vowel) has also been reported in some other languages, such as Japanese (Idemaru & Guion, 2008). This inverse pattern in Japanese might be due to the language being mora-timed, which is not the case with the other languages mentioned above. The vowel-lengthening effect before geminates in Japanese is

argued to be due to the geminated (medial) consonant constituting a bimoraic foot in mora-timed languages. This would not apply to Arabic, which is not considered to be mora-timed.

Another possible explanation for why vowels may tend to be shortened before geminates in non-mora-timed languages is posited by Ridouane (2007). He argues that geminates require greater effort to be produced, and speakers anticipate this added effort by shortening the preceding vowel. Thus, preceding vowel shortening might reflect a more forceful articulation for geminates than singletons.

The other temporal cue that was investigated was the duration of the stop release burst. The results showed that the effect of gemination on the stop release duration was not consistent across consonant voicing categories: the difference in burst duration between geminates and singletons was significant for voiced stops only. The presence or absence of gemination was previously reported to affect the duration of the release burst in some languages, such as Bengali (Mikuteit & Reetz, 2007) and Tashlhiyt Berber (Ridouane, 2010), but in these languages it was also attested specifically for voiced stops. The only known exception is Cypriot Greek (Tserdanelis & Arvaniti, 2001), in which the phonemic inventory lacks voiced stops and the duration of the release burst in their voiceless stops was affected by the presence of gemination (longer for geminates). The fact that the stop release burst duration did not vary significantly across both voicing types suggests that this is not a consistent cue to the geminate/singleton contrast in Moroccan and Lebanese.

The acoustic analysis investigated three non-temporal cues (release burst amplitude, preceding vowel offset amplitude, and f0) in order to explore the extent to which they may contribute to contrasting geminates with singletons. To begin with, the release burst amplitude was significantly higher for geminates than singletons only in word-final position. In word-medial position, there was a tendency for the burst amplitude to be higher in the geminate

environment, but the difference was not statistically significant. This suggests that release amplitude plays a more important role than closure duration in distinguishing geminates from singletons in word-final than in word-medial position. It should be noted that Lebanese produced higher values of burst amplitude than Moroccans. In addition, the contrast between amplitude values for geminates and singletons was proportionally higher in Lebanese than Moroccan productions, suggesting that burst amplitude contributes more strongly to geminate/singleton contrasts in Lebanese than in Moroccan. Interestingly, this runs counter to the regional difference in importance of closure duration for gemination contrasts. In addition, the f0 of the preceding vowel offset behaved in a similar fashion to the release burst amplitude. That is, the importance of this cue was particularly witnessed in the word-final position for both dialects as indicated by significantly higher f0 of the preceding vowel offset in the context of geminates as compared to singletons. In word-medial position, despite a tendency for geminates to increase the f0 in the preceding vowel offset, this effect approached significance only for Lebanese (p = 0.07) but not for Moroccan. Although Moroccan speakers produced slightly higher mean values for f0 of the preceding vowel offset than Lebanese speakers overall, the geminate-to-singleton contrast for this parameter was slightly but significantly larger for Lebanese, which might suggest it plays a bigger role in Lebanese than Moroccan realisations of the gemination contrast. As for preceding vowel amplitude, results showed that this cue did vary as a result of the geminate/singleton contrast, despite a slight advantage towards stronger amplitudes in the geminate context.

Taking the acoustic findings as a whole into consideration, it appears that temporal versus non-temporal cues show a different balance in the two dialects of Arabic for contrasting geminates with singletons. While the effect of both acoustic and articulatory closure duration is significant across both dialects and positions, this cue is weighted more strongly in word-medial than -final position and is a stronger cue in Moroccan than in Lebanese. Similarly, the

duration of the vowel preceding the consonant plays a more prominent role in contrasting geminates with singletons in Moroccan than in Lebanese, although there were some individual differences. On the other hand, with the exception of the preceding vowel offset amplitude, the non-temporal acoustic measures appear to play a more prominent role in Lebanese than in Moroccan. The amplitude of the release burst and the fundamental frequency of the preceding vowel offset appeared to be more robust cues in Lebanese, particularly in word-final position. This means that for the most part the non-closure duration cues played a more important role in word-final than word-medial gemination, and particularly for the Lebanese MSA speakers. More importantly, while Moroccan appears to rely more on temporal cues, Lebanese seems to rely more on non-temporal cues, such as release burst amplitude and the fundamental frequency of the preceding vowel offset. Thus, temporal and non-temporal cues to gemination are weighted differently from one dialect to the other as we predicted. In future research we will examine other non-temporal cues, such as formant frequencies of the preceding vowel, in order to investigate their contribution to the geminate/singleton contrast in both dialects.

The presence of acoustic cues that are associated with a forceful articulation (other than closure duration), particularly in word-final position, might be taken as evidence that geminates are not based merely on a distinction of length of closure. Rather, the contrast is enhanced with a more forceful articulation when a large contrast of closure duration is difficult to achieve, i.e., word-finally. Converging evidence that geminates are characterised by a more forceful articulation compared with their singleton counterparts comes from the lack of release burst only for some singleton tokens. Although lack of a release burst affected only a small percentage of tokens (about 30 percent of voiced tokens and about 10 percent of voiceless tokens in Lebanese and 25 percent of voiced tokens in Moroccan), as noted earlier, this involved only singleton tokens, whereas all geminate tokens were produced by all speakers of both groups with a burst irrespective of their voicing type. An unreleased stop is regarded as

weaker version of its released counterpart (Fougeron, Kuehnert, Imperio, & Vallee, 2010). This further supports the account that the geminate/singleton contrast is based on articulatory force differences, with singletons being less forcefully articulated than geminates. The Enhancement Theory (Stevens, Keyser & Kawasaki, 1986) could offer an explanation for the presence and enhancement of non-closure duration attributes in word-final position. According to this theory, the primary distinctive feature (i.e., closure duration for gemination) is often supplemented by other attributes, which serve as enhancing features, particularly when the primary feature is obscured at the periphery of the word. Moraic theory offers another explanation for the forceful articulation associated with geminates (e.g., Davis, 2011; Topintzi & Davis, 2017). According to this theory, geminates are underlyingly bi-moraic, while singletons instead are mono-moraic. In phonology, the mora is the unit representing syllable weight. The moraic theory posits that the prosodic tier is characterised by weight rather than length, i.e., geminates have a single segmental node which makes the syllable they are part of "heavy" because it combines two morae.

On the subject of domain-final strengthening, it was hypothesised that the geminate/singleton contrast would be strengthened word-finally as it aligns with the final position of the prosodic domain. The results partially supported our hypothesis. In terms of closure duration, word-final stops were substantially longer in word-final than word-medial positions. However, we noticed that the geminate-to-singleton contrast ratio was slightly larger in the medial position compared to the final position. This result is not consistent with the strengthening in the domain-final position. Nonetheless, this result would be less surprising if we take into account that the role of closure duration in signalling gemination is mostly attested in the word-medial position. That is, medial gemination is primarily based on differences in closure duration. In the word-final position, we have noticed that other non-closure cues, such as preceding vowel duration and amplitude of the burst, played a stronger role in signaling the

geminate-singleton contrast. Thus, it is possible domain-final strengthening is achieved through various parameters that are language-specific.

The articulatory signals generated by TRACTUS from the ultrasound data analysis mirrors the findings from acoustic analysis in that geminates in these both Moroccan and Lebanese are produced with longer articulatory closure than singletons, which is in agreement with most articulatory investigations on gemination. The longer closure for geminates relative to singletons was maintained across both stop consonant voicing types and more importantly, in both word positions. This finding was statistically confirmed for all speakers of both dialects. The ultrasound analysis was particularly useful for data from the word-final position where there were tokens whose closure duration could not be measured from the acoustic data because their closure was not followed by a measurable burst. The TRACTUS analysis reliably generated an articulatory signal for these tokens, allowing for closure duration to be estimated directly from the articulatory signal, even for the "burstless" tokens. This further proves that this TRACTUS and the articulatory closure duration function we created offer an efficient and robust method for quantifying duration differences for contrasts whose temporal changes are difficult or impossible to discern from the acoustic signal alone.

One limitation worth noting is the limited number of words used in this production study. The aim of this PhD project was to investigate how the native speakers of Morocco and Lebanese produce initial (Chapter 6), medial and final gemination. Knowing the acoustic and articulatory similarities and differences between the two regional varieties is also crucial to interpreting the results of the subsequent perception experiments (Chapter 8). With these aims in mind, the list of words had to be carefully selected and given the linguistic differences among the two varieties of Modern Standard Arabic (MSA) and the two vernacular dialects, the task of finding words that met our stimulus requirements and were also meaningful words across

all varieties was not an easy one. This limitation was mitigated by increasing the number of repetitions (10 repetitions in total) and recording nonce words that are also comparable to the lexical items. Moreover, the speakers were asked to produce the nonce words in an MSA style of pronunciation Currently, an additional 5 minimal pairs are being examined. Acoustic data are being collected from the same original speakers. The data will be analysed and included before submitting the chapters for publication.

7.6. Conclusion

In conclusion, the results reveal that closure duration (both acoustic and articulatory) is a robust cue in contrasting geminates with singletons in both dialects and across both positions, especially in word-medial position. In word-final position, the novel insights from the present study are that the contrast between geminates and singletons based on closure duration is less strong and the presence of cues other than closure duration is particularly evident. These results might suggest that gemination is not only the result of a lengthening process, but that a more forceful articulation plays an enhancing effect for the contrast particularly when it occurs in word-final position. Domain-final strengthening is thus realised through a more forceful articulation.

CHAPTER 8: NON-NATIVE DISCRIMINATION AND CATEGORISATION OF MEDIAL AND PERIPHERAL GEMINATION IN REGIONAL ARABIC

8.1. Abstract

Research has shown that non-native geminate/singleton contrasts are easier to perceive word-medially than at either periphery (Dmitrieva, 2012). Several accounts have been proposed for this positional effect, but the answer remains unclear. Guided by models of non-native speech perception, particularly the Perceptual Assimilation Model (PAM: Best, 1995), we investigated how variations in native language experience affect discrimination and categorisation of non-native gemination contrasts across word positions. Native Lebanese and Heritage speakers of Lebanese dialect, and naive English listeners, were tested on perception of voiced and voiceless coronal stop gemination contrasts in the Moroccan variety of Standard Arabic and in Moroccan dialect. These listener groups differed in native experience with gemination in each word position. Sensitivity to closure duration differences in one word position did not predict performance in the other positions. Medially, the voiceless contrast was discriminated more accurately than the voiced contrast, but initial and final contrasts showed the opposite pattern, particularly for the Heritage Lebanese and English groups. Even listeners lacking native gemination experience (English) discriminated the medial contrast well, but they had substantial difficulty with peripheral gemination, especially the initial voiceless contrast. Importantly, these results were largely modulated by the amount of perceived phonological overlap in categorisations. Larger overlaps were found at the periphery of the word, which offers another possible explanation for lower discriminability in initial and final position. These results highlight the importance of considering phonological effects as well as contextual variations in sensitivity to phonetic details to improve the predictive power of non-native speech perception theories.

8.2. Introduction

A speaker's native language shapes their perception of non-native speech contrasts. Adults discriminate phonemes employed in their native language with relative ease, but find it difficult to discriminate and categorize many phonemes that do not correspond directly to native phonemic categories or contrasts (e.g., Lisker and Abramson, 1970; MacKain et al., 1980; Miyawaki et al., 1975; Strange, Akahane-Yamada, Kubo, Trent & Nishi, 2001; Trehub, 1976). There is a general agreement that listeners with no prior exposure to another language find it challenging to discriminate between contrasting phonemes in that language which are not contrastive in their native language (L1). This perceptual difficulty has been reported in many studies. For example, native speakers of Japanese, which lacks an /r/-/l/ contrast, find English /r/ and /l/ difficult to discriminate (Goto, 1971). However, not all non-native contrasts fail to be discriminated by listeners. Non-native contrasts may be discriminated with greater or lesser difficulty depending on how their phonetic and phonological properties compare to those of the perceiver's native phonemic categories. Non-native speech perception thus relies on both the abstract phonological features and finer-grained phonetic details of native phonemes as listeners assess the relative fit of the non-native phones to these two levels of information. One way to discern the relative contributions of the two levels of information is through investigating perception of non-native phonological contrasts that behave differently at the phonetic level across different word positions. One such contrast for which there has been little research on either native language or non-native positional effects in perception, is the geminate/singleton contrast (gemination).

Gemination indicates a phonological contrast between geminate and singleton consonants, i.e., it can differentiate words, as it does in Arabic and many of its regional varieties. The phonetic realisation of the contrast is the production of geminates with an

elongated duration relative to their singleton counterparts. Purely phonetic gemination (i.e., consonant lengthening that does not result in lexical change) can also occur in languages that lack singleton-geminate phonological contrasts. Few studies have investigated the perception of gemination, and none has examined the effect of phonetic and phonological factors on non-native perception and how they are modulated by the position of the contrast in the word in light of theoretical models of speech perception in listeners varying in native experience with gemination in different positions. The present study investigates how three groups of listeners (Native Lebanese, Heritage Lebanese and native English) discriminate and categorise the geminate/singleton contrast in two languages/varieties that are non-native to all three groups. Specifically, gemination contrasts occur both word-initially and -medially in Moroccan dialect[15], which is a different language from Lebanese dialect. They also occur word-medially and -finally in the Moroccan variety of Modern Standard Arabic (MSA), which differs from the Lebanese variety of MSA. Although Moroccan and Lebanese are considered to be two dialects of the same language, the stimuli produced by native Moroccan speakers would be considered non-native for both native and heritage Lebanese listeners, especially given that the two Arabic dialects are mutually unintelligible. This use of "non-native" concept is supported by works investigating inter-dialectal speech perception. Non-Australian English regional accent differences have been considered non-native for Australian English listeners in published research (Shaw et. al., 2018).

Several models have been proposed to provide an account for the universal mechanisms that underlie non-native speech perception, through exploration of the phonetic and phonological differences and similarities between the target language and a listener's L1. The

[15] Moroccan Arabic vernacular contains a very limited number of words with final geminated consonants, mostly borrowed from MSA. These geminates do not appear in minimal contrasts with singletons in final position. Thus, Moroccan dialect, like Lebanese dialect, lacks final gemination *contrasts*.

most widely cited and used of these frameworks are the Perceptual Assimilation Model (PAM: Best, 1995; Best & Tyler, 2007), the Speech Learning Model (SLM: Flege, 1995), the Native Language Magnet model (NLM: Kuhl & Iverson, 1995) and the Second Language Linguistic Perception model (L2LP: Escudero, 2005). These frameworks share one core assumption, namely that they all recognize a central influence of the listener's L1 on the perception of non-native speech segments. However, only PAM addresses the importance of systematically considering the relationship between native phonetic and phonological factors in non-native speech perception.

According to PAM, non-native speech perception operates at an abstract phonological level as well as a lower phonetic level. PAM posits that listeners tend to make assimilations on the basis of native phonological categories, but they clearly retain some sensitivity to within category phonetic variations. This sensitivity is determined by the phonetic details of their native language and the perceived fit of the phonetic details of the non-native item. Thus, PAM assumes that non-native speech perception is subject to L1 phonological influences as well as phonetic influences. The magnitude of the phonological and phonetic influences can be determined by how a non-native phone is assimilated to a native category. When a non-native phone is perceived as clearly similar to one L1 phonemic category, it is considered to be *categorised***.** In this case the phonological influence is strong and there appears to be a tendency for it to override sensitivity to lower-level phonetic details. However, when speakers perceive a non-native phone as speech, but they do not perceive it as clearly similar to any single native phoneme, it is considered to be *uncategorised*. Phonological influence is expected to be weaker for uncategorised assimilations. There are three different ways that a non-native phone could be *uncategorised* (see Faris, Best & Tyler, 2016, 2018): 1) a focalized assimilation, in which a non-native phone is perceived as similar to a single L1 category but choices of that native phoneme are below a predefined categorisation threshold; (2) a clustered assimilation, in which

the non-native phone is assimilated below threshold to a small set of L1 categories; or (3) a *dispersed* assimilation in which the choice of native phone category is spread across many L1 categories, but no choice is above chance level. These differences in types of assimilations reflect differences in perceived phonological overlap.

PAM further explains how non-native speech perception is constrained through the listener's native language phonology. Non-native categories that are completely outside the learner's L1 phonology are perceptually assimilated as non-speech sounds (e.g., Zulu click consonants are perceived by some English speakers as fingers clicking or twigs snapping, Best et al., 1988), while non-native contrasts that map equally well or poorly onto a single segment in the L1 pose the greatest difficulty and the discrimination gets more difficult as the learner's L1 fluency progresses from infancy to adulthood. According to PAM, the easiest non-native distinctions to perceive are those that map onto two contrasting phonemes in the learner's L1. Thus, when a listener attempts to categorise and discriminate a non-native contrast, different possible patterns of perceptual assimilation are expected:

1) Two-category (TC). The contrasting non-native phonemes are each assimilated to different native phonological categories. Discrimination is predicted to be excellent because the non-native phones are perceived as instances of native categories that contrast with each other in their native phonological system.
2) Category-goodness (CG). The contrasting non-native phonemes are assimilated to the same native phonological category but differ in that one phoneme is perceived as a good exemplar while the other is perceived as a poor exemplar. Discrimination is predicted to be moderate to good. This reduced level of discrimination, relative to TC assimilations, is due to the absence of a clear native phonological difference between the contrasting non-native phones. So, the listener relies on sensitivity to

phonetic goodness of fit to the native category to discriminate the non-native phonetic distinction.

3) Single-category (SC). Both non-native phonemes are assimilated to the same native phonological category and are perceived as equally good or poor exemplars of the native phoneme. Discrimination is predicted to be poor (difficult) because listeners are unable to detect phonetic differences between the two categories.
4) Uncategorised-Categorised (UC). One non-native phoneme is assimilated to a native phonological category and the other is uncategorised. Discrimination is expected to be very good, although it may vary depending on the perceived phonological overlap between the uncategorised and categorised phone (Best, Avesani, Vayra & Tyler, 2019; Faris et al., 2018).
5) Uncategorised-Uncategorised (UU). Both non-native segments are uncategorised. Discrimination is predicted to be poor to very good, depending on perceived phonological overlap and/or the phonetic distance between the two phonemes.
6) Non-assimilable (NA). Both contrasting non-native phones fall outside the native phonological space and heard as non-speech sounds. Discrimination is predicted to be good to very good, depending on the perceived auditory differences.

PAM generally predicts that TC contrast assimilations are the easiest to discriminate, followed by both CG and UC, while SC assimilations are expected to be the most difficult to discriminate. UU assimilations cannot be ranked easily because their discrimination accuracy is contingent on the similarity between the contrasting phones and/or the L1 phonemes. Recent works have proposed more accurate predictions for UC and UU based on an approach of identifying the presence of perceived phonological overlap (Best, Avesani, Tyler & Vayra, 2019; Faris, Best & Tyler, 2016, 2018). Accordingly, excellent discrimination is expected for non-overlapping contrasts, followed by partially overlapping, and then the poorest

discrimination is expected for completely overlapping contrasts. In line with this argument, we hypothesize in this study that the amount of phonological overlap in categorisations should be a good predictor of the success listeners will have in discriminating the non-native geminate/singleton contrast.

Similarities and differences between native and non-native phonetic details also shape non-native speech perception. Native phonetic influence on perception is most apparent when investigating a non-native contrast that exists in the native language but is not in a familiar context. For example, although /tl/ and /dl/ are permissible consonant combinations in French, they are not phonotactically allowed to occur in the initial position of the word. This results in French listeners perceiving the initial /tl/ and /dl/ clusters of Hebrew as the legal French onset clusters /kl/ and /gl/ (Hallé et al., 1998; Hallé & Best, 2007). The authors' interpretation was that /dl, tl/ underwent a phonotactic perceptual assimilation to the most phonetically similar permissible clusters through reliance on native phonetic details. Thus, native phonotactic constraints appear to hinder the perception of phonotactically impermissible non-native contrasts. However, this result is inconsistent across languages. Although Dutch disallows voiced obstruents in the final position of the word, Dutch listeners' performance in discriminating final voiced versus voiceless obstruents in nonwords exceeded that of native English listeners (Broersma, 2005). Similar to Hallé and Best (2007), the author suggested that Dutch listeners rely on native phonetic cues to aid perception of the non-native contrast. These phonetic cues are language-specific and differ from those that native listeners of the target contrasts employ. Importantly, the inconsistency between the findings of the two studies makes it unclear whether familiarity with a contrast facilitates or hinders discrimination of the contrast in an unfamiliar phonotactic context. Thus, more research is warranted into native phonetic influences on non-native speech perception as modulated by positional effects.

PAM-based research proposed that native speech perception entails identifying phonetic differences that signal a lexical distinction (phonological distinctiveness), while being able to recognise a given lexical item despite irrelevant within-category phonetic details that do not signal a change in the word (phonological constancy) (Best, Tyler, Gooding, Orlando & Quann, 2009; Best, 2015). Developing phonological constancy contributes to detecting phonological contrasts with relative ease and efficiency. There is evidence that phonological constancy is modulated by the functional load of a contrast in the native language of the listener. Native Russian listeners had difficulty discriminating between consonant + palatal glide (Cʲj) and consonant + high vowel + palatal glide (Cʲij) (Diehm, 1998). The author explained this result by the fact this contrast carries a low functional load in Russian. Similarly, a recent set of experiments showed that English contrasts with a high functional load (e.g., /s/-/ʃ/) were more perceptually distinct than those with a low functional load (e.g., (/z/-/ʒ/) (Lin, 2019). Research on gemination concerned only the effect of functional load on the production of gemination. Certain phonetic aspects, such duration ratio of the geminate/singleton contrast, were shown to be positively corelated with the varying functional load of the contrast (Harris, 2014). However, to the best of our knowledge there are no other published studies investigating the effect of functional load on the perception of gemination.

Production of coronal stop gemination across different word positions has been examined in native Moroccan versus Lebanese Arabic speakers (Frej, Carignan & Best, 2017; Frej, Carignan, Proctor & Best, 2018; see Chapter 6 and 7). In Modern Standard Arabic (MSA), as produced by Moroccan and by Lebanese speakers, both geminates and singletons are longer in word-final than in word-medial position. However, the length contrast between a geminate and a singleton is more pronounced word-medially, as evidenced by a significantly higher geminate-to-singleton closure duration ratio (Frej, Carignan & Best, 2017). Similarly, word-initial gemination in Moroccan and Lebanese dialects was realised with a lower geminate-to-

singleton ratio despite both singletons and geminates being produced with substantially longer closure durations than in word-medial position. Clearly, closure duration plays a stronger role in cuing gemination in the medial position of the word. However, at the periphery of the word, other non-closure duration acoustic cues, both temporal and non-temporal, were enhanced. Also, because there was a tendency for Moroccans to rely more on temporal cues than non-temporal cues and vice versa for Lebanese speakers, we tentatively suggest that temporal and non-temporal cues to gemination are weighted differently from one dialect to the other.

Given these positional differences in production, word position should correspondingly affect perception of the geminate/singleton contrast. Word-medial consonant gemination is regarded to be easiest to discriminate and categorise, particularly when the target consonants are intervocalic (i.e., flanked by vowels). Studies have shown that even speakers of languages with no geminate-singleton contrasts can detect durational differences in consonant closure for non-native gemination contrasts that occur intervocalically (e.g., Huggins, 1972; Pajak, 2009; Pickett & Decker, 1960). However, detecting duration differences at the absolute periphery of the word is challenging. In an identification task, native speakers successfully identified a geminate only slightly above chance (54%) in absolute phrase-initial position in the Swiss German dialect of Thurgovian (Kraehenmann, 2001). Compatible with these findings, word-initial gemination in Berber was more challenging for French listeners to discriminate than final gemination was (Hallé, Ridouane & Best, 2016). Conversely, in a recent study comparing American English, Russian and Italian[16] listeners' categorisation of non-native geminate/singleton distinctions across different word positions (as produced by a native Russian speaker), word-final contrast was categorised the poorest relative to word-initial and especially intervocalic gemination contrasts (Dmitrieva, 2018). Nonetheless, there seems to be

[16] Gemination is phonologically contrastive only in Russian and Italian, but not in American English.

an agreement among scholars that word-medial (i.e., intervocalic) gemination is the easiest to discriminate whether native or non-native to the listener.

Several explanations have been offered for the greater challenges of perceiving/categorising peripheral gemination as compared to intervocalic gemination. First of all, peripheral gemination is prone to neutralisation. The peripheral positions of a word that are also at the periphery of an utterance can render the closure inaudible, particularly for utterance-initial voiceless stops, due to the lack of vocal fold vibration to provide acoustic information about temporal duration of the closure. Similarly, word-final geminated stops can be produced without a release, which makes the end of the closure auditorily indiscernible. This obscures the closure duration acoustically. However, these weak acoustic cues to gemination at the periphery of the word may apply more to certain segments (e.g., voiceless stops) but not others (e.g., voiced stops). Nevertheless, although these missing acoustic cues to closure duration are likely to make the contrast less discriminable, this does not seem to prevent native listeners' perception of peripheral contrast. For example, Pattani Malay (Abramson, 1991; 1999) and Kelantan Malay (Hamzah, 2013) listeners were clearly successful at categorising the geminate/singleton contrast in absolute initial position, even for voiceless stops in their native languages. Therefore, due to the reduced acoustic cues to closure duration differences at the periphery of the word, native listeners seemingly attend to other differences between a geminate and a singleton. The second explanation that was put forward for the difficulty with discriminating and categorising peripheral contrasts is that word-peripheral gemination can yield a lower geminate-to-singleton closure duration ratio than medial gemination (Pajak, 2009; Frej, Carignan, Proctor & Best, 2018; Frej, Carignan & Best, 2017). The higher the ratio, the easier it is to discriminate the contrast (Dmitrieva, 2012). However, the increase in duration ratio in the medial position is not consistent across all languages. In Tashlhiyt Berber, the duration ratio is instead higher in word-initial and word-final than in the intervocalic position:

intervocalic < initial < final (for fricatives) and intervocalic < final (for stops, which were not measured in word-initial position) (Ridouane, 2007). Hence, the closure duration ratio criterion cannot be used as a generalised explanation for the reduced discriminability of peripheral gemination. Thirdly, some researchers attribute ease of perception of word-medial gemination to the contrastive effects of the surrounding environment of the geminate/singleton contrast. To explain, the flanking vowels offer a good contrast for intervocalic gemination to be perceived (Padgett, 2003). According to this account, the geminate /k:/ in *takka* /tak:a/ "fireplace" is easily differentiated from its singleton counterpart /k/ in *taka* /taka/ "back" in Finnish, because the preceding and following short vowel /a/ highlight the onset and offset of the medial /k:/~/k/. However, these studies fail to draw links to existing cross-language perception theories to offer a more principled, broader explanation for why the intervocalic environment renders intervocalic gemination more discriminable. We hypothesise that the differential discriminability of gemination across different contexts can be explained more systematically by PAM's principles.

Prior research on PAM has focused primarily on the discrimination and categorisation of segmental properties of phonemes, i.e., vowels (e.g., Tyler et al., 2014) and consonants (Best et al., 2001). However little to no attention has been given to how suprasegmental properties of non-native consonants, e.g., consonantal length, are perceived. The only study that has investigated non-native perception of gemination in light of PAM principles, examined only the discrimination of Berber gemination by French native listeners (Hallé, Ridouane & Best, 2016), without looking into how the contrasts are categorised. It is necessary to investigate both categorisation and discrimination in order to fully utilise the predictive power of PAM. Thus, this study builds on findings from previous research to offer new insights into non-native discrimination and categorisation of consonantal length across different contexts, by three groups of listeners with varying experience with consonantal length. Two perceptual tasks are

commonly used to test PAM predictions: categorisation and AXB discrimination tasks. The AXB discrimination task is used to determine potential difficulties with non-native contrasts. The task does not impose heavy memory load and by using different speakers and tokens for the A, X and B elements, the participants are prevented from responding solely on the basis of low-level auditory-acoustic differences and similarities (Hallé, Ridouane & Best, 2016). In conjunction with an AXB task, a categorisation task is often employed. The categorisation task is usually a simple forced choice procedure. The participants are asked to categorise each token presented separately (all tokens used in the AXB discrimination task), judging whether it is a long or a short consonant, and then to rate the goodness-of-fit of the token to the chosen category. The categorisation task allows the interpretation of the discrimination performance. For this reason, an AXB paradigm along with a categorisation task are employed in order to test the discrimination and categorisation of gemination in this study.

In order to address the unresolved theoretical issues discussed above, namely determining the relative contributions of abstract phonological features and finer-grained phonetic details of native phonemes on the perception of non-native gemination, we needed three non-native listener groups that varied according to their phonological and/or phonetic-only experience with consonantal gemination contrasts in the three critical word positions (medial, initial and final). Accordingly, we chose Arabic-naïve native English listeners (henceforth, English listeners), Moroccan-naïve native Lebanese listeners who spoke both Lebanese vernacular dialect and the Lebanese variety of standard Arabic (MSA) (henceforth, Lebanese listeners) and Lebanese heritage listeners who learned Lebanese dialect as their home language in Australia but had not learned standard Arabic (henceforth, Heritage Lebanese listeners). We investigated how these groups discriminated voiced and voiceless coronal stop[17]

[17] Only coronal consonants that can be geminated word-initially in Moroccan and Lebanese.

gemination contrasts in word-medial and -final positions in the Moroccan variety of MSA and word-initial position in Moroccan dialect.

The choice of these listener groups was motivated by their differential exposure and familiarity with gemination contrasts at different word positions in their native languages. English speakers are not familiar with phonologically contrastive gemination in any word position but do have purely phonetic-level experience with so-called "fake gemination", which can occur when the same phoneme occurs on either side of a word boundary (e.g., <top pick> vs. <topic>). The geminated segment is phonetically long, but it is spread across two syllables falling on either side of a word boundary (Hayes, 2001; Pickett & Decker, 1960). English speakers lack even phonetic-level experience with initial and final gemination. Lebanese speakers have native phonological experience with initial and medial gemination contrasts in their vernacular dialect, and with final gemination contrasts in MSA, in which they carry a relatively high functional load. Our Heritage Lebanese speakers raised in Australia have native experience with initial and medial gemination in the Lebanese dialect that they learned at home as their first language, but no experience with final gemination contrasts as they have not learned MSA.

We predict that the three groups' success with discriminating the medial, initial and final gemination contrasts in the Moroccan target stimuli will reflect their varying experience with the contrast in each word position. Phonological-level experience with gemination contrasts in a listener group's native language(s) (the two Lebanese groups) should yield better discrimination than experience only with phonetic-level "fake" gemination across word boundaries (English group). Phonological experience with medial gemination should also be more likely to support generalisation to non-experienced word-final gemination distinctions (Heritage Lebanese group) than does only phonetic-level experience (English listeners).

Moreover, the stimuli are non-native for all three groups of listeners, not only for English listeners but also for the two Lebanese groups, as they are not native speakers of the Moroccan vernacular dialect and the Moroccan variety of MSA. Testing Lebanese listeners' discrimination and categorisation of these Moroccan stimuli allows us to test hypotheses concerning the effects of native language experience on perception of a different, geographically distant vernacular dialect (a non-native language for both Lebanese groups) and of experience with a different regional variety of MSA (a non-native regional accent of standard Arabic for the Lebanese group).

8.3. Method

8.3.1. Participants

We recruited three groups of listeners consisting of 24 Arabic-naïve native English listeners (M_{Age} = 22.5 years; 12 females, 12 males), 24 Moroccan-naïve native Lebanese listeners (M_{Age} = 31.2 years; 14 males, 10 females) and 24 Australian Lebanese heritage listeners (M_{Age} = 23.8 years; 13 males, 11 females). The three groups were recruited from among students at Western Sydney University and from the general community of greater Sydney, Australia. The English and Heritage Lebanese participants were all born and raised in Australia, whereas the native Lebanese participants were all born in Lebanon and arrived in Australia within 4 years of the test session date. Heritage Lebanese listeners were required to have had no formal instruction in MSA, and they all reported themselves to be fluent or near-fluent speakers of the Lebanese dialect, which was their native language (home language from birth). None of the native or heritage Lebanese listeners had ever been to Morocco, and all reported little to no experience with the Moroccan vernacular dialect. An additional 11 participants (four English, three Lebanese and four Heritage Lebanese listeners) were tested

but their data were removed from the final dataset for not having successfully followed the task instructions.

8.3.2. Stimulus materials

The stimuli for this experiment were extracted from the productions of three male Moroccan speakers (see Chapter 6 and 7). Four nonce disyllabic minimal pair words contrasting coronal stops /t/ and /d/ with /t:/ and /d:/, respectively, in word-medial and -final position in Modern Standard Arabic (MSA) were used. In addition, since initial gemination is not permissible in MSA, two word-initial minimal pairs contrasting geminate with singleton were extracted from Moroccan vernacular dialect recordings. The stimulus items are presented in Table 8.1. Only the isolated word productions were chosen for the perception experiment to avoid coarticulation effects from surrounding words that would have been present had they been extracted from a carrier sentence.

Table 8.1: The voiced and voiceless singleton nonce stimulus word pairs[18]

	Voiced	**Voiceless**
Initial	/’**d**amah/-/’**d:**amah/	/’**t**amah/-/’**t:**amah/
Medial	/’a**d**am/-/’a**d:**am/	/’a**t**ab/-/’a**t:**ab/
Final	/qa’ma**d**/-/qa’ma**d:**/	/qa’ba**t**/-/qa’ba**t:**/

Nine tokens of each target word were used in the experiment (3 speakers x 3 tokens). All tokens were realised with a closure duration comparable to the mean closure duration for each contrast in the previous production experiment (Frej et al., 2017, 2019). Word-initially, geminates had longer closure durations than singletons for both voiced (*M* = 286 ms vs. 197

[18] It was not possible to use the same context in all three positions, because some of the words would be meaningful.

ms) and voiceless cases (*M* = 297 ms vs. 214 ms). Word-medially, geminates also had longer closure durations than singletons for both voiced (*M* = 220 ms vs. 95 ms) and voiceless cases (*M* = 254 ms vs. 115 ms). Word-finally, geminates had longer closure durations than singletons for both voiced (*M* = 325 ms vs. 193 ms) and voiceless cases (*M* = 335 ms vs. 201 ms). See Chapter 6 and 7 for detailed durational and non-durational acoustic results (see also Frej et al., 2017, 2019).

8.3.3. Procedure

All the participants were tested individually. Some were tested in a sound-attenuated testing booth at the MARCS Institute at Western Sydney University. Others were tested in quiet rooms at a local mosque. Stimulus presentation and response collection was controlled using Psyscope X B57 (Cohen, MacWhinney, Flatt & Provost, 1993) installed on a Macbook Pro computer. The stimuli were presented via Sennheiser HD 650 headphones.

A discrimination task for all three word positions was completed by the listeners followed by a categorisation task for all the word positions. Because the concept of gemination is unfamiliar to the English and heritage Lebanese listeners, they were introduced to it by an explanation that Arabic distinguishes between a long and short version of the same consonant and that this occurs word-initially, -medially and -finally. In a pilot experiment that had not included this brief explanation, Heritage and English listeners found the discrimination task extremely difficult and indicated they were frustrated because did not understand what they were supposed to be listening for. Examples of similar words used in this study were given to demonstrate the concept. All participants reported having understood the concept of gemination prior to starting the experiment.

For the categorisation task, these participants were told that one highlighted letter represented a short consonant, while a long consonant was represented by two highlighted letters. For each word-position discrimination and categorisation task, the participants were instructed to focus on durational differences in the consonants and on one specific word position.

8.3.3.1. Discrimination

On each AXB trial, participants were presented with three stimuli. Participants were instructed to decide whether the stimulus X matched stimulus A or B in consonant duration by pressing 1 for A or 3 for B using the keyboard. The three stimuli were separated by a 1-s inter-stimulus interval and 3.5-s inter-trial interval between trials. Every combination of A, X and B appeared four times based on the four possible arrangements of the AXB triplets (AAB, ABB, BBA and BAA). The participants completed six discrimination blocks (2 voicing types x 3 word-positions). There were 16 trials per block. Each block was designed to test the discrimination of the geminate/singleton contrast. The blocks were presented in a random order for each participant.

8.3.3.2. Categorisation

The categorisation task was a simple four-choice forced choice procedure. The participants were asked to categorize each token presented separately in random order as being a /t/ or /d/ or their geminated counterparts /t:/ or /d:/, respectively, by clicking with a mouse on buttons on the screen, each representing one target word with the target geminate/singleton consonant highlighted. Target words were presented both in Arabic script and Roman alphabet (see Figure 8.1), so that even Heritage Lebanese and English participants, who could not read the Arabic script, could participate in the experiment. Immediately after responding,

participants heard the same stimulus a second time and rated the goodness-of-fit to their chosen response category based on a 7-point scale via a keypress (1 = poor, 7 = perfect). The categorisation task was divided into three blocks, one for each word position and the blocks were counterbalanced across participants in each group.

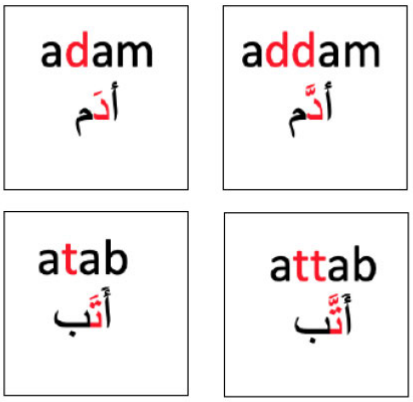

Figure 8.1: Word-medial categorisation task choice buttons as displayed on a screen.

8.4. Results

Throughout the results sections, the medial contrast results will be presented first, followed by final, then initial contrast. The medial results would be considered the baseline for comparison, since the three groups are familiar with gemination in the word-medial position. The discrimination task was run first, then followed by the categorisation task for listeners. However, we discuss the categorisation results first in order to determine the assimilation, which helps predict the discrimination accuracy.

8.4.1. Categorisation results

Table 8.2 shows the mean percentage categorisation and mean goodness of fit ratings for each item, for each of the three groups. Selecting a pre-determined threshold has been

commonly practised in perception studies to determine whether a phone is deemed categorised or uncategorised to a native category. Accordingly, if the total number of selections for category is above the pre-set threshold, the phone is deemed categorised, or otherwise uncategorised. Depending on the purpose of the study, previous research has applied various thresholds, ranging from 50% (e.g., Faris, Best and Tyler, 2018), 70% (e.g., Tyler, Best, Faber & Levitt, 2014) to 90% (e.g., Harnsberger, 2001). In this study, we applied a 70% criterion as this is frequently used for consonant contrasts (e.g., Antoniou, 2010). If one label is selected 70% or more of all other responses for a given target word, it is considered *categorised*. Otherwise, it is considered *uncategorised*. To determine perceived overlapping category choices, we followed the method adopted by Faris and colleagues (Faris, Best and Tyler, 2018). Accordingly, a contrast was deemed non-overlapping when the target stimuli were each categorised into separate response categories without any shared above chance response. A contrast was partially overlapping when the two non-native stimuli shared one above chance category, but there was also one or more above-chance category that was not shared. If the all of the above-chance categories were shared by both contrast members, then the contrast was considered to be completely overlapping.

Table 8.2: Mean percent categorisation and mean rating of goodness-of-fit of Moroccan Arabic voiced and voiceless geminate/singleton contrasts for Lebanese, Heritage Lebanese and English groups and across word-medial, final and initial positions. Means are averaged across listeners of each group. Numbers in bold represent categorised categories (70%). Numbers in italics represent significantly above chance response categories (25%). Only above chance responses are displayed. Numbers between brackets represent Goodness-of-fit ratings: 1(poor) to 7(excellent).

	group			Lebanese				Heritage Lebanese				English			
	Voicing			Voiced		Voiceless		Voiced		Voiceless		Voiced		Voiceless	
	Length			Gem	Sing	Gem	Sing	Gem	Sing	Gem	Sing	Gem	Sing	Gem	Sing
POSITION	Initial	Voiced	Gem	**73%** (4.3)				**71%** (4.5)				36% (3.5)	49% (2.8)		
			Sing		**89%** (5.3)				**73%** (4.9)				68% (4.6)		
		Voiceless	Gem			64% (3.6)	34% (3.1)			63% (4)	32% (2.9)			40% (2.6)	53% (2.9)
			Sing				**81%** (5.8)			30% (3.1)	**70%** (4.7)			30% (2.2)	65% (3.1)
	Medial	Voiced	Gem	**94%** (4.9)				**90%** (4.5)				**70%** (4.1)			
			Sing		**96%** (5.3)				**92%** (4.3)				**77%** (4.9)		
		Voiceless	Gem			**93%** (4.6)				**89%** (4.2)				**71%** (3.6)	
			Sing				**95%** (5.8)				**93%** (4.7)				**76%** (4.7)
	Final	Voiced	Gem	**88% (4.6)**				64% (3.7)				58% (5.6)	30% (3.3)		
			Sing		**91%** (5.8)				**79%** (4.1)				69% (4.7)		
		Voiceless	Gem			**81%** (4.9)				62% (4.3)	31% (3.1)			50**%** (4.5)	42% (3.7)
			Sing				**85%** (5.3)				**78%** (4.3)			31% (3.1)	65% (4.3)

Table 8.3: Contrast assimilation patterns (70% for categorisation and 25% for chance level) and their discrimination accuracy scores across medial, final and initial positions by the three groups of listeners (Lebanese, Heritage Lebanese and English). The first two letters (e.g., TC) refer to the assimilation type: TC: Two-Category, UC: Uncategorised-Categorised, UU: Uncategorised-Uncategorised, and the letter after the dash (e.g., N) refers to the degree of phonological overlap in assimilation: N: no overlap, P: partial overlap and C: complete overlap.

				Assimilation Type	Discrimination Accuracy scores
GROUP	**Lebanese**	**Initial**	**voiced**	TC-N	87%
			voiceless	UC-P	74%
		Medial	**voiced**	TC-N	94%
			voiceless	TC-N	96%
		Final	**voiced**	TC-N	92%
			voiceless	TC-N	90%
	Heritage Lebanese	**Initial**	**voiced**	TC-N	70%
			voiceless	UC-C	62%
		Medial	**voiced**	TC-N	88%
			voiceless	TC-N	92%
		Final	**voiced**	UC-N	79%
			voiceless	UC-P	69%
	English	**Initial**	**voiced**	UU-P	60%
			voiceless	UU-C	46%
		Medial	**voiced**	TC-N	77%
			voiceless	TC-N	82%
		Final	**voiced**	UU-P	71%
			voiceless	UU-C	58%

As can be seen in Table 8.2 and 8.3, all medial geminate and singleton target words were categorised into the correct voicing x gemination category by all three groups of listeners. All correct category choices exceeded the preset 70% threshold with the highest categorisation mean percent shown by the Lebanese group as indicated by all the target words being correctly categorised more than 90% of the time.

Since all target stimuli were categorised above the 70% threshold and significantly into their equivalent response categories, the voiced and voiceless gemination contrasts were Two Category (TC) assimilations for all listener groups with no overlap between response categories, according to PAM.

The final voiced and voiceless geminate/singleton contrasts were more challenging than their medial counterparts, consistent with our predictions. For the Lebanese group, all target words were categorised correctly without overlap, which makes both contrasts Non-overlapping TC (TC-N). For the Heritage group, however, while the singletons /qabat/ (*M* = 71%) and /qamad/ (*M* = 79%) were both Categorised, both geminates /qabat:/ (*M* = 62%) and /qamad:/ (*M* = 64%) did not reach the 70% categorisation threshold despite approaching it. Moreover, the voiceless geminate was assimilated with partial overlap: /qabat:/ into <qamat> (*M* = 33%), which makes voiced contrast non-overlapping Uncategorised-Categorised (UC-N) and the voiceless contrast partially-overlapping Uncategorised-Categorised (UC-P). For the English group, only the singletons approached the threshold for categorisation but were not deemed categorised: /qamad/ into <qamad> (*M* = 69%) and /qabat/ into <qabat> (*M* = 65%). Furthermore, the voiced contrast shared one above chance level category choice, which makes it partially-overlapping Uncategorised-Uncategorised (UU-P). Conversely, the voiceless contrast shared more than one above chance level category choice. This makes it completely-overlapping Uncategorised-Uncategorised (UU-C).

In line with our predictions, categorisation of initial gemination was challenging for all groups, particularly with regard to the voiceless contrast. Listeners of the Lebanese group assimilated the voiced initial contrast /damah/-/d:amah/ into the intended response categories <damah> and <ddamah> with no overlap. Conversely, while /tamah/ was significantly categorised correctly (*M* = 81%) with no overlap by the Lebanese group, its geminated

counterpart /t:amah/ was assimilated as intended (<ttamah>) only 64% of the time. This makes the latter uncategorised, and since the singleton response category <tamah> was selected above chance level (*M* = 34%) for the target voiceless geminate /t:amah/, /tamah/ /t:amah/ would be considered a partially-overlapping Uncategorised-Categorised assimilation type (UC-P).

Heritage Lebanese listeners categorised the initial voiced and voiceless singleton/contrast with more overlap than the Lebanese group. The voiced contrast was categorised into its equivalent categories with no overlap as TC-N. While the initial voiceless singleton target word /tamah/ was categorised correctly as <tamah> (*M* = 70%), its geminated counterpart /t:amah/ was not categorised above the threshold (*M* = 63%). Moreover, the voiceless contrast shared more than one overlapping above chance category choice: /tamah/ into <ttamah> (*M* = 30%) and /t:amah/ into <tamah> (*M* = 32%). This makes the voiceless initial contrast a completely-overlapping Uncategorised-Categorised (UC-C) assimilation type.

The categorisation results yielded more uncategorised and more overlap between categories for the English group compared to the other two groups. According to the 70% threshold, none of the target words was be deemed categorised despite all response categories being selected above chance level. This can reflect the difficulty English listeners encountered with categorising the initial geminate/singleton contrasts. Furthermore, the voiced contrast was realised with a partial overlap. Thus, it is considered partially-overlapping Uncategorised-Uncategorised (UU-P). Meanwhile, the voiceless contrast shared more than one above chance level category choice, which makes it completely-overlapping Uncategorised-Uncategorised (UU-C).

8.4.2 Predictions

The assimilation types obtained from the categorisation experiment can be used to predict how the voiced and voiceless medial, final and initial contrasts should be discriminated according to PAM principles.

In the medial position, since all the contrasts were categorised as TC contrast assimilation types, discrimination is expected to be excellent. For final contrast categorisation results, only the Lebanese group categorised /qabat/-/qabat:/ and /qamad/-/qamad:/ as a non-overlapping TC assimilation (TC-N). Discrimination by Lebanese listeners should thus be excellent for both contrasts, according to PAM principles. As for the Heritage Lebanese group, both contrasts were assimilated as Uncategorised-categorised with a partial overlap for the voiceless contrast (UC-P) and no overlap for the voiced contrast (UC-N). Discrimination is predicted to be good but not near ceiling. The voiceless contrast is expected to be discriminated less accurately than the voiced contras, given the partial overlap. As for the English group, /qabat/-/qabat:/ was assimilated as partially overlapping Uncategorised-Uncategorised (UU-P). /qamad/-/qamad:/ was a completely-overlapping Uncategorised-Uncategorised (UU-C) assimilation type. Poor discrimination for the final contrast by English listeners would therefore be predicted, particularly for the voiceless contrast given the degree of overlap.

As for the initial contrasts, /damah/-/d:amah/ is the only contrast which yielded a non-overlapping TC-N for both Lebanese groups. Discrimination by listeners of these two groups is expected to be excellent. /tamah/-/t:amah/ was assimilated as partially overlapping UC (UC-P) by the Lebanese listeners and completely-overlapping UC (UC-C) by the Heritage group. Discrimination is expected to be good for UC-P and moderate for UC-O. The initial voiceless contrast should be discriminated less accurately than the voiced contrast by both groups. The lower percentage of categorisation of the target words to the correct response categories by the

English listeners yielded partially overlapping Uncategorised-Uncategorised (UU-P) for the voiced contrast and completely-overlapping UU (UU-C) for the voiceless contrast. Discrimination is expected to be poor for the initial contrasts by the English group, particularly for the voiceless contrast.

8.4.3. Discrimination results

We constructed a linear mixed effect model (LME) with discrimination accuracy scores as a dependent variable, and Group (Lebanese, Heritage Lebanese, English), Position (Medial, Final, Initial), Voicing (Voiced, Voiceless), and Gemination (X in the AXB = Geminate vs Singleton) as fixed factors. Listener was fitted as random intercepts, since random slopes did not improve the model fit. *p*-values were generated using library lmerTest (Kuznetsova et al., 2017). Any *p*-value of .05 or below was deemed significant.

The mean percent accuracy scores are presented in Figure 8.2. The LME analysis revealed a significant main effect of Position, $F(2, 243) = 12.6$, $p < .001$. Tukey HSD Post-hoc comparisons showed that the overall discrimination scores were highest in the medial position ($M = 88\%$) followed by final position ($M = 77\%$) (*ps* < .001), then initial position ($M = 67\%$), (*ps* < .001). Also, there was a significant main effect of Group, $F(2, 243) = 19.9$, $p < .001$. Comparisons of group means indicated that the Lebanese group's discrimination accuracy ($M = 89\%$) was higher than Heritage Lebanese group's (M =77%), (*ps* < .001), who performed significantly better than the English group ($M = 66\%$), (*ps* < .001). These main effects were accompanied by a significant interaction between group and gemination, $F(2, 243) = 15.2$, $p = .011$. Breaking this interaction down using a Tukey HSD Post-hoc test revealed that the English group performed more accurately when the target in the AXB was a singleton ($M = 72\%$) than a geminate ($M = 60\%$), ($p < .001$), but this effect was not significant for both

Lebanese groups. Another two-way interaction between Group and Position was significant, $F(2, 243) = 33.8$, $p < .001$. This effect can be seen in Figure 8.2, where the relative difference in accuracy for medial versus initial and final contrasts appears to be smaller for the Lebanese group than the Heritage and English groups. It seems to be particularly apparent for the medial versus final contrasts. Please refer to Table 8.3 for the mean percent accuracy scores for each contrast.

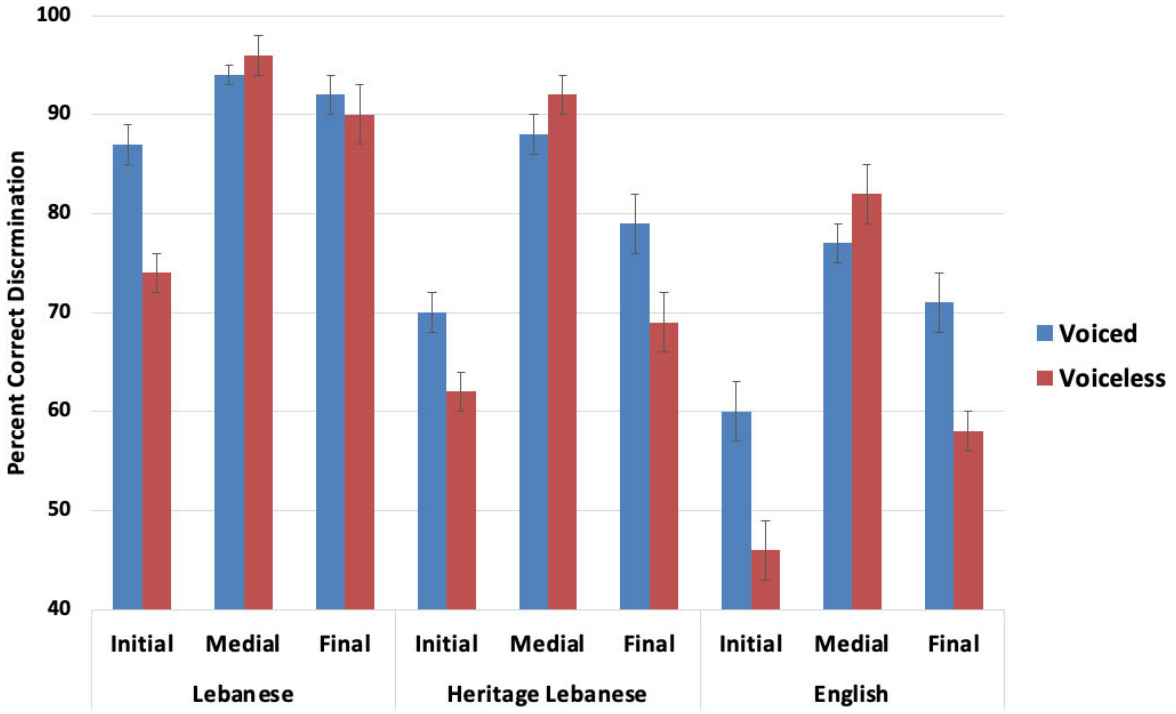

Figure 8.2: Mean percent correct discrimination scores of the voiced/voiceless geminate/singleton contrasts for the Lebanese, Heritage Lebanese and English groups

In addition, there was a significant interaction between voicing and position, $F(2, 243) = 13.4$, $p = .010$, and another 3-way interaction between Voicing, Position and Group, $F(2, 243) = 16.8$, $p < .001$. Since the two interactions involved position, we decided to break them down by running separate LME models for each word-position type with discrimination accuracy scores as a dependent variable, Group (Lebanese, Heritage Lebanese, English),

Voicing (Voiced and Voiceless) and Gemination (X in the AXB = Geminate vs Singleton) as fixed factors. Listener was treated as a random intercept.

The LME analysis of the medial discrimination results revealed a significant effect of group, $F(2, 79) = 20.9$, $p < .001$. Post-hoc multiple comparisons showed that while the difference between the Lebanese and heritage Lebanese (M = 95% vs 90%) only approached significance ($p = .052$), both Lebanese ($p < .001$) and Heritage Lebanese ($p < .001$) discriminated the medial contrasts more accurately than the English group (M = 80%). There was also a significant effect of voicing, $F(1, 79) = 12.1$, $p = .011$, and a significant interaction between group and voicing, $F(2, 79) = 19.5$, $p = .003$. Breaking this interaction down indicated that the voiceless contrast was discriminated more accurately than the voiced contrast for the Heritage group, M = 92% vs. 88%, $p < .001$) and the English group, M = 82% vs. 77%, $p = .031$, but not for the Lebanese group, M = 94% vs. 96%, $p = .09$. Although there was no significant main effect of gemination, there was a significant interaction between group and gemination, $F(2, 79) = 9.6$, $p = .042$, i.e., only the English group performed slightly better when the target in the AXB was a singleton (M = 85%) than a geminate (M = 75%), ($p < .001$).

The LME analysis of the discrimination accuracy for final contrasts indicated a significant effect of group $F(2, 83) = 78.7$, $p < .001$, with the Lebanese group's performance (M = 91%) exceeding that of the Heritage Lebanese (M = 74%), ($p <.001$) and with the English' performance coming close third (M = 65%), ($p <.001$). Moreover, the voiced contrast (M = 81%) was discriminated more accurately than the voiceless contrast (M = 72%) in the word-final position, as indicated by a significant effect of voicing, $F(1, 83) = 29.4$, $p < .001$. This was also modulated by a significant interaction involving group and voicing, $F(2, 83) = 13.8$, $p = .009$. The only group whose discrimination accuracy of the voiced contrast did not significantly surpass that of the voiceless contrast was the Lebanese group, M = 90% vs. 92%.

Last, there was a significant interaction between group and gemination $F(2, 83) = 8.9$, $p = .049$. The English group showed an advantage in discrimination when the target in the AXB was a singleton ($M = 71\%$) than a geminate ($M = 59\%$), $p < .001$. The difference was not significant for both Lebanese groups.

As for the initial contrast results, there was a significant main effect of group, $F(2, 81) = 59.1$, $p < .001$. Multiple comparisons between groups showed that the Lebanese group performed the most accurately ($M = 81\%$), compared to the Heritage group ($M = 66\%$), $p < .001$, and then the English group ($M = 53\%$), $p < .001$. There was also a significant effect of voicing, $F(1, 81) = 67.3$, $p < .001$, i.e., the voiced contrast ($M = 72\ \%$) yielded higher overall discrimination accuracy scores than the voiceless contrast ($M = 61\%$). Last, there was a significant main effect of gemination, $F(1, 81) = 33.3$, $p < .001$. The overall discrimination accuracy was higher when the target in the AXB was a singleton than a geminate. This was modulated by a significant interaction involving group. There was no difference between the singletons and geminates for the Lebanese group in the initial position of the word, $M = 83\%$ vs. 79%, $p = 0.70$. The discrimination of the initial contrast was significantly more accurate when X in the trial was a singleton than a geminate by the Heritage, $M = 70\%$ vs. 62%, and English groups, $M = 59\%$ vs. 47%.

8.5. Discussion

In this experiment, we investigated the categorisation and discrimination of word-medial, -final and -initial voiced and voiceless geminate/singleton contrast in Moroccan as perceived by Lebanese, Heritage Lebanese and English listeners.

To help interpret the categorisation and discrimination results, a quick reminder should be given about the status of gemination in the native languages of the target listener's groups.

Gemination is not contrastive in English. Thus, English speakers are not familiar with phonologically contrastive gemination in any word position. They do, nonetheless, have experience with purely phonetic "fake" gemination across word boundaries, that is analogous to word-medial gemination. English speakers lack even phonetic-level experience with initial and final gemination. Lebanese speakers, on the one hand, have native phonological experience with initial and medial in their vernacular dialect, and medial and final gemination contrasts in MSA, in which they carry a relatively high functional load. On the other hand, Heritage Lebanese speakers have native phonological experience with initial and medial gemination in the Lebanese dialect that they learned at home as their first language, but no experience with MSA final gemination, since they never learnt MSA. We predicted that the three groups' success at discriminating the medial, initial and final gemination contrasts in the Moroccan target stimuli would depend on their familiarity and experience with the contrast in each word position. In particular, we predicted that phonological-experience with gemination (the two Lebanese groups) would yield better discrimination than phonetic-only experience with gemination (English group).

To begin with, since all groups have phonological or phonetic-only experience with medial gemination, they were expected to be able to attend to differences in consonant duration in the medial position of the word. The discrimination results showed the three groups were successful in discriminating the Moroccan medial contrasts: Lebanese (M = 95%), Heritage Lebanese (M = 90%) and English (M = 80%). The difference between the Lebanese and Heritage groups only approached significance, but both groups performed better than the English group in discriminating the medial contrasts. This superior performance over the English group can be attributed to the fact that medial gemination is contrastive and highly functional in the Lebanese groups' native language. Thus, phonological experience resulted in more success with discrimination than purely phonetic experience with gemination. The high

success with discriminating the medial contrasts by the three groups was predicted by the non-overlapping Two-Category assimilation types that the categorisation results revealed for the three groups of listeners. However, the difference between the groups' accuracy scores is not accounted for through the categorisation results.

The difference between the performance of the Lebanese and Heritage Lebanese is more pronounced in discriminating the final contrasts as indicated by the significant main effect of group. The Lebanese group discriminated the final contrast the most accurately (M = 91%), followed by the Heritage Lebanese group (M = 74%), then the English group (M = 65%). The English group's poorer performance compared to the Lebanese groups was expected due to their lack of both phonological and phonetic experience with final gemination in their native language. Although gemination is contrastive in Lebanese, its occurrence in the final position of the word is strictly limited to a small number of words containing final geminated consonants. These words are borrowed from MSA and their singleton counterparts rarely exist in the lexicon of Lebanese. Thus, the productivity of final gemination is very limited in Lebanese compared to MSA, where final gemination carries a relatively high functional load. Since final gemination is structurally and phonetically different from both initial and medial gemination (see Chapters 6 and 7), Heritage Lebanese listeners' lower performance compared to that of the native Lebanese listeners was expected. Another factor that could have influenced these results is the two groups' language use on a daily basis. The Heritage Lebanese listeners are dominant in English. They grew up and live in an English-speaking environment. Native Lebanese listeners, on the other hand, grew up in Lebanon and were exposed to Lebanese as their primary language in their day-to-day life until they moved to Australia as adults. They also learned MSA in Lebanon, whereas the Heritage Lebanese participants did not learn MSA. These difference in language learning and contact could offer an explanation for the lower

performance of the Heritage Lebanese group compared to the Native Lebanese group in discriminating the final contrast.

These results also demonstrate the influence of lower-level phonetic details on the perception of non-native contrasts, as proposed by PAM. These results could also be explained through the categorisation results. The Lebanese group was the only group that assimilated the final contrasts as non-overlapping TCs (whose discrimination is expected to be excellent). The Heritage group assimilated the contrasts as UC-N for the voiced and UC-P for the voiceless contrasts. UC assimilation is expected to be lower than a TC, which explains the Heritage Lebanese group's performance being significantly lower than the Lebanese group. The presence of a partial overlap for the voiceless contrast by the Heritage group was also expected to significantly decrease the discrimination accuracy scores for the voiceless contrast relative to the voiced contrast (see Table 8.3). The English categorisation of the final contrast yielded a partially overlapping Uncategorised-Uncategorised (UU-P) for the voiced contrast and a completely overlapping Uncategorised-Uncategorised (UU-C) for the voiceless contrast. UU assimilation is traditionally discriminated worse than a UC and more so than a TC, especially when realised with a complete overlap. Thus, PAM's predictions were upheld.

It is not entirely clear from the literature whether initial gemination is more discriminable than final gemination or vice versa. Though they are rare, studies comparing their discriminability showed conflicting results as we explained in the introduction. Therefore, it was important to investigate the perception of the initial contrasts and compare it with that of the final contrasts. Of particular importance are the results of the Lebanese listeners, since they have a native phonological experience with the contrast at both word-periphery. Lebanese listeners' discrimination of the initial contrasts was significantly lower (M = 81%) than their discrimination of the final contrasts (M = 91%), as revealed by the main effect of position. This

finding is compatible with results reported in (Hallé, Ridouane & Best, 2016) for the discrimination of Berber initial and final gemination by French naive listeners. Although initial gemination is contrastive in Lebanese dialectal Arabic, Heritage Lebanese listeners discriminated the initial contrast significantly less accurately (*M* = 66%) than the Lebanese group (*M* = 81%). This lower performance by the Heritage group should not be entirely surprising. It is very likely due to the language use patterns, which has been proven to influence the perception of phonemes (e.g., Flege & MacKay, 2004). Our Australian Heritage speakers are L2-English dominant, and presumably have more restricted use and knowledge of Lebanese dialect, even though it was their first/home language, because they had been educated in English and grown up in an English environment, speaking mostly English. Therefore, it would not be implausible to expect them to perform less proficiently than native listeners. This scenario is in accord with findings from other research studies comparing heritage language listener's performance to that of native listeners (e.g., Chang, 2016 & Lee-Ellis, 2012). Importantly, Heritage listeners fared significantly better than the English group (*M* = 53%). This result further highlights the importance of native phonological experience with a contrast and its advantage in enhancing perception of non-native speech, as PAM advocates. In terms of categorisation, the discrimination results largely supported PAM's predictions. The English group categorised the initial contrasts as Uncategorised-Uncategorised. Thus, their poorer performance in discriminating the initial contrasts compared to the other groups was to be expected. Both Lebanese groups assimilated the initial voiced contrast as non-overlapping TC. Although PAM predicts highly successful discrimination for TC-N, this was the case for the Lebanese group (*M* = 87%), but less so for the Heritage group (*M* = 70%).

The initial voiceless contrast was the most challenging for the Lebanese groups: Lebanese (*M* = 74%) and Heritage Lebanese (*M* = 62%). These low accuracy scores are accounted for in terms of the categorisation results. The only contrast that the Lebanese

assimilated as partially-overlapping UC was the voiceless contrast. The heritage group assimilated the initial voiceless contrast as UC too but with complete overlap, which explains the significant difference between both groups in discriminating the contrast. Moreover, an attempt to explain the lower performance with the initial voiceless contrast should recall the reduced phonetic cues to gemination in the absolute initial position of the word: the lack of the preceding vowel renders the closure inaudible for voiceless stops. Although native listeners of word-initial geminating languages have been shown to rely on other non-closure cues to discriminate the initial voiceless contrast (e.g., Kelantan Malay: Hamzah, 2013), non-native listeners should not necessarily be able to attend to subtle phonetic details in a non-native contrast. Since Moroccan stimuli are non-native to both Lebanese and Heritage Lebanese listeners, regional dialectal differences in realising initial gemination in Moroccan and Lebanese dialects (Chapter 6) should not be ruled out as an explanation for the challenge with discriminating the Moroccan initial voiceless contrast by the Lebanese listeners. This further emphasises the contribution of lower level phonetic details in non-native speech perception, as PAM posits.

Voicing, was therefore an important factor in the discrimination accuracy score results. Although not significantly different for the Lebanese group, overall accuracy scores were significantly higher for the voiceless contrast than voiced contrast in medial position for the Heritage Lebanese and English groups. One plausible explanation for this result is that voiceless stops are realised with a significantly longer closure duration than voiced stops, and the flanking vowels in the medial position provide an ideal context for this difference in closure duration to be perceived. At both word peripheries, higher discrimination accuracy was observed for voiced contrast compared to the voiceless contrast. The initial voiceless contrast was discriminated the poorest among all contrasts, and the English group failed to discriminate this contrast. Since closure duration is acoustically inaudible, listeners should attend to other

acoustic cues differentiating geminates and singletons. One such cue is the higher amplitude of the burst in the geminate context (see Chapter 7). English listeners are speculatively not sensitive to this acoustic cue in contrasting geminates to singletons word-initially. Heritage Lebanese group and especially Lebanese group were presumably more attentive to those non-closure duration acoustic cues. However, this research does address the perceptual relevance of acoustic cues that signal the Moroccan geminate/singleton contrast to non-native listeners. This would certainly be an interesting and valuable avenue for future research on this topic.

Most important, however, is the finding that sensitivity to consonant duration distinctions in one word position does not guarantee the capability to perceive consonant duration distinctions in other word positions. This effect is particularly apparent for the English group. Despite having shown good discrimination of the medial contrast, where they do have some phonetic experience in the form of cross-word "fake" gemination effects in their native language, the English listeners showed substantial difficulty with peripheral gemination, which they do not experience even phonetically in English. Importantly, this result reveals the importance of low-level phonetic details alongside the abstract phonological level in the categorisation and discrimination of non-native phones, as posited by PAM.

A note should be made about how other non-native speech perception models would have predicted some of the findings in this study. Of particular importance is how English listeners discriminated and categorised geminates. Not only did English listeners categorise geminates more accurately than singletons, but their discrimination of the contrasts significantly improved in the singleton context. This is not entirely surprising since English has only singletons and not geminates (phonologically contrastive, phonetically lengthened consonants) as native categories. Hence, according to PAM, less accurate discrimination is expected for geminates compared to singletons. Conversely, SLM's notion of the *new* versus

similar non-native phones would have predicted geminates to be perceived more accurately than singletons by the English group, since singletons are similar to native English categories. Our findings contradict SLM predictions in this regard.

Another important finding from the current experiment is that phonological overlap is a considerable factor in determining the level of success in discriminating the geminate/singleton contrasts. For example, the geminate/singleton contrasts in the medial position were the only contrasts that were categorised with the least amount of overlap. We demonstrated that the larger the level of overlap, the more poorly the contrast would be discriminated. Contrasts with partial overlap were discriminated significantly less accurately than contrasts with no overlap. For example, the Heritage Lebanese group assimilated both final contrasts as a UC. The only difference was that a partial overlap was witnessed for the voiceless contrast, and no overlap for the voiced contrast. This was also reflected in significantly lower performance in discriminating the final voiceless contrast compared to the final voiced contrast. Similarly, completely overlapping contrasts were discriminated more accurately than partially overlapping contrasts. For example, the English group discriminated voiced contrasts at the periphery of the word more accurately voiceless contrasts. This was due to the presence of complete overlap in categorising the word-initial and -final voiceless contrasts compared to their voiced counterparts. Perceived phonological overlap has been shown in previous research to influence the discrimination of non-native phones (Best, Avesani, Tyler & Vayra, 2019). However, this is the first study to confirm that it also has an effect on the perception of consonantal length contrast across listener groups differing in phonological and phonetic experience with gemination.

Finally, it is important to acknowledge that the number of stimuli used in these perception experiments is small and that the results reported should therefore be interpreted

with some caution. Future research should tackle this issue by testing a larger number of stimulus items from a wider range of consonant types. Another limitation concerned the fact that the listener's attention was directed to durational but not to non-durational aspects of the geminate/singleton contrast. The production component (Chapter 6 and 7) of this thesis found significant non-durational acoustic contributions to the peripheral gemination contrasts.

8.6. Conclusion

This study was conducted to tease apart the contribution of higher level phonological and lower level phonetic information on the discrimination and categorisation of gemination across three different word positions (medial, initial and final) by three groups of listeners varying in their phonological (Lebanese and Heritage Lebanese groups) and phonetic-only (English group) experience with gemination. The results showed that the three groups' success with discriminating the medial, final and initial gemination contrasts in the Moroccan target stimuli was dependent on their differing experience with the contrast in each word position. Phonological-level experience with gemination contrasts in a listener group's native language(s) (the two Lebanese groups) resulted in better discrimination than experience only with phonetic-level "fake" gemination across word boundaries (English group). Phonological experience with medial gemination was more readily transferable to support non-experienced word-final gemination distinctions by the Heritage Lebanese group than was only phonetic-level experience by English group. Voicing along with word-position were considerable predictors. The voiceless contrast in the medial position was discriminated more accurately than its voiced counterpart (non-significant for the Lebanese group), but the opposite was witnessed in the word-final and especially word-initial position. Proficiency and experience also affected the discrimination accuracy scores. The Lebanese group fared the best in all contrasts followed closely by the Heritage Lebanese group. Therefore, word-position, voicing

and proficiency were all important factors affecting the categorisation and discrimination of non-native gemination contrast. Importantly, these are factors that highlight the importance of low-level phonetic details within the same phonological contrast and how they influence non-native perception of gemination. This has implications for PAM, whose application had not been utilised to account for positional effects on perception of consonantal length.

9. GENERAL DISCUSSION

This PhD project reported on a comprehensive articulatory, acoustic and perceptual investigation of voiced and voiceless coronal stop gemination in regional varieties of Modern Standard Arabic (MSA) and Arabic vernacular dialects, as produced by native Arabic speakers from two geographically distant regions of the Arab world: Morocco and Lebanon. Various acoustic cues, both durational and non-durational, were assessed in how they signalled the geminate/singleton contrast in word-medial position as compared to two much rarer types of gemination: contrasts in initial and final word positions. The Moroccan production data provided the stimuli and the basis for the subsequent perception experiment involving three groups of non-native listeners — naïve English, heritage Lebanese and native Lebanese speakers — in order to investigate their ability to discriminate and categorise the Moroccan productions.

Two production experiments and one perception experiment make up the experimental component of this thesis. This chapter sums up the principle findings across these three experiments. The results will be linked back to the theoretical background of the thesis, and the implications of the findings, the methodological/analytic advancements offered by the studies, the limitations of the thesis work, and future research directions will be discussed.

9.1. Production findings

9.1.1. Acoustics: Durational cues

Duration of the closure is often described as the most universally consistent cue to consonantal geminate/singleton contrasts (Ham, 1998). Quantifying this duration is an important step towards determining the extent to which the closure duration differs between a singleton and a geminate and whether this difference is indicative of a phonological length contrast. While several linguistic factors have been previously shown to affect the duration of the closure — e.g., stress (Dmitrieva, 2012), voicing (Al-Tamimi & Khattab, 2018), syllable structure (McCrary, 2002) and speaking rate (Hirata & Whiton, 2005) — of particular interest in the current study is how this closure duration varies across word and utterance position and from one dialect or one regional variety of a language to another. Word-initial gemination in Moroccan and Lebanese dialectal Arabic was compared across two utterance positions: utterance-initial and -medial positions. Medial and final gemination in the Moroccan and Lebanese varieties of MSA were compared only across word-position: medial versus final word-positions.

Our acoustic results indicate that coronal voiced and voiceless stop geminates are realised in the target language and dialects with a significantly longer closure duration than singletons across utterance/word positions and voicing types, with the exception of initial voiceless stops, for which acoustic closure duration could not be measured. Overall, however, closure duration of both geminates and singletons was longer for initial contrast in utterance-initial position and for word-final contrast as compared to word-medial contrast and word-initial contrast in utterance-medial position. The increase in duration affected both geminates and singletons, indicating that closure duration was affected more by prosodic and positional effects than by the effect of the geminate/singleton phonological length contrast.

The geminate-to-singleton closure duration ratio varied across different utterance/word positions. The highest geminate-to-singleton acoustic closure duration ratio was found in word-medial position (*M* = 2.25 for Moroccan and *M* = 1.85 for Lebanese), followed closely by word-initial gemination in utterance-medial position (*M* = 2.10 for Moroccan and *M* = 1.74 for Lebanese), followed by word-final position (*M* = 1.67 for Moroccan and *M* = 1.56 for Lebanese), followed by initial gemination in utterance-initial position (*M* = 1.35 for Moroccan and *M* = 1.40 for Lebanese). Although investigations of gemination across the three main word positions are rare, the finding that the ratio was the highest in the medial position is consistent with other studies. For example, a ratio of 1.7 for word-initial positions and 2.06 for word-medial positions was observed in Russian (Dmitrieva, 2018) and a ratio of 2.1 for word-medial and 1.4 for word-final contrasts was observed in Rural Jordanian Arabic (Al-Deaibes & Rosen, 2019), which is typologically closer to Lebanese and Moroccan Arabic. However, this finding is not consistent across all languages. In Tashlhiyt Berber, a language that also allows gemination across the three word positions, the duration ratio was reported to be higher in word-initial and -final than in -medial position: medial < initial < final for fricatives, and medial < final for stops, which were not measured in word-initial position (Ridouane, 2007). This finding has not been replicated in any other language, to the best of our knowledge. Although the duration ratio at the periphery of the word did not differ significantly between Moroccan and Lebanese, Moroccan speakers consistently produced higher duration ratios across all three word positions. This suggests that closure duration is employed more robustly by Moroccan than Lebanese speakers to contrast geminates with singletons.

This thesis project also investigated the role of durational acoustic cues that are not directly related to closure duration, such as the duration of the stop release burst. The results showed that while gemination did not affect burst duration in word-medial position, voiced stops were affected by gemination in word-final position, with voiced geminates having a

longer burst than voiced singletons. However, in initial position, there was an overall decrease in release burst duration in the geminate compared to the singleton context. This difference in release burst duration was not significant for voiceless stops in Lebanese dialect, whereas gemination affected the release burst duration in both voicing categories for Moroccan dialect. We observed that release burst duration was used more consistently as a correlate of gemination in the Moroccan speakers than in the Lebanese speakers, particularly in word-initial position. Importantly, release burst duration did not differ between geminates and singletons in medial position for either group.

Acoustic investigations of gemination often examine whether there is any effect of temporal compensation between consonant duration and flanking vowel durations, frequently reporting an inverse relationship between the durations of the consonant and the vowel(s). Four out of five of our Moroccan speakers produced shorter durations of the vowel before geminates in word-medial and -final positions, which we interpret as temporal compensation, whereas the Lebanese participants did not shorten pre-geminate vowels. Thus, shortening of the vowel before geminates (i.e. temporal compensation) is a reliable cue to the geminate/singleton contrast in Moroccan but not in Lebanese, particularly in word-final position.

For word-initial gemination, we investigated the duration of the following vowel, which is not commonly examined in the literature. We observed that post-geminate vowels were longer than post-singleton vowels. This is consistent with the results for vowels following medial geminates in languages such as Japanese (Idemaru & Guion, 2008). Lebanese vowel durations were not only longer following the word-initial stops, but the contrast between post-geminate and post-singleton vowels was larger than in Moroccan. The fact that Lebanese speakers produced longer vowel durations following word-initial coronal stops might be indicative of trading relations between the two different durational acoustic cues to gemination,

in which the weighting of the cues appears to differ between Moroccan and Lebanese. Closure duration and preceding vowel duration is weighted more heavily in Moroccan, and preceding vowel duration is weighted more heavily in Lebanese. A larger contrast between post-initial-geminate vowel durations for Lebanese than Moroccan suggests that Lebanese speakers rely more on vowel than consonant closure duration to implement the geminate/singleton stop contrast word-initially, conversely to Moroccan speakers who show the opposite weighting.

9.1.2. Acoustics: Non-durational cues

Another important question that was addressed in the production experiments is how the geminate/singleton contrast is reflected in non-durational acoustic cues. Non-durational cues are investigated less frequently in the literature than durational cues because gemination is commonly believed to be a distinction based purely on the duration of the consonant (e.g., Maltese: Galea, 2016). Interestingly, the acoustic analysis (Chapter 7) revealed significant contributions of release burst amplitude, as well as the amplitude and f0 of the flanking vowels, to the realisation of the geminate-singleton contrast. The release burst amplitude was significantly higher for geminates than singletons in word-final position but not in word-medial position. Moreover, for both groups, release burst amplitude differentiated the word-initial contrast in utterance-initial position, but not in utterance-medial position. This result suggests that speakers produce initial gemination in utterance-medial position similarly to word-medial gemination, at least with respect to this stop release property.

With regard to the non-durational acoustic measures associated with the flanking vowels, significantly higher f0 values were found for word-initial geminates in utterance-initial than in utterance-medial position. Similarly, the difference in f0 values between geminates and singletons was significant in word-final position. In word-medial position, despite a weak tendency for geminates to increase the f0 of the preceding vowel offset, this effect approached

significance only for Lebanese (p = .07) but not for Moroccan. In addition, the vowels preceding word-medial and -final geminates were generally produced with a higher amplitude in the context of geminates than singletons, but the difference was not significant. However, in word-initial contrast, significantly higher values of the following vowel amplitude were observed only for the Lebanese group and only in utterance-initial position.

In summary, two general patterns emerge from the acoustic results. The role of closure duration, which is often regarded as the 'primary' cue to the geminate/singleton contrast, is robust for medial gemination and for initial gemination in utterance-medial position. In word-final and word-initial geminates in utterance-initial position, closure duration differences are less robust but other non-closure duration cues are more enhanced than in medial position. Secondly, the acoustic cues seem to be weighted differently between the two dialects: whereas Moroccan speakers favoured durational cues (closure and vowel durations) to realise the gemination contrast, Lebanese speakers instead favoured non-durational cues.

9.1.3. Articulatory results

One innovation in this study was the use of ultrasound to investigate the temporal dynamics of the geminate/singleton contrast, given that ultrasound has traditionally been used in speech research only to observe lingual articulation at static time points. One possible reason for the dearth of ultrasound research on gemination is that the amount of work and time associated with manually tracing tongue contours can make this technique quite labour-intensive. In order to address this problem, we analysed the ultrasound data using Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; Carignan, 2014) software and LDA modelling. With minimal manual adjustments, this method allowed us to visualise and measure changes in closure duration between geminates and singletons irrespective of voicing type and word position.

To obtain values for the articulatory closure duration, we created a Matlab function to estimate duration values of the closure semi-automatically through determining gestural landmarks from the velocity signal. In order to evaluate the validity of this articulatory closure metric from the articulatory data, we performed a standard correlation analysis on the results from utterance-medial data obtained from the acoustic analysis and using the articulatory closure calculation function. There was a strong positive correlation, $r = 0.96$, between the articulatory duration and the acoustic duration of the closure in the utterance-medial position, thus validating the closure duration estimation technique we created. Next, we corrected the values generated by the articulatory analysis, because it was showed that articulatory and acoustic closure durations do not have the same range of values despite the strong positive correlation between the estimations of duration generated by both methods. This was done to be able to use the articulatory closure duration values as accurate proxy for acoustic closure duration. This was particularly useful for word-initial voiceless stops and word-final burstless stops, whose closure duration could not be measured using acoustic means.

The results we obtained showed that geminates were produced with a longer articulatory closure duration than singletons in word-medial, -final and -initial positions. In line with the acoustic results, articulatory closure duration values were longer for initial and final gemination, but the contrast between geminates and singletons was proportionally smaller than medial gemination (please refer to sections 6.6.1 and 7.4.1 for more details about the articulatory results). Importantly, the articulatory results support the use of ultrasound and the innovative analytic techniques we devised in offering a reliable solution for visualising and quantifying duration differences for contrasts whose closure durations cannot be measured from the acoustic signal.

9.1.4. Implications of the production findings

Prior to this thesis work, it was not clear whether gemination is maintained word-initially in the Moroccan and Lebanese dialects, as well as word-finally in the Moroccan and Lebanese varieties of MSA. Impressionistic studies had argued that initial gemination has been neutralised in dialectal Arabic and final gemination is not contrastive in MSA, but there was previously no systematic instrumental acoustic and/or articulatory evidence to empirically address this claim. The durational and non-durational acoustic cues that we measured and compared across coronal stop singletons and geminates of both voicing types, and across the three main word positions, indicate that gemination is contrastive and maintained at the periphery of the word in MSA and in these vernacular varieties of Arabic. This finding is significant and unprecedented because the geminate-singleton contrast has been shown to be prone to neutralisation at word boundaries in other languages which had historically used the contrast in these peripheral positions, e.g., initial gemination in Maltese (Hoberman & Aronoff, 2003) and final gemination in Hungarian (Curtis, 2003).

Our acoustic results also revealed that closure duration is not the only cue to signalling the geminate/singleton contrast in these languages. This finding is incompatible with some research that has claimed that increased closure duration for geminates is the single effective acoustic cue that distinguishes geminates from singletons. For example, geminates in Maltese were reported to be realised with a longer closure duration than singletons, but gemination did not affect any of the other acoustic cues examined (e.g., VOT and vowel duration; Galea, 2016). Conversely, our results have shown that, in some word and utterance positions, non-durational cues in fact play a stronger role than closure duration in contrasting geminates with singletons. For example, for two Lebanese speakers, articulatory and acoustic closure duration values did not vary significantly between geminates and singletons in utterance-initial position.

Interestingly, these two Lebanese speakers realised the gemination contrast using acoustic cues that were also used by the other Lebanese speakers: they produced geminates with higher burst amplitude than singletons, and f0 values were significantly higher at the vowel onset following a geminate than following a singleton.

Prosodic position has systematic and significant effects on the phonetic realisation of gemination. Variant phonetic realisation of gemination is often linked to the position of the contrast in the word (Pajak, 2010). Our production results are compatible with this account. We showed in Chapter 7 that both geminates and singletons have longer closure durations in word-final than in word-medial position, for both Moroccan and Lebanese speakers. However, the duration difference between geminates and singletons is proportionally larger in word-medial than -peripheral positions, as evident in a significantly larger geminate-to-singleton ratio for medials than peripherals (Frej, Carignan & Best, 2017). Likewise, word-initial gemination in Moroccan and Lebanese dialects was realised with a smaller geminate-to-singleton closure duration ratio despite being produced with substantially longer average closure duration values than word-medial gemination. These results indicate that closure duration is a more robust cue in word-medial position than in either of the peripheral positions. Moreover, we have shown that other, non-closure durational cues are more enhanced at the edges of the word than word-medially; however, this finding was not consistent across Moroccan and Lebanese varieties.

Although this variability in the phonetic realisation of gemination seems at first glance to be caused by the word position of the contrast, we would like that to argue that the higher-level prosodic position of the word within an utterance has an even stronger influence. By placing the initial geminate/singleton contrast at two different positions in the utterance, i.e., utterance-initially and -medially, and testing how Moroccan and Lebanese speakers produce

the different acoustic cues we have examined, we demonstrated that the phonetic implementation of phonetic cues changes dramatically from one utterance position to the other; the consistency of this change across both varieties of Arabic suggests that this variation is systematic. For example, geminates and singletons alike were produced with significantly longer closure duration in utterance-initial position compared to utterance-medial position. Nonetheless, the geminate-to-singleton ratio was larger in the utterance-medial position of word-initial stops, analogous with the results found for word-medial gemination. Furthermore, the prominence of absolute utterance position effects over word position effects alone was also be found in some speaker-specific results. Although Lebanese speakers L2 and L5 failed to differentiate between geminates and singletons through closure duration, this concerned only their productions of initial geminates in utterance-initial position. In utterance-medial position, both L2 and L5 contrasted geminates and singletons by producing significantly longer closure duration values for geminates in the same way as the rest of the speakers had for both word-initial and word-medial targets. Thus, utterance position can lead to either maintenance or neutralisation of certain acoustic distinctions to gemination contrasts in word-initial position.

In addition, we evaluated the status of the epenthesis vowel preceding word-initial geminates. we noted that vowel epenthesis was observed for word-initial geminates in utterance-initial position for the same two Lebanese speakers. Around 20% of the geminated tokens in the utterance-initial position produced by L2 and L5 (and only these two speakers) were preceded by a schwa-like vocalic element. Vowel epenthesis before word-initial geminates has, indeed, been reported in other languages. For example, in the acoustic investigation of word-initial gemination in Maltese (Galea & Ussishkin, 2018), native speakers inserted an [i]-like vocalic element before geminates in 96% of cases. The authors posited that initial gemination never truly occurs syllable-initially in Maltese, and that the vowel insertion thus serves to re-syllabify word-initial geminates. We cannot make the same assertion for our

results given that only two Lebanese speakers produced an epenthetic vowel, and even for those speakers only ~20% of closures in absolute initial position were preceded by this vocalic element. Thus, pre-closure vocalic insertion is not a consistent feature for the Lebanese dialect and not even consistent or highly frequent for these two speakers. However, this vowel epenthesis is nonetheless compatible with the underlying phonological form of initial gemination in Arabic. Gemination in our two target dialects arises as a result of a morphological process whereby the underlying form of the definite article in Arabic /al/ is reduced to /l/ which then assimilates in manner and place to the coronal onset of the stem morpheme. As a result, three different possible scenarios can be outlined as the cause of the vocalic insertion: 1) The epenthetic vowel is a "reappearance" of the article's /a/, which was deleted, suggesting a morphological process; 2) The epenthetic vowel is a way of restructuring the syllable, so that the syllable does not begin with a consonant, suggesting a phonological process; 3) The epenthetic vowel is a way of providing phonetic context for realising the geminate in production, suggesting a phonetic process. Any of these scenarios, or some combination of them could be happening at the same time. However, the current study does not provide a definitive way of deciding among these possibilities. This would be an interesting and informative issue to address in future research.

Differences between the dialects in realising peripheral gemination may be attributed to their different syllable structures. Previous linguistic studies have classified Moroccan and Lebanese dialects into different syllable structure groups: C-dialects (Moroccan: Harrell, 2004) and VC-dialects (Lebanese: Blanc, 1953 & Cowell, 2005). VC-dialects are those in which morphologically derived CCC clusters are syllabified as CVCC and C-dialects are those in which CCC clusters remain syllabified as CCC. One important implication of this difference in syllable structure is that clusters at the periphery of the word occur more freely in Moroccan than in Lebanese. A study on the frequency of initial clusters in a number of Arabic dialects

including Moroccan and Lebanese found that initial consonant clusters are three times more frequent in Moroccan than in Lebanese (Hamdi et al., 2005). Thus, if peripheral gemination patterns with peripheral clusters, this would increase the functional load of peripheral gemination in Moroccan compared to Lebanese, which in turn would increase the contrastiveness of peripheral gemination in Moroccan relative to Lebanese. This could explain why the contrast in closure duration was found to be larger in Moroccan than in Lebanese. We interpreted this finding as that peripheral gemination is more likely to undergo neutralisation in Lebanese than in Moroccan. More support for this proposal was the inability of the two Lebanese speakers to systematically contrast geminates with singleton in absolute initial position using closure duration. Another potential contributor in these dialectal differences could be the different way the two dialects assign stress, as was mentioned in Chapter 6. However, confirming these possibilities would require additional research to address these theoretical speculations more directly.

The question of whether gemination is a duration distinction or a distinction in articulatory force was also addressed. Our findings indicated that while initial gemination in utterance-medial position and word-medial gemination were mainly distinguished through differences in closure duration, initial gemination in utterance-initial position and word-final gemination were realised with strengthening of characteristics that are associated with forceful articulation. We would posit that these effects, which imply a distinction of articulatory force, were caused by prosodic factors rather than by the gemination contrast alone. To explain, strengthening of those non-closure duration cues occurred only at both peripheral positions of the utterance, which are prosodic positions that have been shown to cause articulatory strengthening/enhancement of segments. This is compatible with the contrast maximization principle (Cho & Jun, 2000; Cho & Keating, 2001), which predicts segments to be elongated and the contrast between a segment and another contrastive segment in the sound system of a

language (e.g., geminate vs. singleton) is maximised. Our findings, however, provided only partial support for this account. Closure duration was substantially longer in utterance-initial/word-initial and utterance-final/word-final positions as compared to word- and utterance-medial positions. However, the geminate-to-singleton contrast ratio was slightly larger for word-medial contrast word-initial contrast in utterance-medial position compared to word-final contrast and word-initial contrast in utterance-medial position.

After having established that gemination is contrastive and maintained through durational (closure, release burst and vowel durations) and non-durational (burst and vowel amplitude, vowel offset/onset f0) acoustic cues both word-initially in Moroccan and word-medially and -finally in the Moroccan variety of MSA, we then turned to the other main goal of this thesis: determining whether and how these acoustic cues may be relevant to the perception of gemination by non-native listeners who vary in terms of their phonological or phonetic-only experience with gemination in various word positions.

9.2. Perception experiment

In the perception experiment, we investigated how variations in native language experience affect discrimination and categorisation of non-native gemination contrasts across word positions. Native speakers of Lebanese dialect and the Lebanese variety of Standard Arabic, Heritage speakers of Lebanese dialect, and Arabic-naïve English speakers were tested on perception of voiced and voiceless coronal stop gemination contrasts in the Moroccan variety of Standard Arabic and in Moroccan dialect. These listener groups differ in native experience with gemination in each word position. English listeners are not familiar with phonologically contrastive gemination in any word position, but do have purely phonetic-level experience with so-called “fake gemination”, which can occur when the same phoneme occurs

on either side of a word boundary (e.g., <top pick> vs. <topic>), akin to word-medial gemination. English speakers lack even phonetic-level experience with initial and final gemination. Lebanese speakers have native phonological experience with initial, medial and final gemination. Heritage Lebanese speakers have native experience with initial and medial gemination in the Lebanese dialect, but no experience with final gemination contrasts as they did not learn MSA. All the categorisation and discrimination results are summarised in Table 9.1.

Table 9.1: Contrast assimilation patterns (70% for categorisation and 25% for chance level) and their discrimination accuracy scores across medial, final and initial positions by the three groups of listeners (Lebanese, Heritage Lebanese and English). The first two letters (e.g., TC) refer to the assimilation type: TC: Two-Category, UC: Uncategorised-Categorised, UU: Uncategorised-Uncategorised, and the letter after the dash (e.g., N) refers to the degree of phonological overlap in assimilation: N: no overlap, P: partial overlap and C: complete overlap.

				Assimilation Type	Discrimination Accuracy scores
GROUP	Lebanese	Initial	voiced	TC-N	87%
			voiceless	UC-P	74%
		Medial	voiced	TC-N	94%
			voiceless	TC-N	96%
		Final	voiced	TC-N	92%
			voiceless	TC-N	90%
	Heritage Lebanese	Initial	voiced	TC-N	70%
			voiceless	UC-C	62%
		Medial	voiced	TC-N	88%
			voiceless	TC-N	92%
		Final	voiced	UC-N	79%
			voiceless	UC-P	69%
	English	Initial	voiced	UU-P	60%
			voiceless	UU-C	46%
		Medial	voiced	TC-N	77%
			voiceless	TC-N	82%
		Final	voiced	UU-P	71%
			voiceless	UU-C	58%

One central question behind this research was to know how phonological experience or phonetic-only experience with durational differences in a specific word position affects performance in the other word positions. As can be seen from Table 9.1, sensitivity to position-specific gemination distinctions does not guarantee success with categorising and discriminating the distinction in other word positions. This failure of carry-over across word

positions applies especially to the English group, but also somewhat to the Heritage group, who despite having shown good discrimination of medial gemination, were less accurate with the final contrast and most poorly with the initial contrast. The Lebanese group slightly outperformed the heritage group, especially with regards to the word-final contrast.

Unsurprisingly, word-medial contrasts were the easiest for all three groups of listeners to categorise and discriminate. This is very likely due to the fact that the three groups are familiar with gemination either phonologically for the Lebanese groups or phonetically for the English group in their native languages. In terms of discrimination, all three groups registered high discrimination accuracy scores: Lebanese (*M* = 95%), Heritage Lebanese (*M* = 90%) and English (*M* = 81%). These results were predicted by the categorisation results in light of PAM. All groups assimilated the medial contrasts as non-overlapping Two-Category (TC), whose discrimination is predicted to be excellent.

Word-final contrasts were more challenging to perceive than medial contrasts. The Lebanese group discriminated the final contrast the most accurately (*M* = 91%), significantly better than the Heritage Lebanese group (*M* = 74%), who in turn performed significantly more accurately than the English group (*M* = 65%). The lower discrimination scores for the Heritage group, and especially the English group, were predicted by the assimilation types found in the categorisation results. The Lebanese group was the only group that assimilated the final contrasts as non-overlapping TC, hence their significantly higher discrimination scores than the other two groups. The Heritage group assimilated the contrasts as Uncategorised-Categorised (UC), whose discrimination is expected to be lower than a TC. The English group categorised the final contrast as Uncategorised-Uncategorised assimilation type. Poorer discrimination is predicted for UU assimilation type than UC or TC assimilation types. Hence, they performed significantly lower than both Lebanese groups.

The initial voiceless contrast was by far the most challenging to perceive. Nonetheless, Lebanese group's success with its discrimination exceeded that of the Heritage Lebanese group and the English group (see Table 9.1). These results were also predicted by PAM. The Lebanese assimilated the initial voiceless contrast as their only UC partially-overlapping contrast. Similarly, the heritage group assimilated the initial voiceless contrast as UC too but with complete overlap. PAM-based research has shown that a completely-overlapping assimilation type is predicted to discriminated less accurately than the same assimilation type with no overlap or partial overlap (Best, Avesani, Tyler & Vayra, 2019). The English group's categorisation of the voiceless initial contrast was completely overlapping Uncategorised-Uncategorised (UU-C), which predicts very poor discrimination, given the assimilation type and the degree of overlap. Therefore, PAM's predictions were upheld.

In the following section, these results and their implications are discussed in relation to predictions and principles of the Perceptual Assimilation Model (PAM) theoretical framework (Best, 1995, Best & Tyler, 2007), as compared to the other main non-native speech perception models addressed in the introductory chapters.

9.3. Theoretical implications of the perceptual findings

The results from the perception experiment (Chapter 7) indicated that the listener groups were sensitive to acoustic differences between a geminate a singleton, particularly the Lebanese and Heritage Lebanese groups. This was evidenced by their high accuracy scores in discriminating the word-medial contrasts. Conversely, the three groups of listeners fared less successfully with the peripheral contrasts, particularly word-initial voiceless contrasts. Taken together, these results indicate that non-native listeners are influenced by native phonological and phonetic information that contributes to the accurate categorisation and discrimination of

non-native contrasts, while also being influenced by fine-grained native phonetic information, which results in reduced sensitivity to phonetic differences in the non-native contrast. PAM-based research proposed that native speech perception entails identifying phonetic differences that signal a lexical distinction (phonological distinctiveness), while being able to recognise irrelevant within-category phonetic details that do not signal a change in lexical item (phonological constancy) (Best, Tyler, Gooding, Orlando & Quann, 2009; Best, 2015). Developing phonological constancy contributes to detecting phonological contrasts with relative ease and efficiency. It appears, however, that decreased accuracy by listeners of geminating languages in categorising and discriminating the geminate/singleton contrast at the periphery of words may suggest, furthermore, that phonological constancy is context-dependent. This is very likely dependent on the varying functional load of the contrast at different word positions: in both MSA and dialectal Arabic of both Moroccan and Lebanese varieties, word-medial gemination is more productive and has a higher functional load than both initial and final gemination. This might have contributed to Lebanese and Heritage listeners' more developed phonological constancy for medial contrasts than contrasts at the peripheries of the word.

Lebanese and Heritage Lebanese groups performed better than the English group in discriminating and categorising the geminate/singleton contrasts at all word-position. This result is also accounted for in PAM's principles. PAM holds that non-native speech perception operates at an abstract phonological level as well as at a lower phonetic level. Thus, PAM assumes that there are both native phonological influences and phonetic influences, but that the phonological influences are stronger than lower-level phonetic influences. Since both Lebanese groups have phonological experience with gemination in their native language, faring better than the English group was no surprise, given that English listeners have phonetic-only experience with gemination.

Native Lebanese listeners discriminated the final contrast in MSA more accurately than Heritage Lebanese listeners did, and this result can also be explained in terms of PAM's principles. Both Lebanese groups have native phonological experience with initial and medial gemination in their vernacular dialect. However, final gemination is not contrastive in the Lebanese dialect. The Lebanese dialect lexicon contains only a very limited number of words, borrowed from MSA, with final geminated consonants. These words rarely have their singleton counterpart in the lexicon of Lebanese dialect. Thus, heritage Lebanese listeners should have a very limited experience with final gemination contrasts. Unlike native Lebanese listeners, Heritage Lebanese listeners do not know MSA, where final gemination carries a relatively high functional load. Since final gemination is structurally and phonetically different from medial and initial gemination (as we showed in Chapter 6 and 7), Heritage Lebanese listeners' extremely limited (or entirely lacking) experience with final gemination as a phonological contrast is reflected in their lower discrimination of final contrast compared to the Native Lebanese group. The Heritage group's superior performance relative to native English speakers, but poorer performance relative to the Native Lebanese group, may reflect sensitivity to only the phonetic differences between final geminates from singletons, because they lack the position-specific experience with gemination as a phonological contrast. This finding further highlights that experience with the abstract phonological and the phonetic characteristics of the native language can benefit the categorisation and discrimination of non-native phones, consistent with PAM principles.

Although PAM does not directly address positional effects on the perception of non-native contrasts, these findings are in accordance with PAM's principle that native versus non-native phonetic differences play an important role in the perception of non-native speech. Although gemination is not contrastive in English, the English group did categorise and discriminate the geminate/singleton contrast in word-medial position in Moroccan MSA

moderately well, though not as well as the Lebanese groups, who have native experience with phonological gemination. We attribute these results to the phonetic experience of English listeners with between-word gemination that is contextually comparable to word-medial gemination contrasts. Phonetic lengthening, which is often referred to as "fake gemination", exists in English, e.g., <top pick> vs <topic>, arising at the junction of identical consonants in suffixed words and compounds (Kotzor, Molineaux, Banks & Lahiri, 2016). Thus, it is entirely possible that the English native speakers perceived the disyllabic Arabic words as being like two monosyllabic words in medial position, one ending and the other beginning with the same consonant, akin to fake gemination in English. The English group's success with medial gemination in Moroccan Arabic suggests that non-native listeners can rely on native phonetic experience to perceive a consonant duration difference that is not phonologically contrastive in their language.

The advantage of native phonetic experience in supporting perception of non-native speech distinctions appears to be context-dependent, however. The context in which fake gemination occurs in languages like English is at word boundaries. Peripheral gemination does it even occur phonetically at utterance or word peripheries in English. This may be why the English listeners fared significantly worse with the final contrast, and were clearly unsuccessful with the initial voiceless contrast. It appears that reliance on native phonetic experience to perceive a non-native gemination contrast may be restricted to the context in which it is phonetically familiar.

Success with perceiving a familiar contrast in a non-familiar position is more likely to happen at word-final position than word-initial position. Although English listeners have no experience with gemination at word-final position, they discriminated the non-native Moroccan final contrast above chance, particularly the voiced contrast (M = 71%). Importantly,

their discrimination of the initial voiced contrast (*M* = 60%) was significantly lower than their discrimination of the final contrast. Similar results were reported in (Hallé, Ridouane & Best, 2016) for the discrimination of Berber initial versus final gemination by French naive listeners. This preference for the final position over the initial position seems to hold true even for perception of phonological contrasts that exist in the native language but not in a familiar context. Although /tl/ and /dl/ are permissible consonant combinations in French, they are not phonotactically allowed to occur in the initial position of the word. This results in French listeners failing at perceiving the initial /tl/ and /dl/ clusters of Hebrew (Hallé et al., 1998; Hallé & Best, 2007). Conversely, although Dutch disallows voiced obstruents in the final position of the word, Dutch listeners' performance in discriminating final voiced versus voiceless obstruents in nonwords exceeded that of native English listeners (Broersma, 2005). Thus, there seems to be an advantage for the final word-final position compared to word-initial position in facilitating non-native speech perception.

Theoretical debates have long existed between acoustic-phonetic accounts and phonological/articulatory-phonetic accounts of how native language experience affects non-native speech perception. One of the strengths of PAM that made it ideally suited to this thesis project was that it explicitly and systematically considers the relationship between native phonological and phonetic factors in non-native speech perception, rather than focusing nearly exclusively on acoustic-phonetic factors.

PAM posits that adult listeners perceptually assimilate non-native phones to native categories based on the perceived articulatory-phonetic similarities to their native language categories. The degree of perceived phonetic distance between the non-native phones and native phonemes is based on perceived articulatory resemblances between a foreign phone and the articulatory gestures used to produce the most similar native phoneme(s) (Best, 1995; Best

& Tyler, 2007). Since English has only singleton and not geminate consonants as native phonological categories, PAM would predict geminates to be categorised and discriminated less accurately than singletons. This prediction was confirmed by our findings. Not only did English listeners categorise geminates more accurately than singletons, their discrimination of the contrasts significantly improved when X in the AXB was a singleton. SLM principles do less well on this finding, in part because SLM focuses on production more than on perception, but more importantly because the durational ranges for geminate and singleton stops in the stimulus languages differ from those for English stops. Neither singletons nor geminates are phonetically equivalent to native English stops. Both geminate and singleton stops are therefore likely to be "similar" to the corresponding (voiced, voiceless) English stop and thus should be equally poorly categorised and discriminated.

Our perception results are consistent with PAM-based research reporting that perceived phonological overlap in L1 assimilations of contrasting non-native phones can decrease discrimination accuracy. *Non-overlapping* contrasts are better discriminated than *partially overlapping* contrasts, which in turn are better discriminated than *completely overlapping* contrasts (Best, Avesani, Tyler, & Vayra, 2019; Faris et al., 2018; So & Best, 2014). Those authors argued that listeners discriminate overlapping contrasts less accurately than non-overlapping contrasts because they perceive phonological similarity between the two contrasting phones with respect to the same native category(s). This perceived phonological constancy results in the poorer discrimination. Because the aforementioned studies investigated only vowel and consonant, it was important to know how perceived phonological overlap affects discrimination of gemination in Moroccan. We found that contrasts with larger overlaps between categorisation choices were discriminated more poorly than those with smaller overlaps. For example, the Heritage Lebanese group assimilated both final contrasts as a UC. The only difference was that a partial overlap was witnessed for the voiceless contrast, and no

overlap for the voiced contrast. This was also reflected in significantly lower performance in discriminating the final voiceless contrast compared to the final voiced contrast. Similarly, completely overlapping contrasts were discriminated more accurately than partially overlapping contrasts. For example, the English group discriminated voiced contrasts at the periphery of the word more accurately voiceless contrasts. This was due to the presence of complete overlap in categorising the word-initial and -final voiceless contrasts compared to their voiced counterparts. Considering contextual variations, our results showed that medial gemination was categorised with the least phonological overlap compared to final and initial gemination by all groups. Other factors are also likely to have influenced categorisation and discrimination results. First, although initial gemination exists in the native language of Heritage Lebanese listeners, the latter discriminated the initial contrasts significantly less accurately than native Lebanese listeners. This lower performance may be attributable to language pattern use. Australian Heritage listeners are bilingual native listeners of English and Lebanese Arabic dialect. However, due to more limited exposure to their heritage language, they are likely to be less proficient compared to native speakers from Lebanon. So, despite their heritage language being the first language they have learned, it is often regarded as the weaker language (Polinsky, 2018b). Second, the initial voiceless contrast was the most challenging for all three groups of listeners. Although our acoustic analysis showed that initial voiceless geminates are distinguished from singletons through other parameters, such as higher burst amplitude, these non-closure duration cues are clearly not perceptually relevant, thus they do not help the Lebanese listeners surmount the reduced audibility of the voiceless closure in utterance-initial position. Therefore, reduced acoustic cues to gemination are associated with lower discrimination scores. Third, previous research has shown that native listeners are capable of discriminating the geminate/singleton contrast in their native language with ceiling level accuracy even for peripheral contrast (Berber: Hallé, Ridouane & Best 2016). However,

our results showed that although they performed quite well with the Moroccan stimuli, neither Lebanese nor Heritage Lebanese listeners reached ceiling level. A possible explanation for this slightly lower performance is that the Moroccan stimuli were non-native to both groups. Thus, it is possible, indeed likely, that regional dialectal differences somewhat hindered their perception of Moroccan stimuli. However, it is important to note the thesis does not provide a definitive way to confirm these proposals and further research is warranted to compare between native and non-native perception of gemination as modulated by dialectal differences.

9.4. Methodological/analytical advances, limitations, and future directions

One of the methodological advances in this PhD project was the experimental design. Ultrasound and electromagnetic articulography (EMA) data were co-collected with acoustic data, and a subset of the acoustic recordings was then used in a perceptual study with three different groups of non-native listeners. Because we intended to use some of the acoustic tokens as stimuli in the perceptual experiment, our objective was to collect acoustic recordings of clear and unimpeded speech in the production experiment. Ultrasound is ideal for tracking the midsagittal shape of the tongue's upper surface without interfering in any way with speech articulation and thus without affecting the acoustic output, because the ultrasound probe is positioned outside the oral tract. Gluing sensors on the tongue for the EMA method distorts both articulation and the acoustic data, resulting in less than ideal acoustic stimuli for a perception experiment. In order to obtain and preserve clear, high-quality audio recordings of the geminate/singleton contrast, no EMA sensors were glued on the tongue. Sensors were instead placed behind the ears (left and right mastoid) and the upper incisor to allow head position correction of the articulatory data. Additional sensors were glued on the lower gum line and the lower and upper lips to obtain information jaw height, jaw rotation and lip shape, which we could not obtain with the ultrasound technique on its own. The EMA data on jaw

and lip motions during singleton and geminate coronal stop productions will be analysed for future reports.

To collect articulatory and acoustic data, we created an innovative and practical experimental setup. Co-collecting EMA and ultrasound data requires following strict guidelines and steps (see appendix A) before and during the recording session to ensure the data collection proceeds smoothly and without error. It is a common practice for the experimenter to require assistance from another lab colleague. However, collaboration with colleagues who have both availability and the requisite knowledge of technical procedures is not always possible. To avoid this, we created an experimental setup that allows an experimenter to conduct data collection without additional personnel support. This was achieved by increasing the number of monitors and placing them in a way that gives the experimenter more control over the whole setup (see Chapter 6 for details).

Another methodological advance was the highly innovative analytic method used to process the ultrasound data. By submitting the data to Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; Carignan, 2014) and Linear Discriminant Analysis (LDA) modelling, we were able to obtain a time-varying articulatory signal that helps us visualise the difference between geminates and singletons even in contexts that are distinguished by rather subtle and less robust acoustic cues, or even missing acoustic cues, i.e., at the word-periphery. Moreover, we also managed to obtain accurate values for articulatory closure duration using the velocity of the processed ultrasound signal. This was achieved by creating an articulatory closure duration estimation function using ultrasound, a kinematic technique that is relatively accessible to researchers that requires minimal manual input. Future research could develop further applications of this analytic approach to make the best use of ultrasound to investigate temporal research questions. By further refining and improving the

articulatory closure duration estimation function described above, it may be possible to make the generation of more accurate durational values for the geminate/singleton contrast fully automatic. Another line of research to explore is to create software that automatically recognises a segment's place of articulation directly from ultrasound images, which has been done using rtMRI (Hagedorn, Proctor & Goldstein, 2011). Ultrasound offers advantages over other kinematic technologies, such as rtMRI, in terms of portability, safety, non-invasiveness and relatively low cost. However, this technique's full power is far from being exploited at present.

One limitation of this PhD project concerned the number of words used in the production experiment and the number of target contrasts in the perception experiment. Because our objective was to compare Moroccan and Lebanese in how their native speakers produce gemination, the words in the production experiments were selected carefully with the aim of controlling possible factors that could affect the phonetic realisation of the contrast. This allowed us to infer information about the phonetic-phonological relationships. However, given the linguistic differences among the two varieties of Modern Standard Arabic (MSA) and the two vernacular dialects, the task of finding words that met our stimulus requirements and were also meaningful words across all varieties was not an easy one. This limitation was mitigated by recording nonce words that are also comparable to the lexical items. Moving forward, we have plans to supplement the study with more acoustic data by recording a larger range of words, reflecting a wider selection of consonant types, with the same speakers if possible.

Another limitation of this project was the frames per second obtained from ultrasound video recording (30 fps). Increasing the frame rate would result in more accurate temporal measurement of the articulatory results. However, this was not possible with the ultrasound system we used, due to the limitation of the screen-capture hardware. Newer ultrasound

machines that can record image data without using screen-capture devices should be used in future to allow for substantially higher frame rate.

We investigated in this thesis the production of gemination by native speakers and the perception of gemination by non-native listeners. This study did not investigate the relationship between non-native perception and non-native production of gemination. Future research should test how L2 listeners produce non-native contrast and compare the results to their performance discriminating and categorising the same non-native contrasts. Findings are likely to have important implications for PAM and other models of non-native speech perception and L2 learning. Moreover, some studies have tested L2 learners' perception of gemination. They attested to the role of experience in learning L2 on improved perception and production of the geminate/singleton (e.g., Hayes-Harb & Masuda, 2008). However, most of those studies have been conducted on medial gemination only. There is a clear gap in the literature when it comes to the perception and production of peripheral gemination by L2 learners.

9.5. Conclusions

In conclusion, this PhD thesis provided novel insights based on a comprehensive articulatory, acoustic and perceptual investigation of gemination in MSA and dialectal Arabic as produced by native Arabic speakers from two geographically distant regions of the Arab world: Morocco and Lebanon. Acoustic and articulatory results suggested that peripheral gemination is maintained in Moroccan and Lebanese, contrary to previous claims of neutralisation. Importantly, the results showed that gemination behaved differently from one utterance/word position to another and from one dialect to the other, with regard to its phonetic realisation.

Positional differences were also witnessed in the perception results of three groups of listeners. Success with discriminating the medial, initial and final gemination contrasts in the Moroccan target stimuli reflected the groups' varying phonological experience, i.e., Native Lebanese vs. Heritage Lebanese groups, and phonetic-only experience, i.e., English group, with gemination and the different word and utterance positions. The perceptual findings have important implications for PAM, whose scope had not been applied to address positional effects on the perception of gemination: native phonological experience and native phonetic experience are context-dependent, but phonological experience is more likely to aid non-native gemination perception than purely phonetic experience. Moreover, reduced acoustic cues, dialectal differences, language use patterns and voicing are all factors influencing the categorisation and discrimination of non-native gemination contrast. This range of factors stress the significance of considering low-level phonetic alongside phonological effects on non-native perception of gemination.

In terms of production, this was the first major work to demonstrate that peripheral gemination is maintained in Moroccan and Lebanese Arabic. Furthermore, results showed that the role of closure duration was a more robust cue to gemination utterance-medially, for word-medial gemination and for initial gemination in utterance-medial position. In word-final and absolute-initial positions, closure duration differences are less robust, but other non-closure duration cues are more enhanced than in word/utterance medial position. Although these cues are associated with a forceful articulation, we proposed that his articulatory force is influenced more by prosodic factors than by gemination alone. Secondly, in terms of dialectal differences, the two dialects weight the articulatory and acoustic cues differently. In Moroccan, durational cues are weighted more heavily than non-durational cues, whereas Lebanese shows the opposite weighting. In summary, this thesis' findings constitute important additions to the body

of research on the production of gemination, particularly the typologically rare peripheral geminate/singleton contrasts.

10. REFERENCES

Abdul-Karim, K. (1979). *Aspects of the Phonology of Lebanese Arabic* (Doctoral Dissertation). University of Illinois, Urbana-Champaign.

Abramson, A. S. (1986). The perception of word-initial consonant length: Pattani Malay. *Journal of the International Phonetic Association*, *16*(01), 8-16.

Abramson, A. S. (1987). Word-initial consonant length in Pattani Malay. In T. Gamkrelidze (Ed.), *Proceedings of the 11th International Congres*.

Abramson, A. S. (1991). Amplitude as a cue to word-initial consonant length: Pattani Malay. *In Proceedings of the 12th 15th International Congress of the Phonetic Sciences* , Aix-en-Provence, August 1991, 98–101.

Abramson, A. S. (1998). The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length. In U. Warotamasikkhadit & T. Panakul (Eds.), *Papers from the 4th Annual Meeting of the Southeast Asian Linguistics Society 1994* (pp. 1-20). Tempe, Arizona: Monograph Series Press.

Abramson, A. S. (1999). Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. In J. J. Ohala (Ed.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 591-594). Berkeley: University of California Press.

Abramson, A. S. (2003). Acoustic cues to word-initial stop length in Pattani Malay. In M.J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 387-390). Barcelona: Universitat Autònoma de Barcelona.

Abu-Abbas, K. H., Zuraiq, W. M., & Abdel-Ghafer, O. A. (2011). Geminates and long consonants in Jordanian Arabic. *International journal of linguistics*, *3*(1), 1-17.

Adank, P., Evans, B.G. , Stuart-Smith, J. , & Scott, S.K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. Journal of Experimental Psychology: Human Perception and Performance, 35, 520-529.

Ahmed, M. S. (2016). Difficulties faced by British and Iraqi undergraduate students of English when pronouncing Arabic true geminates and English fake geminates. *Alustath*, *1*(217), 1-20.

Albirini, A., & Benmamoun, E. (2014). Concatenative and nonconcatenative plural formation in L1, L2, and heritage speakers of Arabic. *The Modern Language Journal*, *98*(3), 854-871.

Al-Deaibes, M., & Rosen, N. (2019). Gemination in Rural Jordanian Arabic. *XXX Perspectives on Arabic Linguistics*, 53-76.

Alfozan, A. I. (1989). *Assimilation in classical Arabic: a phonological study* (Doctoral dissertation). University of Glasgow, Scotland.

Al-Kahtany, A. H. (1997). The'problem'of diglossia in the Arab world: An attitudinal study of modern standard Arabic and the Arabic dialects. *al-'Arabiyya*, 1-30.

Al-Tamimi, Y., & I'lawi, N. (2006). Variation in Northern Jordanian Arabic: Manifestations of Standard Origin. *Journal of Language and Linguistics*, *5*(2), 230-245.

Al-Tamimi, F., Abu-Abbas, K., & Tarawnah, R. (2010). Jordanian Arabic final geminates: An experimental clinical phonetic study. *Poznań Studies in Contemporary Linguistics*, *46*(2), 111-125.

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* (pp. 212-215). Hong Kong: City University of Hong Kong.

Al-Tamimi, J., & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, *71*, 306-325.

Altoma, S. J. (1969). *The problem of diglossia in Arabic*. Harvard U. Press.

Amer, F. H., Adaileh, B. A., & Rakhieh, B. A. (2011). Arabic diglossia. *Argumentum*, *7*, 19-36.

Antoniou, M. (2010). *One Head, Two Languages: Speech Production and Perception in Greek–English Bilinguals* (PhD dissertation). University of Western Sydney, Australia.

Aoyama, K., & Reid, L. A. (2006). Cross-linguistic tendencies and durational contrasts in geminate consonants: An examination of Guinaang Bontok geminates. *Journal of the International Phonetic Association, 36*(2), 145-157.

Arvaniti, A., & Tserdanelis, G. (2000). On the phonetics of geminates: Evidence from Cypriot Greek. In B. Yuan, T. Huang, & X. Tang (Eds.), *Proceedings of the 6th International Conference on Spoken Language Processing (Interspeech 2000)* (pp. 559-562). Beijing: International Speech Communication Association.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412.

Baese-Berk, M. (2018). Perceptual learning for native and non-native speech. *Current Topics in Language*, *68*, 1-29.

Bergmann, P. (2014). Reduction and deletion of glottal stops and geminates at pword-boundaries in German–effects of frequency and accentuation. *Syllable and word languages*, *40*, 251-278

Blanc, H. (1953). *Studies in North Palestinian Arabic: Linguistic inquiries among the Druzes of Western Galilee and Mt. Carmel*. Jerusalem: Oriental Notes and Studies, Israel Oriental Society, 4.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK:Cambridge University Press.

Blevins, J. (2008). Explaining diversity in geminate consonant inventories: An evolutionary approach. In *Lecture delivered during Leipzig Spring School on Linguistic Diversity. Max Planck Institute for Evolutionary Anthropology: Leipzig, Germany.*

Beckman, M. (1986). *Stress and non-stress accent*. Dordrecht: Foris Publications.

Belazi, H. B. H. (1984). *Diglossia and the situation of Arabic in Tunisia* (Doctoral dissertation, Cornell University).

Bell-Berti, F., Gelfer, C., Boyle, M., & Chevrie-Muller, C. (1991). Speech Timing in Ataxic Dysarthia. In *Proceedings of the 12th International Congress of Phonetic Sciences* (Vol. 5, pp. 262-265). Aix-en-Provence: University of Provence.

Berkovits, R. (1991). The effect of speaking rate on evidence for utterance-final lengthening. *Phonetica*, *48*(1), 57-66.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, *167*(224), 233-277.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross language research* (pp. 171-204). Baltimore: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. The Journal of the Acoustical Society of America, 109(2), 775–794.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (Vol. 10389, pp. 13-34). Amsterdam: John Benjamins.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C.A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science, 20*, 539-542. doi: 10.1111/j.1467-9280.2009.02327

Best, C. T., Avesani, C., Tyler, M. D., & Vayra, M. (2019). PAM revisits the Articulatory Organ Hypothesis: Italians' perception of English anterior and Nuu-Chah-Nulth posterior voiceless fricatives. In A. M. Nyvad, M.Hejná, A. Højen, A. B. Jespersen & M. H. Sørensen [Eds.] *A Sound Approach to Language Matters: In Honor of Ocke-Schwen Bohn.* Pp. 13-40. Aarhus University Press: Aarhus DK.

Bohn, O. S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, *14*(2), 131-158.

Bombien, L., Mooshammer, C., & Hoole, P. (2013). Articulatory coordination in word-initial clusters of German. *Journal of Phonetics*, *41*(6), 546-561.

Bouarourou, F., Vaxelaire, B., Ridouane, R., Hirsch, F., Sock, R. (2008). Gemination in Tarifit Berber: X-ray and acoustic data. In R. Sock, S. Fuchs, & Y. Laprie (Eds.), *Proceedings of the 8th International Seminar on Speech Production* (pp. 117-120). Strasbourg: The National Institute for Research in Computer Science and Control.

Bouarouroua, F., Vaxelaire, B., Ridouane, R., Hirsch, F., Fautha, C., & Sock, R. (2011). Gemination in Tarifit Berber: Doing one or two things at once? In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 344-347). Hong Kong: City University of Hong Kong.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *The Journal of the Acoustical Society of America*, *117*(6), 3890-3901.

Boersma, P., & Weenink, D. (2007). Praat (Version 4.5. 25)[Software]. *Latest version available from www. praat. org. URL (accessed 19 November 2010): http://www. fon. hum. uva. nl/praat*.

Burgel, R., & Faehndrich, M. (2005). The effect of mora-timing on the duration of vowels preceding geminate consonants. In *Working Paper in Linguistics*, *36*(3), 1-20.

Carignan, C. (2014). TRACTUS (Temporally Resolved Articulatory Configuration Tracking of UltraSound) software suite. URL: http://christophercarignan.github.io/TRACTUS. DOI: 10.13140/RG.2.2.27085.03047

Carr, P. (2008). *Glossary of phonology*. Edinburgh: Edinburgh University Press.

Chang, C. B. (2014). Transfer effects in perception of a familiar and unfamiliar language. *Journal of the Acoustical Society of America*, *135*(4), 2355.

Chang, C. B. (2016). Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition*, *19*(4), 791-809.

Christodoulou, C. (2007). Phonetic geminates in Cypriot Greek: the case of voiceless plosives. In *Eighth Annual Conference of the International Speech Communication Association*.

Cho, T., & Jun, S.-A. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society*, *36*(1), 31-44.

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*, 155-190.

Church, K. (2012). *Phonological parsing in speech recognition* (Vol. 38). Springer Science & Business Media.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers, 25, 257–271.

Cohn, A. C., Ham, W. H., & Podesva, R. J. (1999). The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. *In Proceedings of the 14th 15th International Congress of the Phonetic Sciences* (pp.587-590).

Cowell, M. W. (2005). *A Reference Grammar of Syrian Arabic with Audio CD:Based on the Dialect of Damascus*. Georgetown University Press.

Curtis, E. K. J. (2003). *Geminate weight: Case studies and formal models* (Doctoral dissertation). University of Washington, Seattle.

Davis, S. (2011). Geminates. *The Blackwell companion to phonology*, 1-25.

Derrick, D. , Best, C. T. , and Fiasson, R. (2015). Non-metallic ultrasound probe holder for co-collection and co-registration with EMA. in *Proceedings of the 18th International Congress of Phonetic Sciences(ICPhS)*, 1–5

Derrick, D., Carignan, C., Chen, W., Muayiwath, S., & Best, C. T. (2018). 3D Printable ultrasound transducer stabilizer. *Journal of the Acoustical Society of America Express Letters*, 144, EL392. https://doi.org/10.1121/1.5066350

Diehm, E. E. (1998). *Gestures and linguistic function in learning Russian: Production and perception studies of Russian palatalized consonants* (Doctoral dissertation, The Ohio State University).

Dmitrieva, O. (2012). *Geminate typology and the perception of consonant duration* (Unpublished doctoral dissertation). Stanford University.

Dmitrieva, O. (2017). Production of geminate consonants in Russian: Implications for Typology. In H. Kubozono (Ed.), *The Phonetics and Phonology of Geminate Consonants* (pp. 34–65). Oxford: Oxford University Press.

Dmitrieva, O. (2018). The role of perception in the typology of geminate consonants: Effects of manner of articulation, segmental environment, position, and stress. *Language and Speech*, 61(1), 43-70.

Doty, C. S., Idemaru, K., & Guion, S. G. (2007). Singleton and geminate stops in Finnish Acoustic correlates. In H. van Hamme & R. van Son (Eds.), *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)* (pp. 2737-2740). Antwerp: Curran Associates, Inc.

Dufour, S., Brunellière, A., & Nguyen, N. (2013). To what extent do we hear phonemic contrasts in a non-native regional variety? Tracking the dynamics of perceptual processing with EEG. *Journal of psycholinguistic research*, *42*(2), 161-173.

Dunn, M. H. (1993). *The Phonetics and Phonology of Geminate Consonants: A Production Study* (UMI Dissertation Services, Ann Arbor, MI).

El Saaran, M. 1951. *A critical study of the phonetic observations of the Arab grammarians*. (Unpublished PhD dissertation, University of London.)

Engstrand, O., & Krull, D. (1994). Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica*, *51*(1-3), 80-91.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorisation*. Netherlands Graduate School of Linguistics.

Falahati, R. (2013). *Gradient and categorical consonant cluster simplification in Persian: An ultrasound and acoustic study* (PhD Dissertation, University of Ottawa, Ottawa).

Faris, M. M., Best, C. T., & Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorised. The Journal of the Acoustical Society of America, 139(1), EL1-5. https://doi.org/10.1121/1.4939608

Ferguson, C. A. (1959). Diglossia. *Word*, *15*(2), 325-340.

FFmpeg Development Team, 2016. FFmpeg. Computer software program available from https://ffmpeg.org .

Fintoft, K. (1961). The duration of some Norwegian speech sounds. *Phonetica*, *7*(1), 19-39.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*(6), 3728-3740.

Fougeron, C. (1999). Prosodically conditioned articulatory variations: A review. *University of California, Los Angeles Working Papers in Phonetics*, *97*, 1-73.

Fougeron, C., Kuehnert, B., Imperio, M., & Vallee, N. (2010). Laboratory phonology. Mouton de Gruyter, Berlin.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. Studies in Second Language Acquisition, 26(1), 1–34. doi:10.1017/S0272263104261010

Flege, J. E. (2005). Origins and development of the Speech Learning Model. *Retrieved December*, *13*, 2005.

Frej, M. Y., Carignan, C., & Best, C. T. (2017). Acoustics and articulation of medial versus final coronal stop gemination contrasts in Moroccan Arabic. (F. Lacerda, S. Strombergsson, M. Wlodarczak, M. Heldner, J. Gustafson, & D. House, Eds.)*Proceedings of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH 2017), August 20-24, 2017, Stockholm, Sweden*, 210-214. doi:10.21437/Interspeech.2017-1292.

Frej, M. Y., Carignan, C., Proctor, M., & Best, C. (2018). Acoustic and ultrasound investigation of word-initial gemination in Moroccan Arabic. 144(3), 1871-1871.

Fujimoto, M., & Shinohara, S. (2015). Articulation of voiced geminate and word-initial geminate obstruents in the Miyakojima Ikema dialect of Ryukyuan: A Real-time MRI analysis. In *International Workshop on Geminate Consonants (GemCon) 2015*.

Galea, L., Grice, M., Hermes, A., & Mücke, D. (2014). Lexical and surface geminates in Maltese. Proc.10th ISSP, 128-133.

Galea, L. (2016). *Syllable structure and gemination in Maltese*(Doctoral dissertation, University of Cologne).

Galea, L., & Ussishkin, A. (2018). Onset clusters, syllable structure and syllabification in Maltese. *The languages of Malta*, *18*, 55.

Ghalib, G. B. M. (1984). *An experimental study of consonant gemination in Iraqi colloquial Arabic* (Doctoral dissertation, University of Leeds).

Ghosh, A. (2015). Acoustic Correlates of Voicing and Gemination in Bangla. In *Proceedings of the 12th International Conference on Natural Language Processing* (pp. 413-418).

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". Neuropsychologia, 9 (3): 317–323, doi:10.1016/0028-3932(71)90027-3.

Griffen, T. D. (1990). Degemination strategies in new High German and Swabian. *Lingua*, 82, 131-149.

Hagedorn, C., Proctor, M., & Goldstein, L. (2011). Automatic analysis of singleton and geminate consonant articulation using real-time magnetic resonance imaging. In *Twelfth Annual Conference of the International Speech Communication Association*.

Hallé, P. A., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation?. Journal of experimental psychology: Human perception and performance, 24(2), 592.

Hallé, P. A., & Best, C. T. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/clusters. The Journal of the Acoustical Society of America, 121(5), 2899-2914.

Hallé, P. A., Ridouane, R., & Best, C. T. (2016). Differential Difficulties in Perception of Tashlhiyt Berber Consonant Quantity Contrasts by Native Tashlhiyt Listeners vs. Berber-Naïve French Listeners. *Frontiers in psychology*, *7*.

Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. The Journal of the Acoustical Society of America, 110(1), 489-503.

Harrell, R. S. (2004). *A short reference grammar of Moroccan Arabic: With audio CD*. Georgetown University Press.

Hall, N. (2013). Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics*, *41*(2), 133-143.

Ham, W. (1998). *Phonetic and phonological aspects of geminate timing*. Doctoral dissertation, Cornell University.

Ham, W. (2001). *Phonetic and phonological aspects of geminate timing*. New York: Routledge.

Ham, W. (2013). *Phonetic and phonological aspects of geminate timing*. Routledge.

Hamdi, R., Barkat-Defradas, M., Ferragne, E., & Pellegrino, F. (2004). Speech Timing and Rhythmic structure in Arabic dialects: a comparison of two approaches. In P*roceeding of InterSpeech*.

Hamzah, H. (2010). Durational properties of initial geminate consonants in Kelantan Malay. In M. Tabain, J. Fletcher, D. Grayden, J. Hajek, & A. Butcher (Eds.), *Proceedings of the 13th Australasian International Conference on Speech Science and Technology* (pp. 18-21). Melbourne: Australian Speech Science and Technology Association.

Hamzah, H., Fletcher, J., & Hajek, J. (2011). Durational correlates of word-initial voiceless geminate stops: The case of Kelantan Malay. In *Proceeding of the 17th 15th International Congress of the Phonetic Sciences* (pp.89-92).

Hamzah, M. H., Fletcher, J., & Hajek, J. (2012). An acoustic analysis of release burst amplitude in the Kelantan Malay singleton/geminate stop contrast. In: *Proceedings of the 14th Australasian international conference on speech science and technology* (pp. 85–88).

Hamzah, H. (2013). The acoustics and perception of the word-initial singleton/geminate contrast in Kelantan Malay (Unpublished doctoral dissertation). The University of Melbourne.

Han, M. (1994). Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America*, 96, 73-82.

Hankamer, J., Lahiri, A., & Koreman, J. (1989). Perception of consonant length: Voiceless stops in Turkish and Bengali. *Journal of Phonetics*, *17*(4), 283-298.

Hansen, B. B. (2004). Production of Persian geminate stops: Effects of varying speaking rate. In *Proceedings of the 2003 Texas linguistics society conference* (pp. 86-95). Somerville, MA: Cascadilla Proceedings Project.

Harris, Z. S. (1942). *The phonemes of Moroccan Arabic. Journal of the American oriental society*, 62(4), 309-318.

Harris, K. T. J. (2014). A Functional-Load Account of Geminate Contrastiveness: a Meta-Study. *Linguistics*.

Hayes, R. L. (2001). The perception of novel phoneme contrasts in a second language: A developmental study of native speakers of English learning Japanese singleton and geminate consonant contrasts. *Coyote Papers: Working Papers in Linguistics, Language in Cognitive Science*.

Hayes-Harb, R. (2005). Optimal L2 speech perception: Native speakers of English and Japanese consonant length contrasts. *Journal of Language and Linguistics*, *4*(1), 1-29.

Hayes-Harb, R., & Masuda, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research*, 24(1), 5-33. doi:10.1177/0267658307082980.

Heath, J. (1987). *Ablaut and Ambiguity: Phonology of a Moroccan Arabic Dialect*. Albany: State University of New York Press

Hedia, S. B. (2019). *Gemination and degemination in English affixation: Investigating the interplay between morphology, phonology and phonetics* (Vol. 8). Language Science Press.

Hirata, Y., & Whiton, J. (2005). Effects of speaking rate on the single/geminate stop distinction in Japanese. *The Journal of the Acoustical Society of America*, *118*(3), 1647-1660.

Hirst, D. (2011). The analysis by synthesis of speech melody: from data to models. *Journal of speech Sciences*, *1*(1), 55-83.

Hisagi, M., & Strange, W. (2011). Perception of Japanese temporally-cued contrasts by American English listeners. *Language and Speech*, *54*(2), 241-264.

Hoberman, R. D., & Aronoff, M. (2003). The verbal morphology of Maltese. *Language Acquisition and Language Disorders*, *28*, 61-78.

Hueber, Thomas, Guido Aversano, G. Cholle, Bruce Denby, Gérard Dreyfus, Yacine Oussar, Pierre Roussel, and Maureen Stone. "Eigentongue feature extraction for an ultrasound-based silent speech interface." In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 1, pp. I-1245. IEEE, 2007.

Huggins, A. W. F. (1972). Just noticeable differences for segment duration in natural speech. *The Journal of the Acoustical Society of America*, *51*(4B), 1270-1278.

Idemaru, K. (2005). *An acoustic and perceptual investigation of the geminate and singleton stop contrast in Japanese* (PhD thesis). University of Oregon, Oregon.

Idemaru, K., & Guion, S. G. (2008). Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association*, *38*(2), 167-186.

Jakobson, R., & Waugh, L. R. (1979). *The Sound Shape of Language*. Bloomington: Indiana University Press.

Shaw, J., Best, C. T., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., & Mulak, K. E. (2018). Resilience of English vowel perception across regional accent variation. *Laboratory Phonology*, *9*(1).

Jessen, M. (1999). *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam & Philadelphia:John Benjamins Publishing.

Kasatkin, L. L., & Choj, M. C. (1999). *Dolgota/kratkost' Soglasnogo na Meste Sochetanij Dvuh Bukv v Sovremennom Russkom Literatunom Jazike*. [Duration of the doubly-spelled consonants in the Contemporary Standard Russian]. Dialog-MGU, Moskva.

Kawahara, S. (2005). Voicing and geminacy in Japanese: An acoustic and perceptual study. *University of Massachusetts Occasional Papers in Linguistics, 31*. 87-120.

Kawahara, S. (2012). Amplitude changes facilitate categorization and discrimination of length contrasts. *The Institute of Electronics, Information and Communication Engineers Technical Report*, *112*, 67-72.

Kawahara, S. (2015). Geminate devoicing in Japanese loanwords: Theoretical and experimental investigations. *Language and Linguistics Compass*, *9*(4), 181-195.

Khattab, G. (2007). A phonetic study of gemination in Lebanese Arabic. In *Proceeding of 16th International Congress of The Phonetic Sciences in Sarrebrücken*, 153-158.

Khattab, G., & Al-Tamimi, J. (2008). Durational cues for gemination in Lebanese Arabic. *Language and Linguistics, 22*, 39-55.

Kim, J.-Y. (2019). Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism: Language and Cognition*, 1-18. DOI: 10.1017/S1366728918001220

Kiparsky, P. (2003). Syllables and moras in Arabic. *The syllable in optimality theory*, 147-182.

Klatt, D. H., & Cooper, W. E. (1975). Perception of segment duration in sentence contexts. In A. Cohen & S. Nooteboom (Eds.), *Structure and process in speech perception* (pp. 69-89). Heidelberg: Springer.

Kochetov, A., Sreedevi, N., Kasim, M. & Manjula, R. (2012). A pilot ultrasound study of Kannada lingual articulations. *Journal of Indian Speech and Hearing Association*, 26, 38–49

Kohler, K. J. (1984). Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica*, *41*(3), 150-174.

Kotzor, S., Molineaux, B. J., Banks, E., & Lahiri, A. (2016). "Fake" gemination in suffixed words and compounds in English and German. The Journal of the Acoustical Society of America, 140(1), 356-367.

Kraehenmann, A., Fitzpatrick, J., Lahiri, A., & Mohanan, T. (2000). Word-initial geminates in Swiss German and Malayalam: Production and perception. Poster presented at the Laboratory Phonology 7, Nijmegen.

Kraehenmann, A. (2001). Swiss German stops: Geminates all over the word. *Phonology*, *18*(1), 109-145.

Kraehenmann, A., & Lahiri, A. (2008). Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops. *The Journal of the Acoustical Society of America*, *123*(6), 4446-4455.

Kupisch, T., Einfeldt, M., Lleó, C., & van de Weijer, J. (2016). Voice Onset Time and gemination in Italian heritage speakers. In *Proc EUROSLA.*

Kuhl, P. K., & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect". *Speech perception and linguistic experience: Issues in cross-language research*, 121-154.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, *82*(13), 1-26.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lahiri, A. & J. Hankamer. (1988). The timing of geminate consonants. *Journal of Phonetics.16*(3), 327-338.

Lee-Ellis, S. (2012). *Looking into Bilingualism through the heritage speaker's mind* (Doctoral dissertation).

Lehtonen, J. (1970). Aspects of Quantity in Standard Finnish. *Studia Philologica Jyväskyläensia VI*, Jyväskylä.

Letterman, R. (1994). Nominal gemination in Sinhala and its implications for the status of prenasalized stops. In *Annual Meeting of the Lingusitic Society of America, Boston*.

Lin, I. (2019). *Functional Load, Perception, and the Learning of Phonological Alternations* (Doctoral dissertation, UCLA).

Local, J., & Simpson, A. P. (1999). Phonetic implementation of geminates in Malayalam nouns. *Word*, *4*(92), 46.

Löfqvist, A. (2005). Lip kinematics in long and short stop and fricative consonants. *The Journal of the Acoustical Society of America*, *117*(2), 858-878.

Löfqvist, A. (2006). Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese. *The Journal of the Acoustical Society of America*, *120*(5), 2872-2883.

Löfqvist A. (2007). Tongue movement kinematics in long and short Japanese consonants. *Journal of the Acoustical Society of America*, 1(122), 512–518.

Lunden, S. L. (2006). *Weight, final lengthening and stress: A phonetic and phonological case study of Norwegian*. University of California: Santa Cruz.

Maddieson, I. (2018). Phonetics and African languages In T. Güldemann (Ed.), The Languages and Linguistics of Africa (pp. 546-601). Berlin, Germany: Walter de Gruyter & Co.

Magyar, L. (2014). Learnability of word-final gemination in loan monosyllables. *Squib*, 981(24).

Malécot, A. (1968). The Force of Articulation of American Stops and Fricatives as a Function of Position, *Phonetica*, *18*, 95–102.

Matthews, P. H. (2014). *The concise Oxford dictionary of linguistics*. Oxford University Press: Oxford.

Marçais, W. (1930). La diglossie: un pélérinage aux sources. *Bulletin de la Société Linguistique de Paris*, *76*(1), 61-98.

Markus, E. (2010). Primary and secondary geminates in Ingrian. *Linguistica Uralica*, *46*(1), 38-52.

McCrary, K. (2002). Syllable structure vs. segmental phonotactics: The phonetics and phonology of segment duration in Italian. *Proceedings of International Phonology Meeting*. Vienna, Austria.

Mikuteit, S., & Reetz, H. (2007). Caught in the ACT: The timing of aspiration and voicing in East Bengali. *Language and speech*, *50*(2), 247-277.

Mitchell, T. F. (1993). *Pronouncing Arabic* (Vol. 2). Oxford University Press.

Moisik, S. R., Lin, H., & Esling, J. H. (2014). A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetic Association*, *44*(1), 21-58.

Mooshammer, C., & Fuchs, S. (2002). Stress distinction in German: simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics*, *30*(3), 337-355.

Mooshammer, C., Hoole, P., & Geumann, A. (2006). Interarticulator cohesion within coronal consonant production. *The Journal of the acoustical society of America*, *120*(2), 1028-1039.

Muller, J. S. (2001). *The phonology and phonetics of word-initial geminates* (Unpublished doctoral dissertation). The Ohio State University.

Muller, J. S. (2003). The production and perception of word-initial geminates in Cypriot Greek. In M. J. Sole , D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of the Phonetic Sciences* (pp. 1867-1870). Barcelona: Universitat Autnoma de Barcelona.

Oh, G. E., & Redford, M. A. (2012). The production and phonetic representation of fake geminates in English. *Journal of Phonetics*, *40*, 82-91.

Ohala, M. (2007). Experimental methods in the study of Hindi geminate consonants. In M. J. Sole, P. Beddor, & M. Ohala (Eds.), *Experimental approaches to phonology* (pp. 351–368). Oxford: Oxford University Press.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English, *Journal of the Acoustical Society of America*, 54, 1235- 1247.

Padgett, J. (2003). *Systemic contrast and Catalan rhotics*. Santa Cruz: University of California.

Pajak, B. (2009). Perception of Moroccan Arabic geminates by native English speakers. *The Journal of the Acoustical Society of America*. 125. 2778. 10.1121/1.4784777.

Pajak, B. (2010). Contextual constraints on geminates: The case of Polish. In N. Baier & E. Donnelly (Eds.), *Proceedings of the 35th Annual Meeting of the Berkeley Linguistics Society (BLS)*, 269–280, Berkeley: University of California, Berkeley.

Pajak, B. (2013). Non-intervocalic Geminates: Typology, Acoustics, Perceptibility. *San Diego linguistics papers*, 4, 2-27.

Payne, E. (2005). Phonetic variation in Italian consonant gemination. *Journal of International Phonetic Association, 35*(2), 153–181.

Payne, E. (2006). Non-durational indices in Italian geminate consonants. *Journal of International Phonetic Association, 36*(1), 83–95.

Payne, E., Post, B., Garmann, N. G., & Simonsen, H. G. (2017). The acquisition of long consonants in Norwegian. *The Phonetics and Phonology of Geminate Consonants*, *2*, 130-162.

Percival, M., Kochetov, A., & Kang, Y. (2018). An Ultrasound Study of Gemination in Coronal Stops in Eastern Oromo. In *INTERSPEECH* (pp. 1531-1535).

Pickett, J. M., & Decker, L. R. (1960). Time factors in perception of a double consonant. *Language and Speech*, *3*(1), 11-17.

Pickett, E., Blumstein, S., & Burton, M. (1999). Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica*, *56*, 135–157.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D.R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody* (pp. 90-117). Cambridge: Cambridge University Press.

Pind, J. (1995). Speaking rate, voice-onset time, and quantity: The search for higher-order invariants for two Icelandic speech cues. *Perception & Psychophysics*, *57*(3), 291-304.

Pind, J. (1999). Speech segment durations and quantity in Icelandic. *Journal of the Acoustical Society of America*, *106*(2), 1045-1053.

Polgárdi, K. (2005). Geminates and degemination in Hungarian. *In International Conference on the Structure of Hungarian* Veszprém, Hungary.

Polinsky, M. (2018b). *Heritage languages and their speakers*. **Cambridge: Cam**bridge University Press.

Palo, P., Schaeffler, S., & Scobbie, J. M. (2014). Pre-speech tongue movements recorded with ultrasound. In *Proceedings of the 10th International Seminar in Speech Production (ISSP 10)*, 300-303.

Port, R. F., Dalby, J., & O'Dell, M. (1987). Evidence for mora timing in Japanese. *The Journal of the Acoustical Society of America*, *81*(5), 1574-1585.

Ridouane, R. (2003). Geminates vs. singleton stops in Berber: An acoustic, fiberscopic and photoglottographic study. In *Proceedings of the 15th International Conference on Phonetic Sciences*, 1743-1746.

Ridouane, R. (2007). Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *Journal of the International Phonetic Association*, *37(*2), 119-142.

Ridouane, R. (2010). Geminates at the junction of phonetics and phonology. *Papers in laboratory phonology*, *10*, 61-90.

Ridouane, R., & Turco, G. (2019). Why is gemination contrast prevalently binary? Insights from Moroccan Arabic. *Radical: A journal of Phonology*, *1*, 62-91.

Rinaldi, N. (2014). *Perception and Production of Geminate Consonants by English Speakers and Kuwaiti Arabic Speakers* (Doctoral dissertation), University of Oregon: Oregon.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, *13*(2), 155-163.

Shrotriya, N., Verma, R., Gupta, S. K., & Agrawal, S. S. (1996). Durational characteristics of Hindi consonant clusters. In K. Elenius & P. Branderud (Eds.), *Proceedings of the 13th International Congress of Phonetic Sciences*, 1-4. Stockholm: Royal Institute of Technology.

Slis, I. H. (1971). Articulatory effort and its durational and electromyographic correlates. *Phonetica*, *23*(3), 171-188.

Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. *Invariance and variability in speech processes*, *1*, 426-449.

Takada, S. (1985). Sokuon-no chooon joo-no tokuchoo ni tsuite [articulatory characteristics of geminates]. *Kokuritsu Kokugo Kenkyuujo Houkoku 83 Kenkyuu Hookokushuu*, *6*, 17-40.

Taylor, M. (1985). Some patterns of geminate consonants. *The University of Chicago working papers in linguistics*, *1*, 120-129.

Thurgood, G. (1993). Geminates: a cross-linguistic examination. *Papers in honor of Frederick H. Brengelman on the occasion of the twenty-fifth anniversary of the Department of Linguistics,* CSU Fresno: Fresno, 129-139.

Tomaschek, F., Arnold, D., Bröker, F., & Baayen, R. H. (2018). Lexical frequency co-determines the speed-curvature relation in articulation. *Journal of phonetics*, *68*, 103-116.

Toole, F. E., & Olive, S. (1988). The modification of timbre by resonances: Perception and measurements. *Journal of Audio Engineering Society, 36*(3), 122-142.

Topintzi, N., & Davis, S. (2017). On the weight of edge geminates. *The phonetics and phonology of geminate consonants*, *2*, 260-282.

Tserdanelis, G., & Arvaniti, A. (2001). The acoustic characteristics of geminate consonants in Cypriot Greek. In *Proceedings of the fourth international conference on Greek linguistics*, 29-36.

Tsukada, K., & Hajek, J. (2019). Cross-language perception of Italian and Japanese consonant length contrasts: A comparison of native Italian listeners with and without Japanese language learning experience. In *Proceedings of the International Congress of Phonetic Sciences*, Melbourne, The University of Melbourne.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. Phonetica, 71(1), 4–21. https://doi.org/10.

Whalen, D. H., Iskarous, K., Tiede, M. K., Ostry, D. J., Lehnert-LeHouillier, H., Vatikiotis-Bateson, E., & Hailey, D. S. (2005). The Haskins optically corrected ultrasound system (HOCUS). *Journal of Speech, Language, and Hearing Research*.

Watson, J. C. (2002). *The phonology and morphology of Arabic*. Oxford University Press on Demand: Oxford.

Watson, J. (2007). Syllabification patterns in Arabic dialects: Long segments and mora sharing. *Phonology*, 24(2), 335-356.

APPENDICES

APPENDIX A: RECORDING SESSION STEPS AND GUIDELINES

Ultrasound, EMA and acoustic data collection

1) Before the experiment:

EMA System

-Untangle the sensors (11 sensors)
To minimize strain on the wires, you should start with the protractor that is already connected and by the end of testing session; you should switch to the other protractor
-Disinfect the sensors before attaching them to the participant
-Turn on the sensor boxes of the EMA system
-Turn on the monitoring computer
-Open the NDI WaveFront software
-Load the probe reference file (wave60probe)
-Change the directory and the session name
-Click reference none to be able to test the sensors
-Test the sensors: 1) check the sensor number 2) hold it about 15cm away from the center of the EMA projector long enough to make sure it stays solid green
-Do a quick test run to make sure that the synchronization works
-After testing the sensors and the synchronization, it would be best to shut down the NDI WaveFront software until it is time to start recording.
-Before setting up the ultrasound system, make sure you have all the tools (twisers, gauzes, tongue depressors, tape) and the glues on the cart ready to go.

Ultrasound

-Before running the ultrasound system, you need to shut down any Internet connection as well as any anti-virus software
To do that on a windows computer, right click on my computer, click on manage, then under services and applications, click on services, scroll down the list of all the programs running and stop the anti-virus system (i.e, Sophos)
-Link the ultrasound system and microphone to the computer
-Plug the ultrasound system to the power source
-Plug the ultrasound probe to the ultrasound system
-Have the bat script and ffmpeg program in one directory.

2) During the experiment

-Explain to the participant what they will be doing during the recording and ask them if they are ready.
-Have the participant sit in a chair with the NDI wave EMA projector placed to the right side of their head.

-Strap the non-metallic probe holder on the participant's head.
The strap should be tight enough but should not make the participant uncomfortable.
For some participants, spacers are needed to be inserted at both sides of the jaw platform to have the probe container placed parallel to the larynx. The probe container should not be touching the larynx. The spacers make it move forward a bit.
If the spacers are inserted, the strap should be loosened up a little bit.
-Cut some small pieces of scotch tape.
We use scotch tape to prevent some sensors from falling (mastoid sensors…)

EMA system setup

-Wear some gloves before handling any sensor
Sensor 1: right mastoid
Sensor 2: left mastoid
(apply some dental wax before attaching sensors 3 and 4)
Sensor 3: upper incisor
Sensor 4: lower incisor
-Ask the participant to stick their tongue as far as they can
-Use a marker to mark the positions of sensors to be glued on the tongue
-Use gauzes to dry to the tongue before gluing any sensor
Sensor 5: tongue dorsum
Sensor 6: tongue mid
Sensor 7: tongue tip
Sensor 8: right parasagittal
Sensor 9: left parasagittal
Sensor 10: upper lip
Sensor 11: lower lip
-Mix the powder glue using 1 drop of liquid with 1spoon2/3 of glue powder.
-Apply the glue on the sensors attached on tongue.
-Use the air blower to fast dry the glue.
-After all the sensors are glued, open the NDI software to check whether all the sensors are well centred. Also, make sure no sensor is flashing.

Ultrasound system setup

-Put enough gel on the probe before inserting in the probe container.
-Make sure the audio recording is working.
-Adjust the volume to avoid clipping
-Adjust the width and the depth for better visualization.
-Start recording with the ultrasound system first, then EMA system.
-Have the participant say TA TA TA at the beginning of the data collection and at the end.
-Before stopping the recording, take some recordings of the bite-plane and the palate with both systems.

-For the palate trace with ultrasound, have the participant drink some water and instruct them to drink slowly as if they are drinking from a straw.

-For the occlusal plane with ultrasound, have the participant hold a tongue depressor between their teeth. Instruct them to stick their tongue underneath and press up.
-Stop ultrasound recording and make sure you save the data.

-For the palate trace with EMA, start from as far back in the soft palate and trace along all the way to the front teeth and then trace back and stop there (3 times).

-For the occlusal plane with EMA, unscrew the sensors of the protractor that is on the head rest and replace them with the sensors of the bite plane protractor.
We swap out the protractor that is on the head rest with the bite plane protractor because of the lack of sensor ports.

Sensor 12: occlusal right
Sensor 13: occlusal front
Sensor 14: occlusal left

-Make sure you disinfect the protractor.
-Have the participant bite on the protractor.
-Ask them to move their head up and down, left and right and then in a circle a couple of times while recording.

3) After the experiment

-Remove the probe and put aside in a safe place.
-Use a tissue to wipe off the remaining gel from the probe
-Take the bite plane protractor from the participant
-Put on some gloves and start taking the sensors off the participant.
-Ask the participant if they are more comfortable taking off some of the sensors on the tongue
-Take off the head-mounted probe holder
-Put the sensors away before the participant stands up
-Clean the sensors to be used for next recording session
-Before shutting down the EMA program, make sure you export the data
-Use an external disc to save the data
-Before leaving the lab, make sure you clean the tile on which you mixed the powder glue.

Pilot Checklist

	Day before
	Connect and check all sensors
	Check audio
	Check supplier • Glue • Taps • Tongue depressors
	Check forms & files • Information sheet • Consent form • Background questionnaire • Protocol
	Day of experiment
	Take the glue out of the freezer
	Prepare: • Paper towels • Tongue depressors • Mouthwash
	Check audio
	Check sensors
	Setup file to record experiment protocol
	Participant arrives
	Welcome participant
	Explain entire procedure: sensors, glues
	Participants reads information sheet and signs consent form
	Participant brushes teeth and tongue
	Start gluing
	Beginning recording
	Finish experiment
	Clean-up

APPENDIX B: INFORMATION SHEETS, RECRUITMENT FLYERS, CONSENT FORMS, LANGUAGE BACKGROUND

Participant Information Sheet:
Production

Project Title: The production and perception of peripheral geminate/singleton coronal contrasts in Arabic

Project Summary: The research seeks to advance our knowledge and understanding of regional differences in Arabic, and of why second language (L2) learners have difficulty with certain non-native Arabic consonants.

You are invited to participate in a research study being conducted by PhD student Yassine Frej under the Supervision of Prof Catherine Best, Chair of Psycholinguistics Research at the MARCS Institute for Brain, Behavior and Development, Western Sydney University.

What will I be asked to do?
You will be asked to say some simple words/non-sense words and sentences in Modern Standard Arabic and your respective dialect of Arabic. Your jaw, lip and head movements will be tracked by gluing small electromagnetometry (EMA) sensors to your teeth and/or facial skin using a non-toxic adhesive, which has been approved for use both on the skin and inside the mouth. The sensors will be brand new, and will not have been previously used for any other people. You will also have your tongue movements recorded by ultrasound while you speak. Ultrasound, which is often used to see internal body structures such as the heart (Echocardiography), has no known side effects associated with it, and it causes no physical discomfort. The ultrasound probe covered with ultrasound gel will be placed under your chin and held in place using a specially-designed headset. The jaw, lip and tongue movements data will be recorded to computer, and your voice will be recorded onto computer. Later, these recordings will be analysed by the researchers, and the recordings might also be used as target stimuli in subsequent perception experiments with Australian listeners and heritage speakers of Arabic. Furthermore, they may also be used in future experiments on production and perception of Arabic varieties.

How much of my time will I need to give?
The study will take approximately two hours of your time

What specific benefits will I receive for participating?
You will receive a compensation of $75/hour for your time. You will also gain experience with how a psycholinguistic experiment is run. The results will contribute to our understanding of how speakers of different varieties of a language produce the same consonants. They will also contribute to our understanding of how people perceive non-native sounds.

Will the study involve any discomfort for me? If so, what will you do to rectify it?
Speech production studies involving EMA require tiny wired sensors to be attached to areas in or around the speaker's mouth. For our project, EMA sensors will be glued to the lower lip, upper lip, the corners of the mouth, the upper and lower teeth in order to obtain information about the jaw height and the lip shape as well as head movement. All available scientific evidence indicates that the risks of the EMA procedure and research equipment are

exceedingly small. Because the procedure takes a short time to complete, any low risk of health hazards that might be linked to long-term magnetic field exposure are not likely to apply. All the tiny sensors and the instruments used to attach them to the lips and teeth are properly sterilized and pose no risk to health beyond the levels of risk associated with normal daily activities. Also, because the electromagnetometer uses electrical safety features commonly found in equipment designed for hospital use, there is no significant risk of electrical shock

Similarly, ultrasound, which is often used medically to visualise internal body structures such as the heart (Echocardiography), has no known side effects associated with it, and it causes no physical discomfort. Ultrasound involves no radiation exposure. Although there is no proof of risks of diagnostic ultrasound in pregnant women, pregnant women may wish to exclude themselves from this study to avoid the discomfort of being seated for a long period.

How do you intend to publish the results?

The findings of the research will be presented in a PhD thesis authored by Yassine Frej. They will also be published in journal articles and/or book chapters, and presented at scientific conferences by Yassine Frej and the rest of the research team.

Can I withdraw from the study?

Participation is entirely voluntary and you are not obliged to be involved. If you do participate, you can withdraw at any time without giving a reason.

If you do choose to withdraw, any information that you have supplied will be destroyed at your request..

Can I tell other people about the study?

Yes, you can tell other people about the study by providing them with Yassine Frej's contact details: y.frej@westernsydney.edu.au. They can contact him to discuss their participation in the research project and obtain an information sheet.

What if I require further information?

Please contact Prof Catherine Best should you wish to discuss the research further before deciding whether or not to participate.

Catherine Best
Chair of Psycholinguistics Research, the MARCS Institute
(02) 9772 6760

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through the Office of Research Services on Tel +61 2 4736 0229 Fax +61 2 4736 0013 or email humanethics@westernsydney.edu.au.

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

If you agree to participate in this study, you will be asked to sign the Participant Consent Form

Participant Information Sheet:
Perception

Project Title: The production and perception of peripheral geminate/singleton coronal contrasts in Arabic

Project Summary: The research seeks to advance our knowledge and understanding of regional differences in Arabic, and of why second language (L2) learners have difficulty with certain non-native Arabic consonants

You are invited to participate in a research study being conducted by PhD student Yassine Frej under the supervision of Prof Catherine Best, Chair of Psycholinguistics Research at the MARCS Institute for Brain, Behavior and Development, Western Sydney University.

What will I be asked to do?
You will be asked to complete two listening tasks where you will hear utterances through the headphones and make judgments about them. In the first task, you will hear three utterances on each of many trials, and must decide if the second utterance is the same as the first utterance or the same as the third utterance. In the second task, you will only hear one utterance on each of many trials, but you will hear the same utterance two times on each trial. The first time you hear it, your task will be to identify either the first or the last consonant by clicking on one of the on-screen options with the mouse. After you have made a selection, you will hear exactly the same utterance again. This time, you will rate how good a version the sound is of the category that you just chose on a scale from 1 to 7. Finally, you will complete a survey that will ask questions detailing your language use and demographic information.

How much of my time will I need to give?
The study will take approximately one hour of your time.

What specific benefits will I receive for participating?
You will receive compensation of $20/hour for your time. You will gain experience with how a psycholinguistic experiment is run. The results will contribute to our understanding of how humans perceive non-native sounds.

Will the study involve any discomfort for me? If so, what will you do to rectify it?
The study is very unlikely to cause discomfort for you, and poses no levels of harm greater than everyday life activities.

How do you intend to publish the results
The findings of the research will be presented in a PhD thesis authored by Yassine Frej. They will also be published in journal articles and/or book chapters, and presented at scientific conferences by Yassine Frej and the rest of the research team.

*Please note that the minimum retention period for data collection is five years past publication of the findings.

Can I withdraw from the study?
Participation is entirely voluntary and you are not obliged to be involved. If you do participate, you can withdraw at any time without giving a reason.

If you do choose to withdraw, any information that you have supplied will be destroyed at your request.

Yes, you can tell other people about the study by providing them with Yassine Frej's contact details: y.frej@westernsydney.edu.au. They can contact him to discuss their participation in the research project and obtain an information sheet.

What if I require further information?
Please contact Prof Catherine Best should you wish to discuss the research further before deciding whether or not to participate.

Catherine Best
Chair of Psycholinguistics Research, the MARCS Institute
(02) 9772 6760

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through the Office of Research Services on Tel +61 2 4736 0229 Fax +61 2 4736 0013 or email humanethics@westernsydney.edu.au.

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

If you agree to participate in this study, you will be asked to sign the Participant Consent Form.

WESTERN SYDNEY
UNIVERSITY

Consent Form: Production study

Project Title: The production and perception of peripheral geminate/singleton coronal contrasts in Arabic

I hereby consent to participate in the above named research project.

I acknowledge that:

2 I have read the participant information sheet (or where appropriate, have had it read to me) and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s

3 The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.

I consent to:

[Insert tick box option for each specific activity e.g.

& *Participate in a production experiment*

& *Read aloud from a word list, and a list of sentences*

& *Have your jaw and lip movements recorded by an electromagnetic articulometry (EMA) machine while you speak. Small EMA sensors will be adhered to your teeth and/or skin using a non-toxic adhesive, which has been approved for use both on the skin and inside the mouth.*

& *Have my tongue movements recorded by an ultrasound machine while I speak. The ultrasound probe covered with ultrasound gel will be placed under my chin and held in place using a specially-designed headset.*

& *I may be asked to drink water and to hold a tongue depressor between my teeth (not at the same time) as part of this process, in order to allow the ultrasound machine to create an image of the top of your mouth and your bite plane.*

& *Fill out a questionnaire about my language background and related information.*

I consent for my data and information provided to be used in this project and other related projects for an extended period of time.

I understand that my involvement is confidential and that the information gained during the study may be published and stored for other research use but no information about me will be used in any way that reveals my identity.
Signed:

Name:

Date:

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through Research Engagement, Development and Innovation (REDI) on Tel +61 2 4736 0229 or email humanethics@westernsydney.edu.au.

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Consent Form: Production study

Project Title: The production and perception of peripheral geminate/singleton coronal contrasts in Arabic

I hereby consent to participate in the above named research project.

I acknowledge that:

4 I have read the participant information sheet (or where appropriate, have had it read to me) and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s

5 The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.

I consent to:

[Insert tick box option for each specific activity e.g.

& *Participate in a perception experiment*

& *Listen to spoken words and press buttons to categorize and rate the stimuli.*

& *Fill out a questionnaire about my language background and related information.*

I consent for my data and information provided to be used in this project and other related projects for an extended period of time.

I understand that my involvement is confidential and that the information gained during the study may be published and stored for other research use but no information about me will be used in any way that reveals my identity.

I understand that I can withdraw from the study at any time without affecting my relationship with the researcher/s, and any organisations involved, now or in the future.

Signed:

Name:

Date:

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through Research Engagement, Development and Innovation (REDI) on Tel +61 2 4736 0229 or email humanethics@westernsydney.edu.au.

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Seeking native Arabic speakers from Moroccan/Lebanese origins for a speech production task

You are invited to take part in a study about how Modern Standard Arabic and its dialects are produced. To be eligible, you will need to be:

1) A native Arabic speaker who was born in Morocco and came to live/study/work in Australia as a teenager or adult

   OR

2) A native Arabic speaker who was born in Lebanon and came to live/study/work in Australia as a teenager or adult
3)

If you choose to take part in this study, you will say some simple words and sentences in Modern Standard Arabic and your respective dialect. Your lip, jaw and movements will be tracked and recorded to computer, and your voice will be recorded onto computer and audio-tape.
To compensate for your time, you will be paid $75 per hour. This experiment will

take up to 2 hours. The experiment is conducted at Western Sydney University

(Bankstown Campus).

Project Title: The production and perception of peripheral geminate/singleton coronal contrasts in Arabic.
Ethics Approval ID: H11941

For additional information or to register your interest in participating, please email Yassine at y.frej@westernsydney.edu.au

Participation is completely voluntary

Best Regards,

Mr. Yassine Frej Prof Catherine Best Dr Michael Tyler

Seeking Australian English speakers

You are invited to take part in a study of how Arabic sounds are perceived by English speakers who do not know any Arabic. To be eligible, you will need to:

1) Be born and raised in Australia and speak English only

If you decide to participate, we will ask you to come in for a single visit where we will ask you to listen to passages and answer questions about what you heard. The visit will take approximately 1 hour, and you will be compensated $20.

The experiment is being conducted at Western Sydney University.

Project Title: The production and perception of peripheral geminate/singleton coronal contrasts in Arabic.

Ethics Approval ID: H11941

For additional information or to register your interest in participating, please email Yassine at y.frej@westernsydney.edu.au

Thank you for your interest,

Mr. Yassine Frej (Primary researcher)
Prof Catherine Best (co-investigator)
Dr Michael Tyler (co-investigator)

Participant Identification code _____ Date: _____

NOTE: Information collected on this form is kept confidential, not associated with your name.

LANGUAGE BACKGROUND

1. Date of Birth _____

2. What is the highest level of education you have completed (for example, 'completed 2nd year of high school' or 'enrolled at university now, have completed 2 years')?

_____

3. Do/did you have a ***hearing*** impairment, or ***reading*** difficulties (e.g., difficulties learning to read), or ***language development or speaking*** difficulties (e.g., delayed language development, stuttering, lisping, etc.)?

NO_____

YES_____

If yes, tick any/all that apply: hearing ___ reading ___ speaking ___ language ____

4. What is your native language, i.e., the first language you learned, from birth?_____.

What is your father's native language? _____

What is your mother's native language?_____

5. Have you had formal education in your native language?

Yes No

6. If yes, where and how long?

_____

7. Where did your father grow up (country, city/town/region)? _____

Where did your mother grow up ?_____

8. Please tell us every place you have lived for at least three months or more, starting with your infancy and childhood. Please be specific: if you lived in a city, please also indicate which area or town/suburb.

Country City/Town/Region between which ages?

1. (birthplace) _____

2. _____

3. _____

4. _____

5. _____

9. Were there any other people from a different country (especially Morocco or Lebanon), who lived at your house or frequently spent time with you (for example, a grandmother watched over you while your parents were at work, a housekeeper, a nanny or a roommate)?

If so, who were they, what country were they from, and how much time roughly did you spend with them?

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

10. Have you been to Morocco or Lebanon?

Yes No

11. If yes, where and how long?

_____

11. Please tell us what *other* languages you speak besides English, and when and how long you have studied/learned each language, and how well you speak each one.

	Language	Years Spoken/ & When	How well do you speak it? (1=hardly at all, 5=highly fluent)?
Example:	*Spanish*	*2 yrs (high school)*	*3 (intermediate fluency)*
1.			
2.			
3.			
4.			

APPENDIX C: CONFERENCE PRESENTATIONS

Estimation of closure duration for absolute word-initial geminate/singleton coronal stops in Moroccan Arabic using ultrasound

[1]Mohamed Yassine Frej, [2]Christopher Caringnan, [3]Catherine T. Best, and [4]Michael Proctor

[1,3] The MARCS Institute, Western Sydney University, Australia.

[2] Institute of Phonetics and Speech Processing (Ludwig-Maximilians-Universität München)

[4] Department of Linguistics, ARC Centre of Excellence in Cognition and its Disorders, Macquarie University

The most robust and universal cue distinguishing geminates from singletons is constriction duration (i.e. closure duration for stops) [1]. However, measuring closure duration acoustically can be a challenge particularly for initial voiceless stops in utterance-initial position (absolute initial) due to the relative lack of information in the acoustic signal. Voiced stops are no exception in some cases; the closure of initial geminated voiced stops can be fully devoiced [2]. Articulatory investigations of gemination utterance-initially have employed techniques including electropalatography [3] and oral airflow [1] to measure this closure duration. However, to our best knowledge, there has been no ultrasound investigation of word-initial gemination. In Moroccan Arabic vernacular (MA), word-initial gemination is highly functional. It results from a morphophonological process whereby the definite article /l/ totally assimilates with word-initial coronals making the latter a geminate. Nonetheless, there is no evidence that it is maintained in utterance-initial position. Position in the utterance has been shown to affect the realization of word-initial gemination. Initial geminates were found to be neutralized when stressed utterance-initially in Cypriot Greek [4], maintained but weakened in utterance-initial position compared to utterance-medial position in Kelantan Malay [5], or conversely more enhanced utterance-initially than –medially in in Tashlhiyt Berber [3]. The current study presents a semi-automatic method to estimate closure duration for word-initial geminate/singleton voiceless and voiced stop contrasts directly from ultrasound images, in order to address key theoretical questions about initial gemination in MA. Specifically, whether MA speakers maintain or neutralize the contrast utterance-initially, and if so, what are its articulatory properties and how does position in the utterance condition these properties?

In order to minimize manual processing, ultrasound data acquired from five Moroccan Arabic speakers were analysed through the Temporally Resolved Articulatory Configuration Tracking of Ultrasound (TRACTUS; [6]), which filters and submits ultrasound images to a PCA model; LDA discrimination of the PCs can be used to generate a time-varying articulatory signal that clearly captures the closure difference between geminates and singletons word-initially, even in the absence of acoustic closure duration information (Figure 1) (for details on how to interpret these signals see [7]).

The method we employed to estimate duration values of the closure is akin to that followed by some studies using EMA (e.g. [8]) and rtMRI (e.g. [9]) to determine gestural landmarks from the velocity signal. In this study, after identifying the maximum positive velocity peak (the gestural onset) and maximum negative velocity peak (the gestural offset) for each token, 20% thresholds of these peaks were calculated. The interval of image frames between these gestural landmarks was identified, and stop closure duration was estimated as the duration of the interval (200 ms; Figure 2).

The results show that MA initial geminates are characterized by a closure duration 1.32 times longer (M = 333ms; s.d. 44ms) than singletons (M = 252ms; s.d. 41ms) utterance-initially. These duration differences as well as the articulatory signal generated by TRACTUS offer evidence that initial gemination in MA is maintained articulatorily, rather than neutralized. The calculated duration values are substantially higher than those we found for utterance-medial

data with geminates (*M* = 199ms; s.d. 37ms) being 2.36 longer than singletons (*M* = 84ms; s.d. 29ms). This suggests that speakers might be using a closure-lengthening strategy to enhance the contrast utterance-initially. However, the differences in duration ratio (1.32 vs 2.36) indicates that the difference between geminates and singletons is proportionally smaller utterance-initially than –medially. Interestingly, the same positional effects were found in Swiss German initial gemination [10].

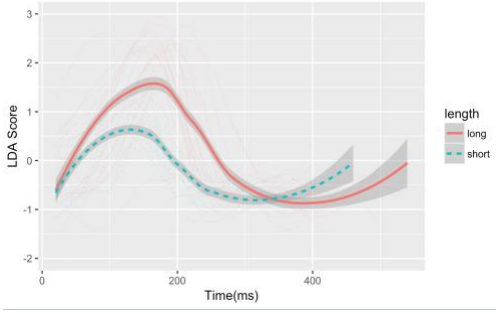

Figure 1: *The LDA class scores over time for geminates and singletons in word-initial position*

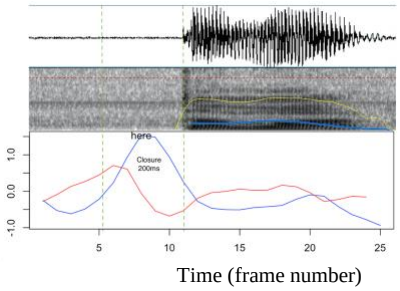

Figure 2: ***Estimation of closure duration of word-intial voiceless stop in [taman]****: articulatory signal (blue line), velocity signal (red line) and time-*

References

[1] Ridouane, R. 2003. Suites de Consonnes en Berbère chleuh: Phonétique et Phonologie. PhD Dissertation. Sorbonne Nouvelle University, Paris 3.
[2] Taylor, M. (1985). Some patterns of geminate consonants. *The University of Chicago working papers in linguistics*, 1, 120–129.
[3] Ridouane, R. 2007. Gemination in Tashlhiyt Berber : an acoustic and articulatory study. *Journal of the International Phonetic Association.* 37 (2), 119-142.
[4] Armostis, S. 2012. An articulatory study of word-initial stop gemination in Cypriot Greek. *In 10th International Conference of Greek Linguistics*.
[5] Hamzah, M. H., Fletcher, J., & Hajek, J. 2016. Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay. *Journal of Phonetics*, *58*, 135-151.
[6] Carignan, C. 2014. TRACTUS (Temporally Resolved Articulatory Configuration Tracking of UltraSound) software suite, URL: http://christophercarignan.github.io/TRACTUS.
[7] Frej, M. Y., Carignan, C., & Best, C. T. 2017. Acoustics and articulation of medial versus final coronal stop gemination contrasts in Moroccan Arabic. *Proc. Interspeech 2017*, 210-214.
[8] Bauer, D., Kannampuzha, J., Hoole, P., & Kröger, B. 2010. Gesture duration and articulator velocity in plosive-vowel-transitions. *Development of Multimodal Interfaces: Active Listening and Synchrony*, 346-353.
[9] Proctor, M., Lammert, A., Katsamanis, A., Goldstein, L., Hagedorn, C., & Narayanan, S. 2011. Direct estimation of articulatory kinematics from real-time magnetic resonance image sequences. In *12 ISCA.*
[10] Kraehenmann, A., and Lahiri, A. 2008. Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops. *Journal of the Acoustical Society of America*. 123(6): 4446–4455.

Acoustic and ultrasound analysis of the intervocalic geminate/singleton coronal contrasts in Moroccan Arabic

Mohamed Yassine Frej[1], Catherine T. Best[2], Christopher Carignan[3]

[1,2,3]*The MARCS Institute, Western Sydney University*

[1]y.frej@westernsydney.edu.au, [2]C.Best@westernsydney.edu.au,
[3]C.Carignan@westernsydney.edu.au

This research carries out a simultaneous acoustic and ultrasound investigation of the intervocalic geminate/singleton coronal contrasts in Moroccan Arabic (MA). The aim of this paper was to demonstrate that temporal differences in the acoustic signal are captured in the articulatory signal using ultrasound, which is a rather novel technique in indexing gemination. One Female (44 years) and one male (40 years) MA speakers produced, 10 times in a carrier phrase and isolated form, 4 real minimal pairs contrasting t and d with tt and dd. Acoustic data was co-registered with ultrasound data (30fps). Acoustic analysis reveals that the participants produced the geminated stops with longer closure duration than their singleton counterparts regardless of their manner of articulation. Although voiceless stops were realized with longer durations than voiced stops, the geminate/singleton duration ratio for voiced stops was larger than the ratio for the voiceless stops, 2.64 /1 versus 2.15/1. In our data, the duration of vowels adjacent to the geminate/singleton contrast was not affected significantly. This is in line with a number of studies, which didn't report a temporal compensation between the consonant and the surrounding vowels (e.g., Zeroual, 2008).

Articulatory investigations of gemination have employed a variety of techniques to determine the articulatory characteristics of gemination. However, to our knowledge, no articulatory investigation of the geminate/singleton contrasts has been conducted using ultrasound. The latter is an ideal technique because of its affordability, safety, and smaller effect on speech articulation and acoustics compared to other kinematic techniques. To avoid the rather tedious and time-consuming task of manually tracing tongue contours, our ultrasound data were analyzed using the Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; Carignan, 2014), which is a suite of Matlab functions designed to submit the pixel intensities of ultrasound images to a PCA model in order to identify a set of principal components (PCs) (94 PCs used for our data) that explain 80% of image variance throughout the ultrasound video recording. An LDA model was then trained on the PC scores for ultrasound frames associated with the consonant closures throughout the recording. The cumulative products of the PC scores and PC loadings were used to generate composite "heatmaps" for both singletons and geminates, which can be interpreted as average models of the lingual configuration for the two consonant types (Figure 1 and 2).

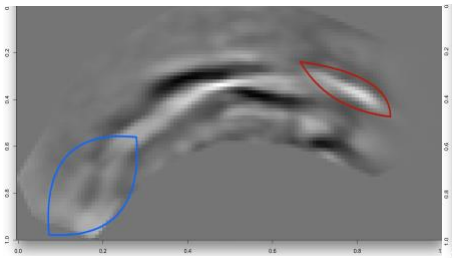

Figure 1: LDA Heatmap for singletons

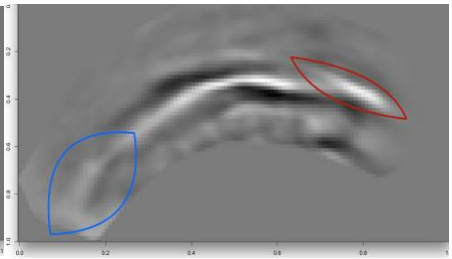

Figure 2: LDA Heatmap for geminates

A close observation of both heatmaps suggests that the gestures responsible for the production of both contrasts are quite comparable with the exception of a noticeable raising of the back of the tongue at the laryngeal area for geminates (see area marked in blue in figure 2). Another interesting observation is that constriction location of the coronal stops in MA is laminal contrary to what was generally assumed (Zeroual, 2007) (see area marked in red). Further investigation is warranted, as these results are only preliminary.

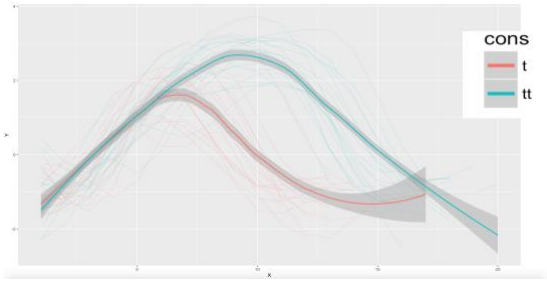

Figure 3: Medial “t” & ”tt” subdata

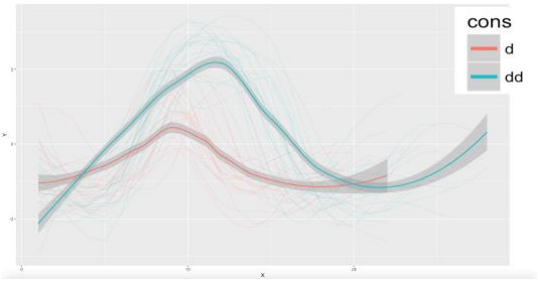

Figure 4: Medial “d” & ”dd” subdata.

Figures 3 and 4 display the articulatory signals for individual tokens, as well as category means and 95% confidence intervals generated using t-based approximation (“loess” smoothing in the *ggplot2* package in R) (Wickham, 2009). Preliminary results suggest that the articulatory signals generated by TRACTUS are capable of capturing the difference between a geminate and a singleton in the word-medial position. The same analysis will be performed at a later stage on the word-initial data in order to demonstrate that even in the absence of clear acoustic information, temporal analysis of gemination in this position of the word is possible through ultrasound and TRACTUS.

References:

Carignan, C. (2014). TRACTUS (Temporally Resolved Articulatory Configuration Tracking of UltraSound) software suite. URL: http://christophercarignan.github.io/TRACTUS.

Ghalib, G. (1984). *An experimental study of consonant. gemination in Iraqi Colloquial Arabic* (Unpublished doctoral dissertation). University of Leeds.

Ridouane, R. (2007). Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *Journal of the International Phonetic Association*, 37(2), 119-142.

Ridouane, R. (2010). Geminates at the junction of phonetics and phonology. Papers in laboratory phonology, 10, 61-90.

Zeroual, C., Hoole, P., Fuchs, S., & Esling, H. (2007). EMA study of the coronal emphatic and non-emphatic plosive consonants of Moroccan Arabic. In *Proc. 16th ICPhS*.

Wickham, H. (2009). Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. URL: http://ggplot2.org.