The production of emphatic fricatives in spoken Arabic dialects

Mahmoud Alsabhi

Doctor of Philosophy

University of York

Language and Linguistic Science

April 2024

Abstract

The current thesis aims at investigating the variations of the emphatic fricative contrast s~sˤ across different spoken Arabic dialects. To achieve this aim, three main studies were conducted, acoustic, perceptual, and articulatory. In terms of the acoustics, eight Arabic dialects from the IVAr corpus were utilised to examine if the differences between the emphatic contrast s~sˤ is the same or different across these eight dialects. Multiple acoustic correlates were measured for the consonants and vowel. The findings suggest that the measures that exhibited significant information to mark the emphatic contrast s~sˤ are observed on the vowel formant information; F1 tends to rise, whilst F2 is consistently lower in the context of an emphatic fricative compared to non-emphatic. The GAMM results reveal that the majority of the emphatic contrast was observable at different points in the trajectory of the vowel following the emphatic contrast s~sˤ, and dialects differ in how large the F2 difference is and how much of the vowel is affected. The findings also reveal that other variables, such as COG and fricative duration, can somehow display difference in the emphatic contrast, though these differences are statistically insignificant.

Drawing from these outcomes, the subsequent perceptual study employed a subset of the acoustic data. F2 trajectories were manipulated based on the results from three Arabic dialects, resulting in real and manipulated stimuli. Results indicate that, in real stimuli, all listeners from each dialect could identify their own native formant cues with higher accuracy in plain condition. Crucially however, all listeners can attend to the cues to emphasis of other dialects with considerable accuracy. However, for manipulated stimuli, the F2 size and trajectory shape was observed to play a significant role for listeners to be able to identify their own cues, in emphatic condition (only). This disparity between real and manipulated stimuli suggests F2 may not be the sole cue relied upon by listeners.

Subsequently, an Ultrasound imaging study was conducted, involving fifteen speakers from five dialects, to investigate any articulatory basis for this phenomenon and whether tongue position and lip movement would distinguish the emphatic contrast. Principal Component Analysis (PCA) results reveal that tongue positioning significantly contributes to distinguishing between emphatic/non-emphatic fricatives, tongue body is retracted during emphasis and advanced during non-emphatic consonants, while there was no association of lip movement observed during the articulation of the emphatic contrast except for four speakers.

List of Contents

List of Figures

List of Tables

List of Abbreviations

egca	Egyptian	**F0**	Fundamental frequency
syda	Syrian	**F1**	First formant frequency
joka	Jordanian	**F2**	Second formant frequency
irba	Iraqi	**F3**	Third formant frequency
tuns	Tunisian	**VOT**	voice onset time
moca	Moroccan	**COG**	Centre of Gravity
kwur	Kuwaiti	**PCA**	Principal Component Analysis
ombu	Omani	**PC1**	First Principal Component
GMM	Geometric Morphometric	**PC2**	Second Principal Component
GPA	Generalised Procrustes Analysis	**UTI**	Ultrasound Tongue Imaging
MRI	Magnetic Resonance Imaging	**AAA**	Articulate Assistant Advanced
GAMMs	Generalised Additive Mixed Models	**TR**	Tongue root
MRI	Magnetic Resonance Imaging	**TD**	Tongue dorsum
IVAr	The Intonational Variation in Arabic	**LPC**	Linear Predictive Coding
LMER	Linear Mixed-Effects Regression model		
GLMER	Generalized Linear Mixed-Effects Regression model		

Symbols of Transcription

Arabic Script	IPA symbol	Description
Consonants		
أ - الهمزة	ʔ	Glottal stop
ب	b	Voiced bilabial stop
ت	t	Voiceless alveolar stop
ث	θ	Voiceless dental fricative
ج	dʒ	Voiced post-alveolar affricate
ح	ħ	Voiceless pharyngeal fricative
خ	χ	Voiceless uvular fricative
د	d	Voiced alveolar stop
ذ	ð	Voiced dental fricative
ر	r	Voiced alveo-palatal trill
ز	z	Voiced alveolar fricative
س	s	Voiceless alveolar fricative
ش	ʃ	Voiceless post-alveolar fricative
ص	sˤ	emphatic voiceless alveolar fricative
ض	dˤ	emphatic voiced alveolar stop
ط	tˤ	emphatic voiceless alveolar stop
ظ	ðˤ, zˤ	emphatic voiced interdental fricative
ع	ʕ	Voiced pharyngeal fricative
غ	ʁ	Voiced uvular fricative
ف	f	Voiceless labio-dental
ق	q	Voiced velar stop
ك	k	Voiceless velar stop
ق	g	Voiced velar stop
ل	l	Voiced alveolar lateral- approximant
م	m	Voiced bilabial nasal
ن	n	Voiced alveolar nasal
ه	h	Voiceless glottal fricative
و	w	Voiced labiovelar glide
ي	j	Voiced palatal approximant
Vowels		
ا	a:	Long low front vowel
ا	ɑ:	Long low back vowel
ي	i:	Long front high vowel
ي	e:	Long front mid vowel
و	u:	Long high back vowel
و	o:	Long mid back vowel
فتحة َ	a	Short low front vowel
كسرة ِ	i	Short Front high vowel
ضمة ُ	u	Short high back vowel

Dedication

To my dad

(*May Allah have mercy upon his soul*)

Acknowledgments

Completing this PhD has been an incredibly long and challenging journey. Reaching this point of submitting my thesis brings a deep sense of relief. This accomplishment, however, would not have been possible without the valuable support and encouragement of several people. As our Prophet, may Allah bless him and grant him peace, beautifully said, "He who does not thank people does not thank Allah," I extend my deepest appreciation to all those who have stood by me throughout this endeavour.

First and foremost, I express my profound gratitude to my supervisors, Dr. Sam Hellmuth and Dr. George Bailey. Dr. Hellmuth's vast knowledge, constant support, and patience were essential in my academic development. Her guidance was not just instrumental but transformative, for which I am forever grateful. Dr. Bailey's support throughout the research years was equally essential. The countless hours spent in the Articulatory Lab, mastering the Ultrasound equipment under his instruction, have been critical to my study's success. His encouragement fostered my independence and confidence in my research capabilities.

To my dear children, Sally, Abdulelah, and Abdulrahman, your boundless joy and patience have been the light through my journey. Your understanding and love made all the difference during times of absence.

To my wife, the cornerstone of my journey, your support during the most challenging periods was my strength. You effortlessly took on dual roles, providing unwavering support and understanding, making our family's journey smoother.

Last but not least, to my father-in-law, who never stopped supporting me and pushing me toward achieving my dreams.

Lastly, to my mother, whose prayers have been my shield and encouragement. Her pride and faith in me, voiced in every conversation, have been a source of endless motivation and strength.

This journey has indeed been a collaborative effort, and my words cannot express the depth of my gratitude. Each of you has played a pivotal role in this significant chapter of my life, and for that, I am forever thankful.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

1 Chapter 1: Introduction

1.1 Aim of the Thesis

One of the phonological characteristics that define the Arabic language is the consideration of being the Lughat AlDhad[1] which means that this language does have sound features which do not exist in any other language other than Arabic (Chejne, 1969). Also, most Semitic languages (excluding Maltese and Modern Hebrew) have a category of segments known as 'emphatics' (Bellem, 2008). These segments exhibit different patterns of manifestations at different levels, acoustic, perceptual, and articulatory in terms of the segment itself and the prosodic word to which it belongs (Sakr, 2023). Examining these patterns is crucial for accurately describing the language varieties they appear within and to better understand the nature of these patterns, phonologically, and their impacts on Arabic varieties (Sakr 2023). This thesis embarks on a journey to unravel the mysteries surrounding emphatic consonants, exploring their multifaceted nature and their implications for Arabic linguistics and phonology.

1.2 Contextualising Emphatic Consonants

The term "emphatics" refers to a category of consonants that possess unique articulatory features, typically involving a constriction or closure at the pharyngeal or uvular region during their production. These consonants are characterised by a heightened degree of muscular tension in the articulatory organs, resulting in a distinctive and perceptually salient sound quality (Al-Ani & el-Dalee, 1984). In Arabic, emphatic consonants play a pivotal role in lexical and morphological distinctions, often serving as markers of grammatical categories or contributing to semantic nuances within words (Sakr, 2023).

The significance of emphatic consonants in Arabic linguistics lies in their multifunctional role within the language. Emphatics are not only essential for maintaining lexical contrasts, but they also reflect broader sociolinguistic and dialectal patterns. For instance, the distribution and realisation of emphatic consonants often vary across Arabic dialects, making them a rich area for investigating dialectal variation. Moreover, their phonetic and acoustic properties have implications for understanding phonological processes, such as vowel

[1] Speakers of this language consider themselves "al-natiqun bil-Dad," meaning the speakers of the letter Dhad [ض], which has the phonetic transcription /dˤ/. This sound is a distinctive feature, believed to be unique to the Arabic-speaking world.

harmony, assimilation, and coarticulation, which are influenced by the emphatic feature (Ghazeli, 1977; Bellem, 2008). The perceptual salience of these sounds also raises significant questions about how native speakers recognise and categorise them, further underscoring their cognitive and linguistic importance (Bellem, 2008).

The study of emphatic consonants extends beyond mere phonetic analysis, encompassing broader linguistic dimensions such as phonology, morphology, and dialectology. Understanding the articulatory mechanisms and acoustic properties of emphatic consonants is crucial for accurately describing the phonetic inventory of Arabic and elucidating the phonological processes that govern its linguistic structure (Al-Ani & el-Dalee, 1984). Moreover, the perceptual salience of emphatic consonants poses intriguing questions regarding their cognitive processing and perceptual categorisation, prompting investigations into the mechanisms underlying their recognition and interpretation by native speakers (Bellem, 2008).

1.3 Rationale for the Study

Despite the significance of emphatic consonants in Arabic linguistics, there remains a paucity of comprehensive studies that delve into their acoustic, perceptual, and articulatory characteristics across different Arabic dialects. Existing research on emphatics often focuses on a single type of methodology, usually acoustic analysis, and few studies are known to use two methodologies—acoustics and perception—such as those by Jongman et al. (2011) and Ali & Daniloff (1972). Furthermore, most research on emphatics tends to focus on only one or two dialects of Arabic and primarily examines emphatic plosives.

This thesis seeks to address these gaps in the literature by undertaking a comprehensive examination of emphatic consonants within the context of Arabic linguistics. By adopting a multidisciplinary approach that integrates articulatory, acoustic analysis, and perceptual experiments, this research aims to provide a holistic understanding of the phonetic, and perceptual dimensions of emphatic consonants in Arabic. By explaining the articulatory gestures, acoustic correlates, and perceptual cues associated with emphatics, this study aims to contribute to a deeper understanding of Arabic dialectal variation, paving the way for future research in this field (Sakr, 2023).

1.4 Outline Summary

Following the introductory chapter, this thesis provides a detailed background on the anatomical aspects of the speech organs, with a focus on the articulators relevant to the production of emphatic sounds. This includes an examination of their musculature and articulatory movements to fully understand two key concepts: the articulation of emphasis and the potential physiological interdependence of tongue regions that could influence it. Additionally, an exploration of the nature of emphasis and its varying definitions by prominent linguists is presented. This is done to comprehend the diverse terminologies used to describe emphasis. Moreover, relevant studies that have discussed the emphatic contrast from articulatory, acoustic, and perceptual perspectives are included. This allows for a connection between them and the findings of this thesis, thereby highlighting its contribution.

In Chapter 3, the thesis conducts an extensive examination of the acoustic results across eight Arabic dialects, aiming to discern whether the phonetic realization of the emphatic contrast between /s/ and /sˤ/ remains consistent, offering a broader and more nuanced perspective. By employing a series of acoustic correlates, including fricative duration, intensity, and spectral properties, in conjunction with vowel formant information, the thesis seeks to uncover potential typological divides and acoustic markers for the emphatic contrast, thereby enriching our understanding of this phenomenon. Also, the dynamic approach of vowel trajectory was also employed to explore the effect of the F2 trajectory in emphatic contrast distinction.

In Chapter 4, building on the acoustic findings, the thesis explores perceptual analysis, investigating how Arabic listeners discern emphatic contrasts when presented with cues from their native dialects as well as unfamiliar cues from other dialects. Utilizing resynthesized stimuli, the study aims to determine whether the findings from the perceptual experiments align with those from acoustics and if the information provided by vowels near emphatics aids in recognizing emphasis. Specifically, the effect of F2 is studied, and its size difference and trajectory slope are manipulated, to determine whether listeners across dialects will attend differently to the manipulated and real signals

In Chapter 5, the thesis advances into the articulatory analysis, employing ultrasound imaging to explore the covert articulation underlying the emphatic contrast. By examining lingual and labial articulatory movements, it aims to uncover hidden nuances in articulatory mechanisms that may not be readily apparent acoustically. This includes analysing tongue movements during the articulation of emphatic sounds and exploring the relationship between lip movements and the articulation of emphatic contrasts.

In chapter 6, the thesis offers a comprehensive overview of the emphatic contrast phenomenon across Arabic dialects. It critically evaluates the hypotheses posited in each chapter, elucidating the interconnections among acoustic, perception, and articulatory analyses. Additionally, it fosters a nuanced discussion on the correspondence between perception, acoustics, and articulation, highlighting the multifaceted nature of the emphatic contrast phenomenon. Finally, the thesis presents its contributions to the existing literature while acknowledging its limitations, advocating for future research endeavours to delve further into individual specificity and methodological nuances.

Through these interconnected studies, this thesis aims to provide a comprehensive understanding of emphatic consonants in Arabic linguistics, shedding light on their articulatory, acoustic, and perceptual properties and their broader implications for Arabic phonetics and dialectology.

2 Chapter 2: Background

2.1 Vocal Tract Organs of Speech

In this thesis, considerable attention is given to the vocal organs in their relation to both acoustic properties and articulatory functions. Specifically, the discussion centres on the articulators essential for producing emphasis or pharyngealized sounds. While the primary focus is on the positioning of these articulators rather than an in-depth examination of their internal anatomical structures, it remains crucial to comprehend their musculature and articulatory movements. Such understanding is essential for fully grasping two significant concepts: the articulation of emphasis and the potential physiological correlations among the tongue’s various portions that may potentially influence this articulation.

2.1.1 The Tongue

The tongue, renowned for its extraordinary muscular flexibility, holds a crucial role in the realm of speech production. Linguists categorise it into four distinct sections based on their positioning within the vocal tract: the tip, blade, dorsum (back), and root (Bin-Muqbil, 2006). Notably, the back and root regions of the tongue interconnect with various fixed structures such as the velum (soft palate), pharynx, epiglottis, and hyoid bone. Comprising entirely of muscles, the tongue possesses a remarkable capacity for executing intricate and finely controlled movements. Operating in the capacity of a muscular-hydrostat (Kier & Smith, 1985), aqueous liquids enrich the tongue's musculature, thereby displaying hydrostats’ fundamental biomechanical attribute, called volume preservation. Consequently, any alteration in the tongue’s dimension (any one of it) prompts compensatory adjustments in one or more dimension. As the tongue root moves forward, it induces a compensatory elevation in dorsum height to preserve the tongue's volume.

Muscles typically feature fixed origin points and movable insertion points during contraction. The muscular architecture of the tongue is notably intricate, comprising two primary categories of muscles based on their origins: intrinsic muscles, which originate and insert into different parts of the tongue; and extrinsic muscles, which stem from surrounding skeletal structures and get inserted into the organ under discussion (Kent, 1998; Zemlin, 1998; Seikel, King, & Drumright, 2000; Gick, Wilson, & Derrick, 2013; Sanders & Mu, 2013; Hixon, Weismer, & Hoit, 2014; Stone et al., 2018; Crumbie et al., 2019). The majority of lingual muscles exist in pairs, facilitating the division of the tongue into two lateral halves.

2.1.1.1 *Extrinsic Muscles*

The styloglossus, genioglossus, hyoglossus, and palatoglossus, are examples of extrinsic lingual muscles (Figure 1), each explained here concerning its structure and role in tongue movement (as per Kent, 1998; Zemlin, 1998; Seikel et al., 2000; Gick et al., 2013; Hixon et al., 2014; Stone et al., 2018).

Being the biggest and the most robust of the extrinsic lingual muscle, the genioglossus plays a crucial role in facilitating various tongue movements and positions. Originating from three fibre bundles attached to the mental spine, a bony projection on the posterior mandible's midline, it inserts into different areas of the tongue: the lower bundle connects to the tongue root, the middle to the junction between the dorsum and blade, and the upper to the tongue tip. Depending on the contracting bundle, the genioglossus enables movements such as the apex's protrusion and the root's advancement elevating the body of the tongue while lowering its anterior portion, and depressing the midline for establishing a longitudinal groove.

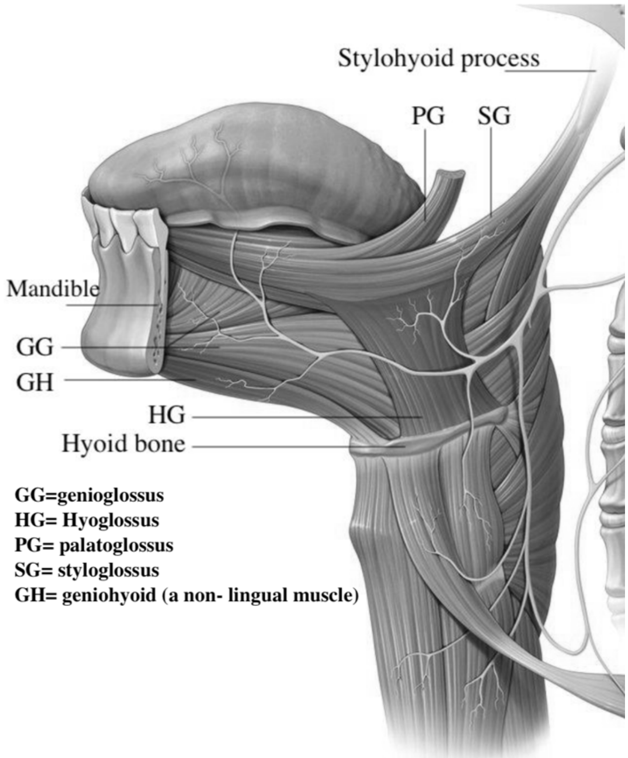


Figure 1 Diagram illustrating the cervical vertebrae and extrinsic lingual muscles in the speech mechanism. Adapted from Crumbie, Salvador, and Rad (2019).

Arising from the styloid process of the temporal bone, the styloglossus muscle extends anteriorly and downward, where its fibres insert into the sides of the tongue root, intertwining with intrinsic and hyoglossus muscles. When the styloglossus contracts, it generates various motor actions, including pushing the dorsum backward, elevating the lateral portions of the tongue, pulling the tip towards the sides of the oral cavity, and lifting the dorsum upward.

There are discrepancies in studies regarding the precise origin and insertion points of the palatoglossus muscle. Some suggest it originates from the lower surface of the soft palate, with fibres extending downward and forward, and then inserting laterally into the sides of the tongue (either into the body or the root). Alternatively, it could originate from the lateral edges of the tongue, inserted into glossopalatine, the soft palate. Differences in attachment points could impact how the palatoglossus functions in movements involving the velum compared to those involving the tongue. This muscle has the capacity to retract the tongue backward and form a longitudinal groove through both sides of the coordinated contraction.

The hyoglossus muscle refers to a quadrilateral muscle sheet stemming from the hyoid bone’s greater cornua. This thin muscle inserts into the posterior half of the tongue's sides. It retracts and depresses the tongue, pulling the sides downward and elevating the hyoid bone, in direct opposition to the palatoglossus.

2.1.1.2 Intrinsic Muscles

Intrinsic lingual muscles play a crucial role in shaping the tongue and are essential for various articulatory movements. Unlike extrinsic muscles, which pull the tongue towards bony structures, intrinsic muscles deform the tongue into different shapes through squeezing actions (Gick et al., 2013). These muscles contribute to the tongue's biomechanical properties, resembling muscular hydrostats, which require the coordination of multiple muscles for deformation (Stone et al., 2018).

There are four intrinsic muscles, organised into two pairs: longitudinal and transverse-vertical muscles. The superior longitudinal muscle forms a broad, thin layer beneath the tongue's surface, extending from the root to the tip. It can curl the sides of the tongue upward, bend the front of the tongue, and shorten it to create a concave shape (Figure 2). This muscle's structure enables independent control of specific tongue regions, minimising conflicts during articulation.

Contrastingly, the inferior longitudinal muscle lies near the tongue's undersurface, adjacent to the genioglossus muscle. Originating from the hyoid bone near the root, it extends forward to insert into the lower surface of the tongue tip. Contraction of its anterior portion results in the downward curling of the tongue tip, while contraction of the entire muscle alters the tongue's shape to a convex form.

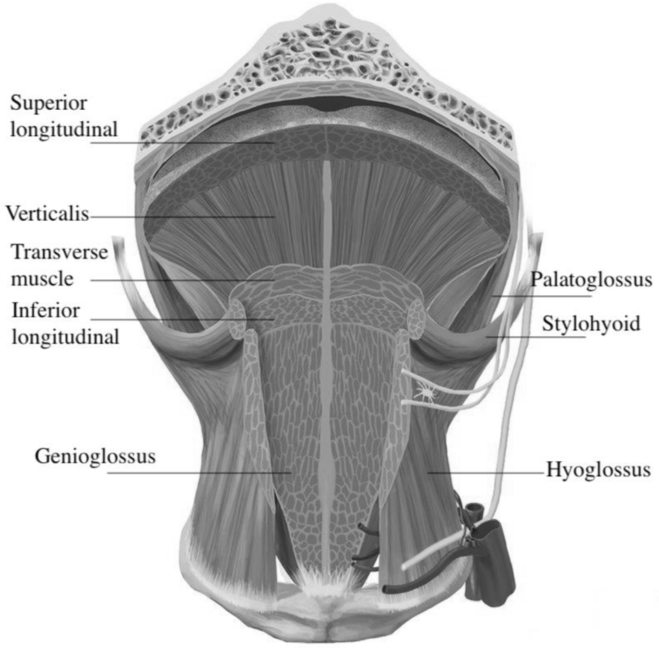


Figure 2 Frontal view that highlights intrinsic as well as select extrinsic lingual muscles. Adapted from Crumbie et al. (2019).

The transverse intrinsic muscle extends laterally as well as horizontally from the midline of the tongue to its sides. Its action involves squeezing the tongue from side to side, which leads to the protraction and thickening of the tongue mass.

In contrast, the vertical intrinsic muscle interlaces with the fibres of the transverse muscle and runs vertically between the inner surfaces of the superior and inferior longitudinal muscles. Contraction of the vertical muscle results in the elongation, widening, and flattening of the tongue.

2.1.2 The Pharynx

Essentially, the pharynx can be described as a tubular structure that stretches from the nasal cavity's back portion to the larynx. Above the velum, its upper part is referred to as the nasopharynx. The segment from the velum to the hyoid bone is referred to as the oropharynx, while the section extending below the hyoid bone to the area above the larynx is known as the laryngopharynx. Figure 3 shows that the pharynx encompasses three constrictor muscles, crucial components. Originating from the pterygomandibular ligament, the superior constrictor muscle extends rearward to converge at the tendinous raphe midline. Similarly, the middle constrictor muscle, originating at the hyoid bone and stylohyoid ligament, moves backward to merge at the tendinous raphe. In contrast, the inferior constrictor muscle, with its expansive fibre sheet, originates from the cricoid cartilage and thyroid lamina, wrapping around to connect at the midline tendinous raphe. Contraction of any of these three pharyngeal constrictor muscles leads to a reduction in the pharynx's diameter at their respective locations. For further elaboration, refer to Zemlin (1968), Lieberman and Blumstein (1988), Palmer (1993), and Seikel et al. (1997).

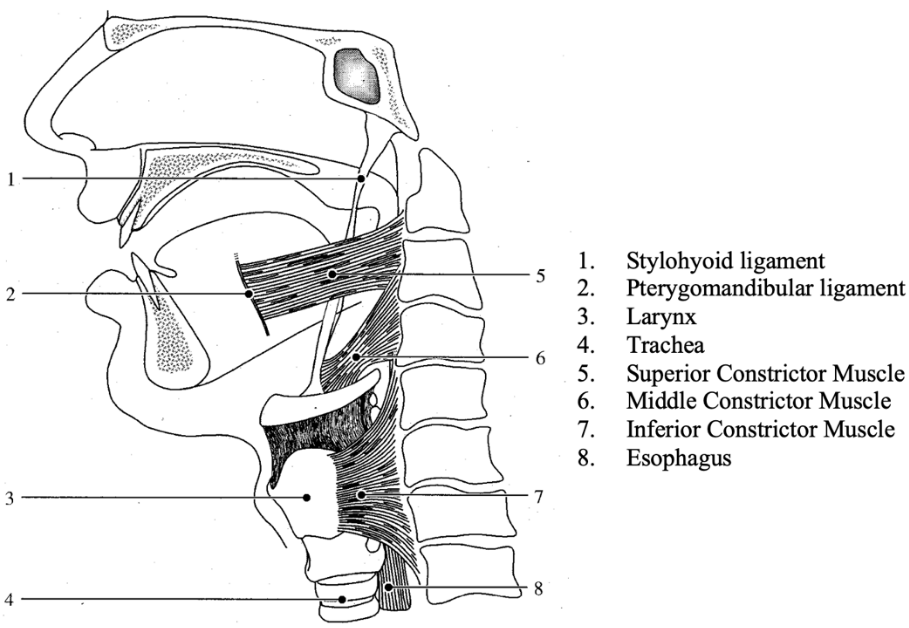


Figure 3 Overview of pharyngeal constrictors and associated anatomical components (Seikel et al. ,1997).

2.1.3 **The Soft Palate (Velum)**

The soft palate, also known as the velum, forms the flexible rear portion of the mouth's roof, comprising muscle fibres and tissues. It connects at the front to the hard palate's end through the palatal aponeurosis and on the sides to the upper pharyngeal constrictor muscles. During speech, the velum is crucial for creating nasal sounds, allowing air to flow via the nasal passages for nasalization when it is lowered. As illustrated in Figure 4, the soft palate is governed by two primary muscles responsible for lifting it, namely the levator veli palatini and the uvular muscle, along with two main muscles tasked with lowering it, namely the palatoglossus and the palatopharyngeal muscle. The genesis of the dual muscle levator veli palatini is from the temporal bone and the Eustachian tube, connecting to the velum's aponeurosis. Its activation results in the upward and backward elevation of the velum. Similarly, the uvular muscle, originating from the posterior aspect of the palatal bones and the palatine aponeurosis, extends backward to the uvula, lifting the velum upon contraction. The palatoglossus muscle, previously discussed as an extrinsic tongue muscle, contributes to reducing the velum's height when contracted or elevating the tongue's rear if the velum remains stationary. Another muscle involved in lowering the velum is the palatopharyngeal muscle, which originates from the soft palate and extends laterally and downward to attach to the thyroid cartilage and the walls of the pharynx, forming part of the posterior facial pillars. Contraction of the palatopharyngeal muscle pulls the faucial pillars inward, leading to a reduction in the pharynx's diameter through a sphincteric action and potentially elevating the larynx. For further insights, refer to Zemlin (1968), Palmer (1993), as well as Seikel et al. (1997).

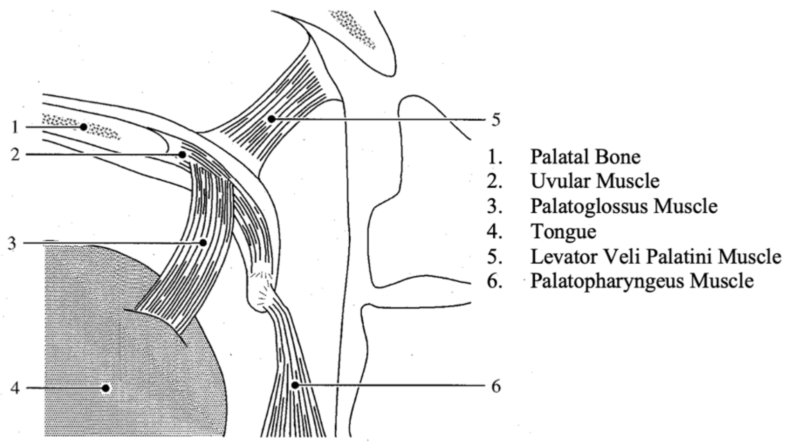


Figure 4 The soft palate muscles and their related structures

2.1.4 The Larynx

The larynx, as depicted in Figure 5, presents a complex structure composed of various cartilages and muscle tissues essential for vocalisation. Among its major cartilages are the cricoid, thyroid, arytenoid cartilages, and the epiglottis. The cricoid cartilage forms a ring-like structure serving as the larynx's base, positioned above the trachea, with a thicker posterior part compared to the anterior. Situated above the cricoid, the thyroid cartilage, the largest in the larynx, connects to the cricoid via pairs of joint facets on each side. The arytenoid cartilages, shaped like pyramids and situated on the upper posterior surface of the cricoid, provide anchorage for the vocal folds, which extend to the inner front surface of the thyroid. Additionally, the epiglottis, resembling a shoe-horn, attaches to the lower front inner surface of the thyroid, extending upwards past the hyoid bone and connecting to the arytenoids via the aryepiglottic folds. While not directly part of the larynx, the arch-shaped hyoid bone lies above it, connecting to the thyroid through lateral hypothyroid ligaments and the hypothyroid membrane, extending from the hyoid's posterior tips to the thyroid's horns and from the hyoid down to the thyroid's upper edge, respectively. Intrinsic muscles within the larynx include the lateral cricoarytenoid, transverse arytenoid, oblique arytenoid muscles, and the cricothyroid muscle. The lateral cricoarytenoid muscle, functioning as a vocal fold adductor, extends from the upper rim of the cricoid to the muscular processes of the arytenoid cartilages, rotating the arytenoids to bring the vocal folds together. Similarly, the transverse arytenoid muscle, positioned between the arytenoids, contracts to pull the vocal folds together. The oblique arytenoid muscles, spanning from the base of one arytenoid to the apex of the other, form the aryepiglottic muscles, inserting into the sides of the epiglottis. Working in tandem with the oblique arytenoids, these muscles pull the epiglottis down over the larynx. Lastly, the cricothyroid muscle, originating from the front and sides of the cricoid, divides into the pars recta and pars oblique, attaching to the lower edge of the thyroid. Contraction of these muscles tilts the thyroid downwards or forwards, respectively, increasing the distance between the thyroid and arytenoids, thereby tensioning the vocal folds. For a more comprehensive understanding, refer to Zemlin (1968), Lieberman and Blumstein (1988), Palmer (1993), and Seikel et al. (1997).

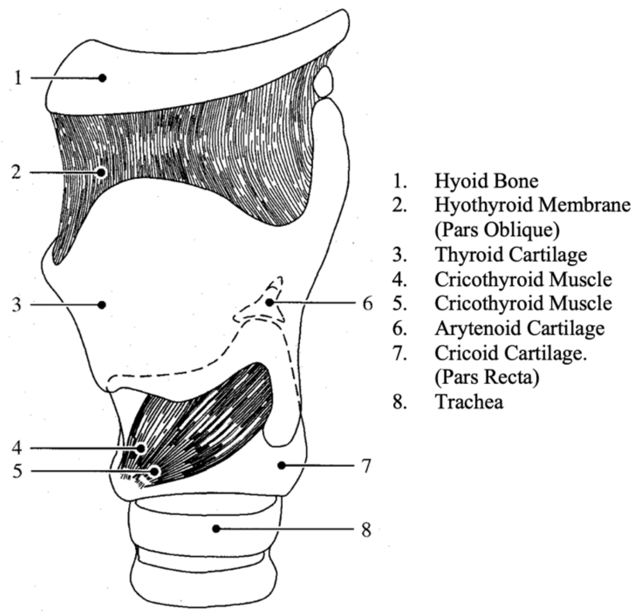


Figure 5 The larynx structure

2.1.5 **The Lips**

According to Perkins and Kent (1986), the orbicularis oris muscle encircles the mouth in an oval-shaped band, serving as a sphincter to close or pucker the lips. Notably, the lips feature distinctive characteristics: the vermilion border, also known as the prolabium, delineates the reddish area surrounding the lips, with the cupid's bow marking the centre of the upper lip and the philtrum indentation located just above it. Moreover, the columella refers to the soft tissue that divides the nostrils above the philtrum. In cases of cleft lip, the philtrum lacks connection to the rest of the lip. Various facial muscles, which insert into the lips, facilitate a wide range of movements for different facial expressions. For speech production, the extrinsic muscles responsible for widening the lips, such as those involved in smiling, are particularly relevant. Among these muscles, the risorius muscle stands out due to its superficial placement and limited strength, contributing to lip spreading. Additionally, the buccinator muscle, a significant transverse muscle, attaches at the corners of the mouth to extend the lips. This arrangement forms a muscular loop with the orbicularis oris muscle at the front and the superior constrictor muscle at the back, encircling the throat and mouth. For a visual representation, please refer to Figure 6 below.

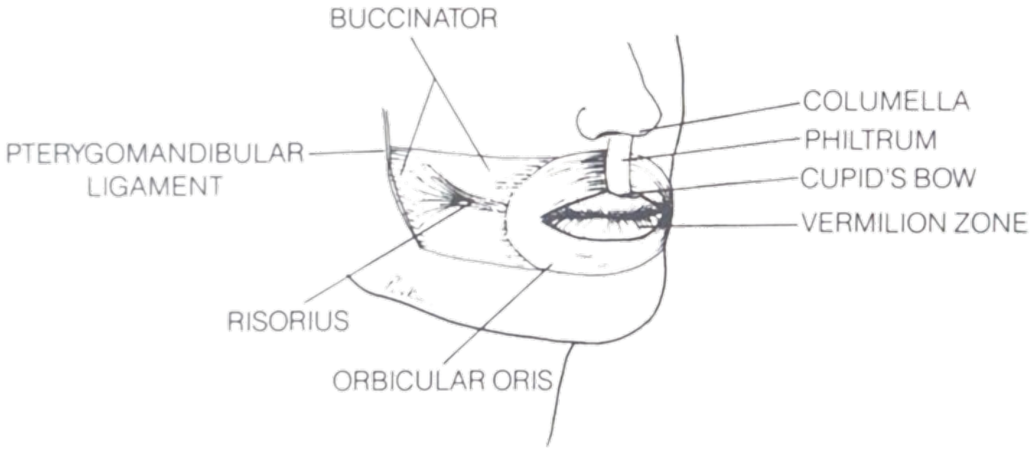


Figure 6 Facial muscles, including lips (Perkins and Kent,1986)

2.2 Emphasis (Definition and Nature)

Emphasis, a distinctive feature of the Arabic language, historically referred to as Lughat Al-Daad (the language of the letter Daad, a voiced emphatic consonant), has been acknowledged since the 8th century, notably by Sibawayh, the author of the first Arabic grammar book (Almuhaimeed, 2022). Sibawayh utilised the term [itʕba:q], meaning 'covering', to describe the articulation of emphatic sounds, where the tongue is brought close to the corresponding area of the palate during pronunciation, as described in Lehn (1963), AI-Nassir (1993), and Al-Tamimi and Heselwood (2011). Meanwhile, the renowned physician Ibn Sina contributed to the scientific understanding of emphasis, using the term [istiʕlaʔ] or 'elevation' to describe this phonetic phenomenon. This demonstrates the early acknowledgment of emphasis within the framework of Arabic linguistics (Card, 1983).

Among various definitions in the literature, Lehn's (1963) provides a comprehensive understanding of emphasis, highlighting its intricate articulatory features. These include slight tongue retraction and lateral spreading, along with concavity and elevation of the tongue's back, akin to velarisation. Additionally, emphasis entails faucal and pharyngeal constriction, known as pharyngealization, as well as slight lip protrusion or rounding, termed labialisation, and an overall increase in oral and pharyngeal musculature tension. This coordinated movement results in emphatic sounds being stronger than their non-emphatic counterparts.

However, it is inaccurate to associate emphasis solely with pharyngealization, as some researchers have suggested. While most emphatics involve pharyngealization, some also exhibit labialisation (Watson, 1999). Delattre (1971) elucidates the articulation of pharyngeal consonants, describing how the tongue's root forms a pronounced bulge, retracting towards the posterior wall of the pharynx to create constriction. While informative, Delattre's (1971) explanation does not fully capture the articulation of emphatic sounds, which involve an additional secondary articulation.

Kahn (1975) defines emphasis as a secondary articulation in the pharyngeal region, primarily affecting stops and fricatives. She suggests that producing an emphatic sound requires not only primary articulatory organs but also an additional pharyngeal articulation. However, her analysis predominantly focuses on stops and fricatives, potentially overlooking other sound segments that may possess emphatic attributes, as suggested by research on Cairene Arabic (Ferguson, 1956).

McCarthy (1994) distinguishes between emphatic sounds and pharyngealized consonants, proposing that the former solely involve emphasis, while the latter should be termed uvularised. Davis (1995) equates emphasis with pharyngealization, stating that it involves producing sounds with a primary articulation at the front of the mouth and a secondary constriction in the upper pharynx, expanding the definition to include bilabial sounds alongside those made with the teeth and alveolar ridge.

2.2.1 Summary of the Aforementioned

The concept of emphasis in the Arabic language, notably recognised since the 8th century by figures like Sibawayh and Ibn Sina, entails a nuanced articulatory process that transcends simple pharyngealization to include features such as velarisation, labialisation, and a complex orchestration of muscular tensions in the oral and pharyngeal regions. This multifaceted approach to articulation, highlighted by the comprehensive analysis of scholars like Lehn, Delattre, Kahn, and McCarthy, underscores emphasis not just as a phonetic characteristic but as a distinct linguistic phenomenon that involves a sophisticated interplay of anatomical adjustments. These adjustments, such as the retraction of the tongue, elevation towards the palate, and the vocal tract’s specific contributions, add to the distinctiveness of emphatic sounds in the language, demonstrating their fortis nature compared to non-emphatic counterparts and challenging the notion that emphasis can be solely equated with pharyngealization.

2.3 Articulatory Studies on Emphasis

The literature discusses that emphasis in Arabic involves two types of constrictions: primary and secondary. Emphatic and non-emphatic sounds share similarities in their primary constriction, typically occurring at the alveolar or dental regions (Aldamen, 2013). However, they differ significantly in the secondary articulation, which occurs at the back of the oral cavity (Al-Solami, 2017). Despite extensive research on the articulation of emphasis in Arabic, a consensus has not been reached regarding the precise nature of the secondary constriction. This lack of consensus can be attributed, in part, to cross-dialectal variation and differences in research methodologies (Aldamen, 2013 & Al-Solami, 2017). Khattab et al. (2006) observed that speakers may employ various articulatory strategies influenced by factors such as dialect, gender, phonological context, and social variables, contributing to this inconsistency. Ghazeli (1977), Bin-Muqbil (2006), and Shar (2012) suggested that during the articulation of emphatic consonants, the tongue body is pulled backward into the upper pharyngeal region, resembling the articulation of uvulars, and is depressed at the palate. However, Ghazeli (1977) pointed out in his study that while the tongue body is pulled backward into the upper oropharynx during the articulation of [S], it is depressed during emphatic consonants but not during plain coronals. To gain further insights into the nature of emphasis, the following sections discuss articulatory studies conducted using various instruments.

2.3.1 Studies with Non-Ultrasound Instruments

In a cinefluorographic[2] investigation on Iraqi Arabic, Ali and Daniloff (1972) observed that emphatics are articulated with simultaneous tongue depression and a backward movement of the tongue dorsum towards the pharynx's posterior wall. They noted that the distinction between emphatics and non-emphatics lies in the retraction of the tongue dorsum, leading to a narrowing of the upper pharynx. Similar findings were reported by Ghazeli (1977) in Tunisian Arabic and by Al-Tamimi and Heselwood (2011) in Jordanian Arabic. However, Ali and Daniloff (1972) also suggested that the velum and posterior pharyngeal wall were not significantly affected during the articulation of emphatic sounds.

Using MRI on Saudi speakers, Shar (2012) discovered that emphatics are produced with dorsal retraction of the tongue, resulting in constant narrowing of the top region of the

[2] Cinefluorography is a radiographic imaging technique that captures continuous X-ray images of internal structures in motion. It has been used in speech studies to visualize articulator dynamics, including the tongue and pharyngeal walls, during speech production. Despite its utility, the method's use has declined due to health risks associated with prolonged radiation exposure (Hardcastle & Hewlett, 2006).

pharyngeal cavity, while the tongue root remains unaffected. However, in a video-fluoroscopic study of emphatics in Jordanian Arabic, Al-Tamimi and Heselwood (2011) observed that the tongue root presses against the anterior surface of the epiglottis during emphatic articulation, pushing the epiglottis towards the back of the pharynx. Additionally, they found that the larynx is raised during the production of emphatics, leading to reduced pharyngeal volume. Similarly, Zawaydeh's (1999) findings indicated that emphatic sounds involve pharyngeal narrowing. However, the study does not specify the precise location of this constriction.

The discrepancy between tongue root movement and pharyngeal volume reduction suggests that tongue root retraction in emphatics may be a mechanical consequence of tongue dorsum retraction (Altairi et al., 2017; Al-Solami, 2017). Consequently, there is inconsistency in tongue root retraction across studies. As a result, researchers have proposed various articulatory mechanisms for emphasis in different Arabic dialects. Emphatics have been suggested to be uvularised in Jordanian Arabic (Zawaydeh, 1999) and Moroccan Arabic (Zeroual et al., 2011), velarised in Lebanese Arabic (Obrecht, 1968), and pharyngealized in Iraqi Arabic (Ali & Daniloff 1972; Gianni & Pettorino 1982) and in Hijazi Arabic (Ahyad, H., & Becker, M. 2020). Focusing solely on the emphatic contrast /s/ and /sˤ/, Hermes et al. (2015) utilised electromagnetic articulography (EMA) to analyse the primary articulation of plain-emphatic /s/-/sˤ/ in Lebanese Arabic. Their investigation revealed that the front of the tongue, up to 1 cm from the tongue tip, maintains a low position behind the front teeth, with a similar location observed in both /sˤ/ and /s/. However, the tongue position during /sˤ/ is even more lower at these points compared to /s/. The disparities between the /sˤ/ and /s/ clusters showed statistical significance across all speakers. Similarly, Embarki et al. (2011) conducted an EMA analysis using a list of words in Modern Standard Arabic (MSA) produced by an Arabic speaker from Tunisia. The aim of the study was to capture speech movements during the articulation of emphatic consonants (/tˤ, dˤ, sˤ, ðˤ/) and their plain counterparts (/t, d, s, ð/). The findings demonstrated differences in the coarticulatory effects between emphatic and non-emphatic consonants, with emphatic consonants exhibiting significant tongue retraction and a minor lowering near the vowel /a/. Using a similar instrument but obtaining different results, Zeroual et al. (2011) examined the emphatic stop /dˤ/ and /tˤ/ and their plain counterparts in the Moroccan dialect. They found that plain sounds exhibit a more laminal contact compared to emphatic sounds, which are more apical. Additionally, they noted slight labialisation during the articulation of emphatic sounds.

Likewise, previous studies have reported some degree of lip protrusion associated with emphasis (see El-Halees, 1985; Hetzron, 2013; Jakobson, 1957). For instance, Lehn (1963) observed slight lip rounding in the Carine dialect, while Bellem & Watson (2014) reported labialisation in Ṣanʿāni Arabic and in southern (gilit) Iraqi Arabic. It is noteworthy that these instances of lip-rounding were based on visual and theoretical analyses of the spread of emphasis and its effect on tongue movements, rather than being supported by instrumental evidence. Therefore, the current study will utilize ultrasound imaging examination to investigate the association of lip movement with tongue movement through video recording for lip-rounding.

2.3.2 Studies using Ultrasound Imaging

A study proposed by Altairi et al. (2017) recruited eight speakers from various Arabic dialects (2 Saudi, 2 Yemeni, 2 Egyptian, 1 Palestinian) to investigate tongue movements across different sound groups (/tˤ/, /sˤ/, /t/, /s/, /ħ/, /ʕ/, /χ/, /ʁ/) using Smooth Spline ANOVA (SS-ANOVA) analysis. They compared these movements to the neutral tongue position, termed the 'inter-speech posture' (ISP), as suggested by Gick et al. (2004), serving as a baseline for comparing and measuring speech sound postures. The study found a significant difference in tongue root (TR) position between emphatic and non-emphatic consonants across all subjects, with emphatic consonants exhibiting substantial TR retraction compared to their non-emphatic counterparts. Additionally, six subjects showed a significant disparity in tongue dorsum (TD) position between emphatic and non-emphatic consonants, with the TD in emphatics positioned further back and higher compared to non-emphatic sounds. Most participants consistently positioned the emphatics posterior to the tongue root of the ISP, indicating a consistent lowering of the tongue body position. However, there were no differences between emphatics and the ISP in terms of tongue dorsum, except for two participants whose emphatics were articulated with the tongue dorsum positioned higher compared to the ISP. Conversely, non-emphatic sounds showed an advancement in tongue root position and a lowering of the tongue dorsum compared to the ISP. Similarly, Al-Solami (2017) conducted an ultrasound imaging study on three participants representing different Arabic dialects (Saudi Arabian, Egyptian, and Palestinian) to explore the mechanism of tongue movements for emphatic, uvular, and pharyngeal sounds. He hypothesised that these sounds involve tongue retraction as a secondary articulatory component, albeit differing in the degree of retraction. The findings regarding the emphatic contrast suggested that the tongue dorsum is more elevated and retracted, with the blade depressed behind the main constriction point during emphatic articulation. Additionally, Zeroual et al. (2011) conducted a study utilising EMA and endoscopic methods, supplemented by ultrasound investigation, to collect data from Moroccan Arabic speakers, aiming to address various questions regarding tongue movements.

In their investigation into the secondary articulation of Moroccan emphatic sounds, Zeroual et al. (2011) compared the characteristics of Moroccan emphatic coronals (/tˤ, dˤ, sˤ/) with their plain counterparts (/t, d, s/), as well as uvulars and pharyngeals. By means of ultrasound research, they enlisted two Moroccan speakers and utilised both words and nonsense words containing emphatic sounds to assess the positions of the tongue and epiglottis. Their findings revealed that the articulation of emphatics more closely resembled uvulars than pharyngeals. They observed a backward movement of the tongue towards the posterior pharyngeal wall during emphatic articulation, whereas pharyngeals involved a backward movement of both the tongue and the epiglottis. Similarly, Alfaifi et al. (2020) conducted an ultrasound study on two Saudi Hijazi participants to examine the primary and the secondary constrictions of the voiceless stop /tˤ/ and the fricative /sˤ/, alongside their plain counterparts /t/ and /s/. Their objective was to explore the impact of emphasis on adjacent vowels and vice versa. Their findings indicated that in emphatics, the tongue root is elevated and more retracted compared to non-emphatic sounds. They also observed that short vowels were more susceptible to emphasis effects than long vowels. Furthermore, they noted that high vowels influenced the shape of the tongue body during the production of emphatics, although this was predominantly observed in one of the participants.

2.3.1 Summary of the Aforementioned

In summary, previous studies have primarily concentrated on exploring distinctions among guttural sounds, analysing tongue movements across various sound classes such as emphatics, pharyngeals, uvulars, and laryngeals. While they have offered insights into emphatic and non-emphatic differences, this aspect wasn't their main focus. The key point of interest lies in understanding the findings regarding the disparities between (s~sˤ) in terms of tongue positions. While there's consensus on the tongue configurations for the primary constriction, there's slight variation in describing tongue positions during the secondary constriction. This variation might lead researchers to categorise Arabic dialects as pharyngealized, uvularised, or velarised.

Overall, observations indicate that in emphatic sounds, the tongue dorsum is retracted and raised, with a depressed area behind the front part of the tongue forming a concavity compared to non-emphatic sounds. Tongue root retraction is more pronounced in emphatic sounds compared to non-emphatics, and among guttural sounds, this varies based on the degree of tongue dorsum retraction. While some studies have noted lip-rounding during emphatic articulation, these observations have mainly relied on acoustic, auditory or visual analysis and have not been explored through ultrasound imaging.

Thus, this study aims to examine the distinction between /s/ and /sˤ/ in terms of tongue movements and lip association. The objective is to determine whether the differences in tongue configurations between emphatic and non-emphatic sounds remain consistent across the target dialects and to what extent ultrasound imaging findings correlate with acoustic outcomes.

2.4 The Concept of Covert Articulation

This section will present related studies from English literature, given that no studies on this topic have been conducted in Arabic literature.

The evolution and diversification of English dialects have been subject of extensive study in the field of sociolinguistics and phonetics. The rhotic/non-rhotic division, which provides insights into mechanisms of historical /r/-loss sound change, is one of the most fundamental divisions in English dialects. However, intriguing complexities are presented by contemporary changes in /r/-loss in rhotic varieties, such as Scottish English. This literature review aims to synthesize recent research on covert articulation in English dialects.

Historically, the loss of /r/ in English dialects has been a well-documented phenomenon but mechanisms behind this sound change remain elusive. A recent study by Lawson, Stuart-Smith and Scobbie (2014) examined weakening of /r/ in contemporary Scottish English suggests /r/ weakening is gesture-timing-based phenomenon which is socially indexical. This implies that the way /r/ is articulated is influenced by social factors.

The investigation utilised a socially-stratified conversational ultrasound tongue imaging speech corpus to examine the impact of boundary context, syllable stress, following-consonant place, and social class on lingual gesture timing in /r/. Results unveiled that in in utterance-final contexts, working-class speakers exhibited a notable delay in anterior lingual gestures for /r/, resulting in less robust or even derhoticized /r/ sounds that were often inaudible. This delay meant the /r/ gesture's peak occurred after voicing had ceased, creating a weakened rhotic quality. In contrast, middle-class speakers displayed earlier and more prominent tongue gestures, with /r/ sounds audibly reinforced. The study underscored the importance of covert articulation, referring to subtle and often imperceptible articulator movements, which can only be accurately observed and analysed through articulatory analysis. This implies that even when acoustic analysis yields different or unclear results, articulatory analysis offers a more precise and detailed depiction of tongue configurations and gestures employed by participants in producing various /r/ sound variants.

Understanding covert articulatory variations provides valuable insights into the dynamics of language change. A study conducted by Mielke, Smith, and Fox (2017) delved into two instances of covert articulatory variation observed in Raleigh, NC, focusing on the tongue shape of /ɹ/ and the location of posterior constriction in /l/. Employing mixed-effects modelling, they sought to uncover connections between the production of covert and overt variables in laboratory speech and their occurrence in spontaneous speech. One significant implication highlighted in their research is the distinct coarticulatory effects of various articulatory gestures, especially pertinent for consonants and vowels involved in phonetic or phonological patterns conditioned by covertly variable sounds. Any sound pattern sensitive to the spectrum of hyper- to hypo-articulated speech will be affected by differences in the articulation of its trigger. The study also suggested that the phonetic ramifications of covert articulatory variation pose challenges for listeners to compensate for, as the relationship between coarticulatory cause and effect may be unclear without insight into the speaker's coarticulatory motivation. Consequently, when acoustic analysis yields divergent or ambiguous results, covert articulation, discernible only through articulatory analysis, often underlies these discrepancies (Mielke, Smith, & Fox, 2017).

Southern British English's high-back vowel fronting has been extensively examined acoustically, but articulatory data offer deeper insights into this phenomenon. Strycharczuk and Scobbie (2017) compared the relative tongue positions of the vowels /uː/ and /ʊ/ in fronting and non-fronting consonantal contexts using ultrasound data. Their findings revealed similar differences between vowels in articulation and acoustics. Interestingly, tongue position in sequences like "fool" and "full" was distinct, yet there was no corresponding discrepancy in F2, challenging the conventional articulatory metaphor associating F2 increase with fronting. Furthermore, the study provided specific recommendations for recording and analyzing ultrasound data in research on vowel variation and change, stressing the significance of considering flanking consonants when estimating vowel distances and the need for caution when interpreting acoustic data in studies of high-back vowel fronting (Strycharczuk & Scobbie, 2017).

2.5 Acoustic Correlates of the Emphatic Consonants

2.5.1 Articulatory-Acoustic Relationship

This section highlights the link between articulation and acoustics that defines emphatic sounds, suggesting that articulatory data might be inferred from acoustic analysis. For instance, the acoustic findings for Modern Standard Arabic (MSA) led Bin-Muqbil (2006) to propose that coronal emphatics are closer to velarised than pharyngealized consonants. This hypothesis stemmed from the negligible impact these consonants had on the first formant (F1) of subsequent vowels. Consequently, Bin-Muqbil challenged the conventional classification of emphatics as pharyngealized in academic discourse, arguing that the articulatory and acoustics evidence did not consistently support this designation. This challenge holds significant implications, as the classification of emphatics as pharyngealized has traditionally been supported by articulatory studies. Bin-Muqbil's proposal raises the possibility that the secondary articulation of emphatic consonants may not uniformly involve pharyngeal constriction and may instead exhibit variation across languages, dialects, or even individual speakers.

Building upon the foundational principles of perturbation theory—which examines how small changes in vocal tract shape affect resonant frequencies—and the acoustic theory of speech production as discussed by Carré and Mrayati (1992), Chiba and Kajiyama (1958), Fant (1960/1971), Howard and Angus (2009), Johnson (2012), Mrayati et al. (1988), and Stevens (1989) - all cited in Al-Tamimi (2017) - the relationship between articulatory movements and their acoustic outcomes can be explored through articulatory-to-acoustic mapping. This analytical approach emphasises that pharyngeal constrictions give rise to a distinctive acoustic pattern characterised by an elevation in the first formant (F1) and a reduction in the second formant (F2). This phenomenon arises because the pharyngeal constriction is positioned near a node for F1 and an antinode for F2, leading to their respective increase and decrease in natural frequencies.

The importance of the first formant frequency (F1) in identifying the location of constrictions within the posterior vocal tract is underscored by studies employing vocal tract modelling. Works by Malmberg (1963), Klatt and Stevens (1969), and Lindblom and Sundberg (1971) establish a direct correlation between a lower-positioned pharyngeal constriction and an elevation in F1. Yeou (2001) further validates this relationship through vocal tract modelling, demonstrating a decrease in F2 alongside the rise in F1 and F3 as the pharyngeal constriction tightens (decreasing from 5 cm^2 to 1 cm^2). This consistent pattern reinforces the reliability of F1 as an indicator for localising constrictions in the posterior vocal tract.

Expanding upon the established connection between F1 and pharyngeal constriction through vocal tract modelling, Jongman et al. (2007) conduct an acoustic analysis of Jordanian Arabic. Their findings reveal a consistent pattern in the behaviour of the first three formants (F1, F2, and F3) during the production of emphatic consonants. This acoustic evidence aligns with predictions from vocal tract modelling and suggests the presence of a pharyngeal constriction in the production of emphatic consonants in Jordanian Arabic.

Kent and Read (1992) also observe a relationship between the pattern of the third formant (F3) and the location of pharyngeal constriction, noting that a constriction positioned lower in the pharynx correlates with a higher F3. Consequently, a low-pharyngeal narrowing results in an elevated F3, while a mid-pharyngeal constriction leads to a decrease in F3. Conversely, a constriction at the upper part of the pharynx either leaves F3 unaffected or causes a minor increase. Lindblom and Sundberg (1971) similarly note an increase in F3 as the tongue transitions from the velar to the pharyngeal region. Stevens (2000) proposes a general trend where all formant frequencies increase with a constriction in the lower pharyngeal area.

The variations in acoustic data for the first and third formants (F1 and F3) across different Arabic dialects may indicate diverse manifestations of the secondary articulation in emphatic sounds. While F1 and F3 are reliable indicators of the position of posterior constriction, the pattern of the second formant (F2) appears less consistent in identifying the location of the constriction compared to F1 and F3. The consistent reduction of F2 during the articulation of both velarised and pharyngealized consonants, as observed by Giannini and Pettorino (1982), suggests a correlation between F2 and tongue retraction, a phenomenon also noted by Delattre (1951), regardless of the specific location of retraction within the vocal tract.

This section emphasises the relationship between articulatory and acoustic aspects of speech, illustrating how articulatory details can be inferred from acoustic signals (Löfqvist, 1990). The physical adjustments involved in producing pharyngealized consonants result in a range of acoustic effects, evident in different Arabic dialects. The act of retracting the tongue towards the pharynx to create a pharyngeal constriction imparts distinct resonant characteristics to pharyngealized consonants, influencing adjacent sounds (Laradi, 1983).

This interaction is crucial for the perceptual distinction between emphatic and plain consonants. In this study, an investigation will be conducted on the F2 trajectories of vowels following the emphatic contrast to explore the mechanism of tongue retraction.

The variability in the articulatory processes of emphatic consonants may result in diverse patterns in formant frequencies, as different articulatory actions can either elevate or reduce the movement direction of these frequencies. For instance, studies have indicated that in addition to other articulatory features, such as tongue retraction, the elevation of the larynx and hyoid bone are associated with emphatic consonants (Laradi, 1983; Maryais, 1948, as referenced in Kriba, 2010). Laryngeal elevation shortens the vocal tract, leading to an increase in the first three formant frequencies (Stevens, 2000). Conversely, research by Kent and Read (1992) suggests that lip rounding can cause a decrease in F3 and other lower formant frequencies by elongating the vocal tract, a phenomenon potentially relevant to the articulation of emphatics. Consequently, the anticipated rise in F1 and F3 linked with emphatics may be counteracted by lip rounding, introducing complexity into the prediction of their articulatory basis. Therefore, the upcoming articulatory analysis will investigate the influence of lip movement on emphatic sounds, particularly through the examination of lip rounding effects.

2.5.1.1 Summary of the Aforementioned

Pharyngeal constriction plays a crucial role in shaping formant frequencies, as indicated by various studies. Vocal tract modelling suggests that tightening the pharyngeal constriction, characterised by a decrease in area, leads to an increase in F1 and F3, while F2 decreases (Bin Mugbil, other studies). This finding is supported by observations indicating that a lower pharyngeal constriction elevates F3, while a mid-pharyngeal constriction lowers F3, with minimal impact observed for an upper pharyngeal constriction (Lindblom & Sundberg, 1971). Additionally, research suggests an increase in F3 as the tongue transitions from the velar to the pharyngeal region (Stevens, 2000).

Studies exploring dialectal variations in Arabic reveal discrepancies in F1 and F3 values, indicating potential differences in emphatic articulation across dialects. However, the correlation between F2 and the location of the posterior constriction appears less direct. The consistent decrease in F2 observed for both velarised and pharyngealized sounds is attributed to tongue retraction towards the posterior region of the vocal tract (Pettorino, 1982; Delattre, 1951).

Moreover, there is some indication of the larynx's involvement in emphatic consonants. Elevating the larynx and hyoid bone is associated with emphasis, likely due to a shortened vocal tract and the subsequent elevation of all three formants. However, this explanation seems contradictory to the hypothesis that F1 and F3 primarily increase during emphatic articulation. Additionally, lip rounding has been linked to a decrease in formants and a lengthening of the vocal tract. Thus, further research is necessary to reconcile these seemingly conflicting effects.

2.5.2 Emphasis Spread Domain Within Arabic Dialects

In Semitic linguistics, a prominent attribute is the influence exerted by emphatic coronals on neighbouring segments, a phenomenon widely acknowledged within Arabic phonology. As highlighted by Almuhaimeed (2021) in her study, this enduring characteristic suggests that the phonetic properties of emphasis have the potential to span from individual syllables to entire phonological words. Consequently, a considerable body of research within Arabic phonology has investigated the extent to which emphasis can affect adjacent segments, yielding consistent findings across multiple studies. In Arabic phonology, the phenomenon of emphasis, known for its distinctive phonetic attributes such as secondary articulation involving pharyngealization or uvularization, has been subject to extensive investigation regarding its domain of influence within the speech stream. Ali and Daniloff (1972) proposed that in certain Arabic dialects, emphasis extends to encompass entire words, potentially affecting multiple syllables. Building upon this notion, Card (1983) employed acoustic analysis to argue that emphasis spreads throughout the entire word, with evidence from F2 analysis of both consonants and vowels indicating a symmetrical distribution, extending leftward and rightward from the word's onset and offset, respectively. Similarly, Hassan (1981) observed in his study that the phonetic characteristics of emphatic sounds transcend segmental boundaries, allowing emphasis to encompass both consonants and vowels.

However, Watson (1999) highlights substantial variability in the scope of emphasis spread across different Arabic dialects. While Classical Arabic exhibits a broad scope of emphasis, affecting entire phonological words, the dialect of Abha Saudi Arabic demonstrates a more limited spread, primarily confined to the adjacent vowel (Younes, 1993). This diversity in emphasis distribution has been documented in various Arabic dialect studies, including those by Zawaydeh (1999), Bin-Muqbil (2006), Al-Khatib (2008), and Jongman et al. (2011).

Furthermore, regional variations in the direction and extent of emphasis spread have been noted. In northern Palestinian Arabic, emphasis tends to extend leftward from the emphatic consonant to the word onset, while rightward spread is typically limited to the vowel within the same syllable (Herzallah, 1990). Conversely, in Qatari Arabic, emphasis may spread bidirectionally across the entire word (Bukshashia, 1985). These findings underscore the intricate nature of emphasis distribution and its manifestation across diverse Arabic dialects. In their acoustic study of Jordanian speakers, Al-Masri and Jongman (2004) found that emphasis spreads to both the right and left of the target syllable. However, this spread is blocked by the vowels /i/ and /u/.

2.5.2.1 Emphasis Impact on Second Formant (F2)

As has been noticed, linguists generally regard consonants as the central focus of emphasis, often referring to them as emphatic consonants. However, in the majority of prior acoustic analyses concerning emphasis, focus has been placed on examining the characteristics of the vowels that are adjacent to the emphatic consonant as opposed to the consonant. Across all Arabic dialects that have undergone instrumental scrutiny, in a consistent manner, emphasis has been identified by a decrease in the second formant (F2) of the vowel following the emphatic consonant, a point noted earlier by Jakobson, Fant, & Halle (1952), Obrecht (1968), and Al-Ani (1970), among others, and later reaffirmed by Card (1983) and Zawaydeh (1999).

Card (1983) observed that the second formant (F2) for a non-emphatic /t/ typically falls within the range of 1600Hz to 1900Hz, whereas for an emphatic /$t^ʕ$/, it is situated between 1100Hz and 1400Hz. Based on her phonological analysis regarding emphasis distribution, Card introduced the acoustic marker [+F2 drop] as a significant correlate indicating emphasis. Similarly, Zawaydeh (1999) discovered that vowels preceding an emphatic consonant (e.g., /$s^ʕ$, T/) consistently displayed a low F2, irrespective of their proximity to the emphatic sound. Conversely, vowels following the emphatic consonant exhibited a gradient in F2 reduction, with those nearer to the emphatic consonant showing a lower F2 compared to those further away. This observation strengthens the connection between the presence of emphatic sounds and the reduction of F2, particularly within the context of Ammani Arabic.

Additionally, Kulikov et al. (2021) addressed the limited exploration of emphasis spread's effects on the temporal acoustic properties of emphatic consonants, particularly in Qatari Arabic. Building on previous research by Mitleb (2001), AlDahri (2013), and Khattab, Al-Tamimi, and Heselwood (2006), which noted that VOT is consistently shorter in emphatic stops than in plain stops, Kulikov et al. (2021) aimed to investigate the impact of emphasis spread on non-emphatic /t/. They concluded that while /t/ produced in an emphatic context experiences spectral mean alterations and affects the first three vowel formants, emphasis spread in Qatari Arabic appears to be more of a phonetic rather than a phonological process. This is evidenced by the absence of VOT shortening and categorical conversion of plain [t] to emphatic [tˤ], indicating that emphasis spread operates as a phonetic phenomenon influenced by coarticulation. This interpretation challenges the notion of VOT as the primary acoustic correlate of emphasis in Qatari Arabic, as suggested by (Sakr, 2023).

In Kulikov (2021), VOT was found to be most relevant for voicing and F2 was mostly associated with emphasis. Al-Khairy (2005) discovered that the formant frequencies of subsequent vowels, particularly F2, were crucial in distinguishing the emphatic contrast, with F2 being significantly lower in vowels adjacent to /sˤ/ compared to /s/ (1288 Hz vs. 1603 Hz, respectively). This finding is consistent with results from earlier research (e.g., Abu-Al-Makarem & Cooper 2006; Jongman & Al-Masri, 2011; McCasland, 1979; Shadle, 1985) that demonstrated acoustic measurements can differentiate between fricative consonants. Furthermore, Al-Tamimi (2017) conducted a study on the plosive /dˤ/ and its plain counterpart /d/ in Jordanian and Moroccan Arabic dialects, exploring the linkage between epilaryngeal constriction and pharyngealization using multiple acoustic indicators (such as voice quality, formant information, and formant distance) to characterize pharyngealzation. He observed that the data from the Jordanian dialect were more indicative of an upper-mid pharyngeal constriction, whereas the Moroccan dialect data suggested a lower-mid pharyngeal constriction.

2.5.2.2 Summary of the Aforementioned

The aforementioned studies primarily focused on investigating the effect of emphasis on the second formant (F2) of vowels, consistently indicating F2 reduction in vowels adjacent to emphatic consonants, thereby distinguishing them from non-emphatic consonants. However, the literature predominantly examined emphatic plosives, with limited attention given to emphatic fricatives, as noted by Zawaydeh (1999) and Al-Khairy (2005). Furthermore, most studies were restricted to one or two Arabic dialects for comparison. Thus, this study will utilise 8 Arabic dialects to compare their emphatic contrast s~sˤ across different acoustic measurements, such as Center of Gravity (COG), which reflects the weighted average frequency of energy in the sound spectrum; Intensity, indicating the loudness or energy of the sound; Peak location, representing the frequency with the highest amplitude within the spectrum; and the Duration of the frication, referring to the length of the fricative sound. Additionally, formant information of the vowel following the emphatic contrast was analyzed to capture vocalic effects associated with the consonantal distinction.

2.5.2.3 Vowel Quality Impact on Second Formant (F2)

While emphasis generally lowers the second formant (F2) of the following vowel, the extent of this decrease is influenced by both the vowel's inherent sound quality and its length. Studies by Card (1983), Alioua (1995), and Yeou (1997) show variations in F2 lowering (measured at the vowel midpoint) for different vowels like /a/, /i/, and /u/. The greatest decrease occurs with the low front vowel /a/, followed by /i/ and /u/. In fact, the F2 of /œ/ can drop considerably in an emphatic context that it becomes a completely different vowel sound, the low ack vowel [a]. This significant change in sound quality might explain why many studies have focused solely on the /œ/ vowel when investigating the effect of emphasis on vowels. Additionally, Yeou (1997) found that short vowels following emphatic consonants tend to have a greater F2 decrease at their midpoint compared to long vowels. However, directly comparing the emphatic effect on long and short vowels in these studies remains an onerous endeavour owing to limitations in the measurement points. F2 measurements were only taken at the midpoint, not throughout the entire vowel duration, making it difficult to draw definitive conclusions about the overall impact of emphasis on vowels of different lengths. Thus, the current study will examine all vowel quality types, utilising ten measurements points, to account for the nature of the vowel formant information following the emphatic contrast.

2.5.2.4 Emphasis Impact on First (F1) and Third Formants (F3)

In the research field, studies identifying acoustic correlates other than F2 often find associations with F2 as well. Laufer and Baer (1988), in their study of Arabic and Hebrew, observed that pharyngeal constriction was invariably linked with a decrease in the second formant (F2) and an increase in the first formant (F1). They referenced preliminary experiments conducted in 1981 using an articulatory synthesizer to demonstrate that their findings in 1988 corroborated the acoustic effects of pharyngeal constriction witnessed in their current study, specifically the reduction of F2 and elevation of F1. Similarly, Al-Tamimi and Heselwood (2011) in their study on the articulatory and acoustic properties of emphatic coronals in Jordanian Arabic, conducted with nine speakers, found that vowels near emphatic coronals exhibited a higher F1 and lower F2 compared to vowels adjacent to plain coronals.

Although research has focused mainly on the second formant (F2) of vowels to pinpoint the location of a constriction in the back of the throat (posterior constriction), the patterns of the first (F1) and third (F3) formants might also offer valuable clues. While both uvular and pharyngeal constrictions lower F2 and raise F1, simulations suggest differences between them. F2 would be lower for uvular constrictions, but F1 would not rise as much compared to pharyngeal constrictions. Conversely, F3 would be lower for pharyngeal constrictions (Klatt & Stevens, 1969; Alwan, 1986). However, few studies have incorporated F1 or F3 measurements, and the findings are mixed; Alioua (1995), Yeou (1997), and Zawaydeh (1999) report an increase in F1, suggesting a pharyngeal constriction. Similarly, Kriba (2010) found that emphasis generally increased F1 and F3 frequencies while decreasing F2, with variations depending on the vocalic context. However, Card (1983) and Norlin (1987) found no consistent effect of emphasis on F1 or F3. Specifically for F3, it was found not to have a consistent pattern since it was found higher in some studies and lower in some others (see McCarthy, 1994; Watson, 2007).

2.5.2.5 Vowel Dynamic of Formant Information

Studies focusing on Arabic as a first language (L1) have predominantly concentrated on the static acoustic attributes of vowels, with limited exploration of dynamic features impacting monophthongal vowel classification. Notably, Al-Tamimi (2007a, 2007b) delved into these dynamic aspects within Jordanian and Moroccan Arabic dialects, as well as French, examining both production and perception. Utilising linear and polynomial regression analyses to model vowel transitions, Al-Tamimi identified that dynamic properties provided a nuanced differentiation, effectively discerning vowels across and within dialects. This methodology substantially enhanced classification accuracy, improving discrimination between Arabic dialects and French by 10–30%. Moreover, dynamic analysis achieved classification rates of 85.68% for Moroccan Arabic and 88.6% for Jordanian Arabic, signifying a 5–8% increase in precision and highlighting the significance of dynamic features in vowel categorisation (Al-Tamimi 2007b).

In another study, Almurashi et al. (2020) delved into Vowel-Inherent Spectral Change (VISC)[3] models—such as offset, slope, and direction models—for the F1, F2, and F3 of Hijazi Arabic (HA) vowels, spoken in various cities within Saudi Arabia, including Jeddah

[3] Vowel-Inherent Spectral Change (VISC) is defined by Nearey and Assmann (1986) as the "relatively slowly varying changes in formant frequencies associated with vowels themselves." This concept highlights the dynamic acoustic trajectories of vowels, which evolve over their duration and contribute to vowel identity.

and Makkah. This dialect encompasses both long and short vowels from Modern Standard Arabic/Classical Arabic and features unique long mid vowels /e:/ and /o:/, originating from diphthongs. Analyzing vowels within /hVd/ syllables in a carrier sentence, the research revealed significant spectral shifts for all HA vowels, indicating greater articulatory flexibility compared to languages with more vowels in the vowel-space. The slope model demonstrated substantial variation, with back vowels exhibiting rising slopes for F2, contrasting with the falling slopes of front vowels. Direction models proved effective in distinguishing tense/lax vowels, suggesting a qualitative distinction alongside the traditional length differentiation in Arabic vowels. Discriminant analysis underscored the superiority of the three-point model over two-point and static models in classifying HA vowels, achieving an average classification rate of 95.5%. This study emphasised the importance of vowel duration, potentially even surpassing F3, for accurate vowel classification, shedding light on the intricate nature of HA vowel articulation and the efficacy of examining internal vowel transitions for linguistic analysis.

In the recent study by Sakr (2024), an examination of the emphatic contrasts $s{\sim}s^{ʕ}$ and $t{\sim}t^{ʕ}$ within Central Mount Lebanon Lebanese was conducted. Sakr conducted an analysis encompassing both static and dynamic facets of vocalic articulation surrounding emphatic sounds. The static examination involved the measurement of the second formant (F2) of vowels at the midpoint both preceding and following the emphatic contrasts. In contrast, the dynamic analysis encompassed the extraction of eleven measurements at intervals of ten percent throughout the entire duration of each vowel segment. This comprehensive process was facilitated by analyzing vowel trajectories utilising Generalised Additive Mixed Models (GAMMs). The results revealed a notable reduction in F2 frequency at the midpoint for both the stop and fricative contrasts in emphatic contexts compared to their non-emphatic counterparts across various vowel qualities. Dynamically, the study uncovered clear indications of emphatic anticipation; specifically, vowels preceding plain consonants and followed by the emphatic fricative /$s^{ʕ}$/ exhibited a consistently lower F2 throughout the vocalic interval compared to other similar vowel contexts. Although the emphatic /$d^{ʕ}$/ consistently presented with lower F2 values than its plain counterpart, the differences were not substantial enough to preclude overlap between the two. Another study conducted by (Almurashi et al., 2024) investigated the dynamic and static aspects of vowel production in Hijazi Arabic, leveraging Vowel-Inherent Spectral Change (VISC) for a deeper understanding of vowel systems. Via an examination of monophthongal vowels in various consonantal contexts, the study emphasises the importance of dynamic cues, such as vowel duration, F0, and F3, in addition to traditional static measurements (F1 and F2) for vowel classification. Through acoustic analysis and discriminant analysis, it was found that dynamic models, especially a seven-point measurement model, provided higher classification accuracy compared to static models. This underscores the significant role of dynamic vowel specifications in linguistic analysis, contributing valuable insights into the phonetic and phonological characterisation of vowels in Hijazi Arabic.

2.4.2.6 Summary of the Aforementioned

The preceding studies employed a dynamic approach to characterize the formant information of vowels adjacent to emphatic contrasts. Al-Tamimi (2007a, 2007b) used two measurement points across two Arabic dialects, while Sakr (2024) and Almurashi et al. (2024) implemented 11 and 7 measurement points, respectively, each focusing on a single Arabic dialect. Despite the utilisation of multiple measurement points in these studies, the current thesis will examine ten measurement points across eight Arabic dialects. This approach aims to identify any typological patterns that these dialects may exhibit.

2.5.3 **Acoustic Measurement of Consonant**

Acoustic properties of emphatic consonants themselves have received scant attention in the literature. For stop consonants, Card (1983) reports F2 values taken from spectrograms, but it is not clear at which point (e.g., at release burst or onset/offset of the formant transition) these measurements were taken. Her data (not subjected to statistical analysis) suggest that F2 in emphatic consonants may be lower than in their plain counterparts. For fricatives, Card focused on the bottom cut-off frequency of the frication noise on spectrograms. Card found no correlation between the cut-off frequency and the presence versus absence of an emphatic fricative and therefore did not report this measure. Kahn (1975) similarly reports no difference between plain /s/ and emphatic /sˤ/ when measuring the bottom cut-off frequency. In his study on Moroccan speakers, Yeou (1997) examined the measurement of locus equations for the purpose of distinguishing between the emphatic /sˤ/ and its plain counterpart /s/ and he found that this measure could distinguish between s and sˤ, with latter having the flattest locus equation slopes. Similarly, the study conducted by Embarki et al. (2011) examined the coarticulatory effect of emphatic contrast between Modern Standard Arabic (MSA) and several dialects of Arabic regions (DAR), involving sixteen participants from Yemen, Kuwait, Jordan, and Morocco. They analyzed locus equation acoustic measurements and found agreement between MSA and DAR, showing that emphatic consonants (/tˤ, dˤ, sˤ, ðˤ/) had a significant effect on slope ($F(3, 63) = 4.86$, $p < .01$), with a flatter slope compared to their plain counterparts (/t, d, s, ð/) ($F(3, 63) = 1.23$, $p = .304$). However, Norlin 1983 found that only peak location could distinguish between s and sˤ while COG values' results were overlapping. He also conducted a study in (1987) on Egyptian speakers and compared the spectral moments correlates among the fricatives (/sˤ, s, z, zˤ/). For this examination, Norlin employed FFT (Fast Fourier Transform) spectra derived from a 26.5 ms window obtained after the initial third of the duration of the fricative. It is noteworthy that Norlin's analysis suggests emphatic /sˤ/ and /zˤ/ in Egyptian Arabic exhibit a concentration of energy at lower frequencies in comparison to plain /s/ and /z/. However, the t-tests conducted by the current researchers on the means (Norlin, 1987) suggest that these comparisons lack statistical significance. Moreover, Al-Khairy (2005) conducted an investigation involving eight male speakers of the Saudi dialect. The study employed spectral measurements such as spectral peak location and spectral moments, along with amplitude and temporal measurements including absolute and normalised frication noise duration.

He applied Fant's source filter theory, proposed in 1960, to interpret the findings related to fricative consonants. According to this theory, speech production involves two distinct components: a source signal (such as the glottal source or noise generated in a compressed state in the vocal tract) and a filter (reflecting resonance in the vocal tract cavities beyond the glottis or constriction). His analysis revealed that spectral measurements, amplitude, and duration did not significantly differentiate between emphatic contrasts, particularly between /sˤ/ and /s/, although emphatics exhibited lower peak location and longer duration.

In a similar vein, Kriba (2010) investigated the acoustic realisation of emphatic consonants in Libyan Arabic (LA), marking the first such study for this dialect. The research, involving twenty native speakers, aimed to compare the acoustic patterns of LA emphatics with those of other Arabic dialects and elucidate their articulatory basis. The study analysed F1, F2, and F3 of vowels (initial three formant frequencies) following plain and emphatic consonants, in addition to locus equations, intensity, and duration measurements of the emphatic contrasts s~sˤ.

The findings revealed that emphasis typically elevated F1 and F3 frequencies while reducing F2, although variations were observed based on the vocalic context. However, there were no notable differences in intensity or duration between the fricative sounds /sˤ/ and /s/. More recently, Sakr (2024) delved into the phonetic subtleties that differentiate emphatic consonants in Central Mount Lebanon Lebanese. In this investigation, spectral moments[4], duration, and intensity served as the primary acoustic parameters for analysing these emphatic consonants. The findings reveal that the emphatic /sˤ/ exhibits lower first and second spectral moments than the plain /s/; Specifically, the emphatic /sˤ/ demonstrates a lower centre of gravity in its spectral distribution, which contrasts with the higher spectral centre of the plain /s/. However, the difference was not found to be statistically significant. Moreover, Sakr (2024) found a notable difference in duration and intensity between the emphatic contrast, with the emphatic /sˤ/ having a longer average duration and higher average in intensity. Again, the study found no significant difference in the duration and intensity levels of the emphatic /sˤ/ and its plain counterpart /s/.

2.6 Studies on Perception

While extensive literature exists on the acoustic analysis of emphasis in Arabic, there is a notable dearth of studies focusing on the perceptual cues associated with emphasis. Jongman et al. (2011), for instance, conducted a study investigating the perception of emphasis (tˤ & sˤ) compared to their non-emphatic counterparts (t & s) using crossed-spliced CVC syllables. In this experimental setup, emphatic and plain consonants were spliced onto plain or emphatic syllable portions, respectively. The findings indicated that words consisting of emphatic target consonants elicited significantly higher emphatic responses when compared with those with plain consonants. Moreover, words with emphatic VC/C̣Ṿ aspects elicited significantly more emphatic responses compared to those with plain VC/CV parts. Notably, having an emphatic ṾC̣ or C̣Ṿ portion resulted in a significantly greater number of emphatic responses compared to having an emphatic consonant alone. These findings indicate that the vowel element exerts a more pronounced influence on the perception of the distinction between emphatic and plain sounds than the consonant itself.

Adopting a different approach, Obrecht (1968) utilised synthetic stimuli to investigate how F2 transition affected the Lebanese dialect, with an emphasis on the transition from an initial consonant to the subsequent vowel. Employing the Haskins pattern playback synthesizer, Obrecht created stimuli modelled after a male speaker of Lebanese Arabic. The study analysed various segments within a monosyllabic word context, systematically varying the

[4] Spectral moments are acoustic correlates used to describe the distribution of energy within a sound spectrum, including measures like mean (center of gravity), variance (spread), skewness (asymmetry), and kurtosis (peakedness) (Johnson, 2012).

F2 value in 120 Hz increments ranging from 1080 to 1800 Hz. The total duration of the syllable ranged from 300 to 400 ms, with the F2 transition duration fixed at 70 ms. Additionally, while F3 remained constant at 3000 Hz, F1 varied based on the vowel used in the experiment. The findings derived from the tˤi: - ti:/ continuum indicated that the perception of the contrast between emphatic and non-emphatic sounds was categorical. Specifically, the endpoints of the continuum were consistently perceived to be /tˤi:/ characterised by low F2 values and as /ti:/ distinguished by high F2 values. The turning point between the emphatic contrast was identified at around 1560 Hz of the F2 value. In other words, while F2 values were about 1560Hz, listeners perceived the stimuli as plain contrast; however, when F2 values were lowered, they categorised the stimuli as emphatic. Yeou (1995) replicated Obrecht's experiments on Moroccan dialect of Arabic with emphatic fricative contrast, /sˤi: and si:/. The results were similar to the Obrecht's. Yeou confirmed the role of F2 in a synthetic /sˤi: and si:/ continuum, further demonstrating that F1 alone was not a perceptual cue to the emphatic/plain distinction.

Ali and Daniloff (1974) conducted a study on naturally spoken Iraqi words, focusing on minimal pairs of emphatic and non-emphatic consonants. They observed that when the target consonant was missing from truncated stimuli, listeners accurately identified it approximately 68% of the time (62% for plain stems and 73% for emphatic stems). The findings indicated that the presence of the emphatic consonant itself was not crucial for fairly accurate perception of the emphatic and non-emphatic contrast. Interestingly, factors such as vowel quality, consonant type, and position of the target consonant did not seem to have an impact on perception. This suggests that the plain and emphatic words' vocalic segment contained adequate cues for listeners to discern the contrast. Therefore, the formant details, in conjunction with the spread of emphasis, played a significant role in guiding listeners' judgments regarding which member of the emphatic contrast to choose.

While some previous studies have explored the sociolinguistic aspects of emphasis perception, examining factors such as social selection and gender (Kahn, 1975; Alwan, 1986; Wahba, 1993; Al-Wer, 2000), they did not primarily focus on perceptual cues. For instance, Kahn (1975) found no significant difference in the perception of emphatic contrast between male and female participants. These studies represent a limited subset that delves into the perceptual correlates of emphasis.

Certain studies have suggested that the F2 transition serves as a reliable perceptual cue, facilitating accurate discrimination of the emphatic contrast. However, listeners have also been found to discern the distinction even in the absence of explicit F2 transition information. Other investigations utilised crossed-spliced syllables to explore the emphatic contrast perception without synthesising specific stimuli. Although many studies reported perceptual findings consistent with their acoustic counterparts, some disparities exist, possibly due to variations in sounds (consonants and vowels) and dialects across different studies.

Overall, it can be inferred that the reviewed literature highlights the tendency of some studies to focus solely on comparing emphasis perception within a single Arabic dialect, while others concentrate primarily on examining emphatic plosives. Moreover, certain studies centred their attention on the consonant itself or the entire word as a perception unit, particularly noting the F2 lowering of the following vowel. However, none of these studies have thoroughly investigated the characteristics of the vowel trajectory, including different height and slope variations. While focusing on consonants offers valuable insights into emphasis, relying solely on the formant information of subsequent vowels may not fully capture the essence of the emphatic contrast. Furthermore, there is a notable absence of studies that manipulate both the height and slope of the F2 vowel trajectory to explore the distinction in emphatic contrast.

The present investigation aims to contribute novel insights into the relative significance of consonantal and vocalic cues in distinguishing between emphatic and non-emphatic sounds. Specifically, the study endeavours to investigate the emphatic fricative /sˤ/ and its plain counterpart /s/ perceptually across four Arabic dialects. It will use a resynthetic stimuli, created by the author to explore the participants responses toward the emphatic and non-emphatic fricative. Two vowel qualities (/a:/ and /i:/) will be used in the present experiment and their F2 trajectories will be manipulated. It's worth noting that F1 and F3 will not be analyzed in this context due to space constraints and the fact that they are not the primary focus of this study.

2.7 Dialects of Arabic

The categorisation of Arabic dialects can vary based on several factors, including historical and contemporary perspectives, cultural contexts, research objectives, and geographical distribution. It is common to encounter Arabic dialects named after countries or major cities, such as Jordanian Arabic, Moroccan Arabic, Egyptian Arabic, and Syrian Arabic in academic discourse and language learning materials. This naming convention is partly justified by the tendency of speakers within a country to converge towards the dialect spoken in major cities or capitals, often perceived as prestigious. One traditional classification divides Arabic dialects into Eastern (mašriq) and Western (maġrib) groups, with the Eastern group spanning from Saudi Arabia to Egypt and the Western group encompassing North African countries like Morocco and Mauritania. This division, extending from the Mediterranean coast along Egypt's western border down to Lake Chad, is based on shared linguistic features within each group.

Another classification proposed by Jastrow (2002) delineates geographic zones reflecting the origin and spread of Arabic, including the Arabian Peninsula (pre-Islam), areas influenced by Islamic conquests, and regions where Arabic arrived through trade, such as parts of Africa and Asia. Dialects from the Arabian Peninsula are considered archaic due to the historical shift in political and administrative focus to newly conquered Islamic territories. Ingham (1982) and Palva (1991) offer a classification of central Arabian dialects into three categories, distinguishing among dialects spoken by various tribes and in different regions, from the northeastern Arabian Peninsula to the Ḥijāz area.

The distinction between Bedouin (Badawī) dialects and sedentary (ḥaḍarī) dialects is another classification method, with Bedouin dialects believed to retain more ancient linguistic features compared to urban (madanī) and rural (fallāḥī/qarawī) sedentary dialects. However, the presence of Bedouin features in sedentary dialects outside the Peninsula complicates this dichotomy, suggesting that certain linguistic features associated with Bedouin dialects may not always indicate conservatism.

Religious factors also influence dialect classification, with variations observed among Muslim, Christian, and Jewish communities in Iraq, as well as between Sunni and Shiʕite Muslims in Bahrain. In Jordan, dialects are classified into urban, rural, and Bedouin types, with further subdivisions within the Bedouin category. However, this nuanced classification raises questions about grouping rural and Bedouin dialects of Jordan under a single category.

3 Chapter 3: The Acoustic Study

3.1 Introduction

In this chapter, the acoustic results of this study will be presented to investigate the main research question "Is the phonetic realisation of the phonological contrast between /s/ and /sˤ/ the same across Arabic dialects?". According to the literature review, the majority of previous studies have mainly focused on comparing the fricatives production only to one or two Arabic dialects. Additionally, a huge body of previous research has focused on examining emphatic plosives and it was found that there is a typological divide which classified Arabic dialects into two groups: some dialects show an acoustic difference in the plosive (VOT) as well as in the following vowel, but other dialects only show a difference in the following vowel (Bellem, 2014). This current study, however, will examine the emphatic fricative /sˤ/ and its plain counterpart /s/ acoustically across a set of eight Arabic dialects, which to the best of the author knowledge no single study has done yet. Additionally, it will see whether we can find a similar typological divide as was found with the plosive consonants. To achieve this aim, multiple acoustic correlates were utilised, 1 for fricative, the duration of the frication, the intensity, the spectral properties which includes Centre of Gravity COG, and spectral peak location. 2 for vowel, the first three formants, F1, F2, and F3.

3.2 Prediction and Hypothesis

For Fricative duration: This measure is expected to be one of the most important metrics, because emphatic fricatives have a feature of co-articulation where there are two sources of energy: the primary one (the vocal tract's turbulence) and the secondary one (the extra energy that the pharyngeal area contains). The actions occurring at the posterior region of the oral cavity result in a reduction of airflow velocity compared to the airflow observed during the production of non-emphatic fricative sounds. This interpretation was influenced by the source-filter theory proposed by Fant (1960). Consequently, it is predicted that the emphatic fricative /sˤ/ will exhibit a longer duration than its non-emphatic counterpart /s/. Furthermore, as discussed in the literature review in Chapter 2, several studies have noted noticeable differences between the emphatic and non-emphatic contrasts, with the emphatic /sˤ/ showing longer frication noise duration compared to the non-emphatic /s/, although the duration disparity was not deemed statistically significant (refer to Al-Khairy, 2005; Kriba, 2010; Sakr, 2024).

For Fricative Intensity[5]: This correlate is expected to be crucial because the energy for the plain fricative sound is expected to be different than it is in the emphatic fricative. Plain /s/ is predicted to have higher amplitude at its, onset when compared to emphatic /sˤ/. However, the overall intensity for plain /s/ is expected to be lower in dB than that of emphatic /sˤ/ due to the fact that the co-articulatory feature found in /sˤ/ is expected to have another region where the energy is increasing, particularly at the mid or the offset points. This prediction is aligned with what (Sakr, 2024) found in his study although the observed increase was statistically insignificant.

For COG in Fricative: It is one of the most important spectral measurements used in the analysis of fricatives (Gordon et al., 2002; Jongman et al., 2000). COG provides a measure of the average frequency in a sound spectrum, weighted by the intensity of each frequency component. Essentially, it pinpoints the "balance point" of energy in the spectrum, indicating whether the energy is concentrated in higher or lower frequencies. COG is calculated by multiplying each frequency by its corresponding intensity and then dividing the sum by the total intensity, showing where the majority of the noise energy is centred. Unlike other spectral moments, such as standard deviation, skewness, and kurtosis, which describe the spread, asymmetry, and shape of the energy distribution, COG directly captures the central tendency of the energy distribution. Both fricatives, plain /s/ and emphatic /sˤ/ share the same place of articulation in the anterior region of the vocal tract, but /sˤ/ involves a secondary articulation which involves a retraction of the tongue toward the pharynx. This co-articulation shifts acoustic energy towards lower frequencies, resulting in a lower COG for /sˤ/. In contrast, **/s/** has its energy concentrated in higher frequencies, resulting in a higher COG. The spectra in Figure 7 visually confirm this difference. For /s/, the energy is concentrated in the upper frequencies, leading to a higher COG, while for /sˤ/, the energy is concentrated in the mid-to-lower frequencies, resulting in a lower COG. This illustrates how emphasis in /sˤ/ shifts the energy downward, compared to the plain /s/.

[5] In this study, I measured **absolute intensity** rather than relative intensity to the adjacent vowel. Although I acknowledge that relative intensity might help control for factors such as speaker loudness and microphone distance, the **controlled recording conditions** used (same speaker, microphone setup, and recording environment) minimized the impact of these variables. Absolute intensity provides an informative baseline for comparing the energy differences between plain and emphatic fricatives, which was the focus of this research. Future studies may explore relative intensity measurements to provide additional insights.

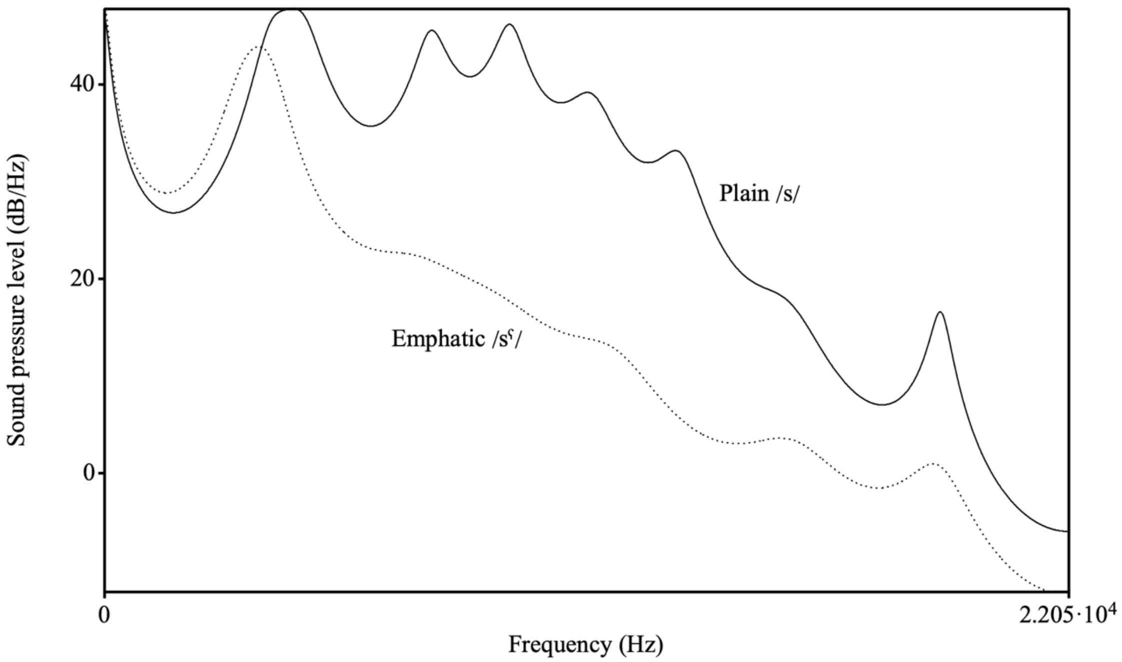


Figure 7 Spectral slices for /s/ (solid line) and /sˤ/ (dotted line) taken from the midpoint of the recorded tokens of a Kuwaiti male speaker. The slice for /s/ shows higher energy concentration in the upper frequencies, while the slice for /sˤ/ shows energy concentrated in the lower frequencies.

For Fricative Peak location: This measure reflects the relationship between the place of articulation for fricatives and the frequency at which the maximum energy in the frication noise occurs. The peak energy may appear at lower or higher frequencies depending on the articulatory properties of the sound. For the plain fricative /s/, the maximum peak is expected to occur at a higher frequency, whereas for the emphatic fricative /sˤ/, the peak is expected to occur at a lower frequency. This difference is due to the emphasis (secondary articulation) in /sˤ/, which involves tongue retraction toward the pharynx, altering the resonance characteristics of the vocal tract and shifting the energy downward (see Norlin, 1983; Al-Khairy, 2005).

For the vowel formant information: With reference to the perturbation theory (see Carré and Mrayati, 1992; Chiba and Kajiyama; 1941; Fant,1960) the pharyngeal constrictions lead to a specific acoustic pattern characterised by increased F1 and a decline in F2. This mechanism would trigger the F2 lowering F1 and F3 increasing. Many studies have reported the same results (the list of studies is found in the literature section 2.4)

This chapter presents the acoustic findings regarding the articulation of the target consonants s~sˤ and their adjacent vowels /a, i, u/ across eight Arabic dialects. The chapter begins with an overview of the methodology and data analysis procedures, followed by the presentation of the results and concluding with a discussion of the findings. It is important to highlight that the data utilised in this chapter is sourced from The Intonational Variation in Arabic (IVAr) project, initiated by the Department of Language and Linguistic Science at the University of York (Hellmuth & Almbark, 2019). The primary objective of the IVAr project is to compile an openly accessible corpus of audio recordings featuring a diverse range of speaking styles from 12 speakers representing each of the eight spoken Arabic varieties.

3.3 Method

3.3.1 Participants

The research identified eight distinct Arabic dialects based on the speech patterns of twelve participants, comprising six males and six, with the exception of the Syrian dialect, which includes three participants of each gender, and the Iraqi dialect, which includes six females and four males, for inclusion in the data analysis. These dialects encompass Egyptian, Jordanian, Kuwaiti, Syrian, Iraqi, Moroccan, Omani, and Tunisian varieties, as outlined in Table 1 below.

code	*dialect*	*F*	*M*	*code*	*dialect*	*F*	*M*
egca	Egyptian (Cairo)	5	7	irba	Iraqi (Muslim Baghdadi)	6	4
joka	Jordanian (Karak)	6	6	moca	Moroccan (Casablanca)	6	6
kwur	Kuwaiti (Urban)	6	6	omba	Omani (Buraimi)	6	6
syda	Syrian (Damascus)	3	3	tuns	Tunisian (Tunis)	6	6

Table 1 Dialects represented in the corpus and number of female/male speakers in each

Each participant selected for the study met specific criteria, ensuring they were born and raised in the designated region or city, with the majority also being current residents. Notably, all speakers originating from Baghdad and Damascus were residents of Amman, Jordan, which is where the recordings were undertaken. Data collection for this paper was conducted alongside the main IVAr project, involving fieldwork in North Africa and the Middle East. A total of 88 participants were recruited, with most dialects represented by at least 12 speakers, although Syrian and Iraqi dialects had smaller participant numbers (6 and 10, respectively). Participants' ages ranged from 18 to 35 years, and further metadata can be accessed from the IVAr corpus (Hellmuth & Almbark, 2017).

3.3.2 Materials

To facilitate comparison across emphatic contrasts, a set of target words was compiled, featuring two fricatives: the emphatic /$s^{ʕ}$/ and the non-emphatic /s/. The methodology for data collection involved systematically varying both the length and quality of the vowel following the target consonants in each word (/a i u a: i: u: /), as well as the position of the target consonant (word-initial or word-medial). The aim was to establish minimal pairs across the emphatic contrast wherever feasible. Table 2 presents a compilation of stimuli, which includes a combination of common and uncommon real words, due to the unavailability of common real-word minimal pairs in some instances. This resulted in 88 speakers producing 20 stimuli each, with one repetition per stimulus, yielding a total of 1,760 tokens.

The Plain fricative /s/			***The Emphatic fricative /$s^{ʕ}$/***				
Arabic	*item*	*status*	*Arabic*	*item*	*status*	*position*	*Vowel*
سين	siin	R/uc	صين	$s^{ʕ}$iin	R/uc	initial	i:
سور	suur	R/c	صور	$s^{ʕ}$uur	R/c	initial	u:
سار	saar	R/c	صار	$s^{ʕ}$aar	R/c	initial	a:
سلسال	silsaal	R/c	صلصال	$s^{ʕ}$il$s^{ʕ}$aal	R/c	initial	i
سُبّ	subb	R/c	صُبّ	$s^{ʕ}$ubb	R/c	initial	u
سَدّ	sadd	R/c	صَدّ	$s^{ʕ}$add	R/c	initial	a
يسيبو	yisiibo	R/c	يصيبو	yi$s^{ʕ}$iibo	R/c	medial	i:
نَسَفو	nasafo	R/uc	نَصَفو	na$s^{ʕ}$afo	R/uc	medial	a
مسابح	masaabiħ	R/c	مصابح	ma$s^{ʕ}$aabiħ	R/c	medial	a:
محسودين	maħsu:di:n	R/c	محصودين	maħ$s^{ʕ}$u:di:n	R/c	medial	u:

Table 2 List of items used in the study for all dialects
* **Real/uc** : uncommon real word ***Real/c** : Real common word

3.3.3 **Procedure**

A local fieldwork assistant, fluent in the target dialect, oversaw each recording session. The assistant was selected based on their native language being the specific dialect under investigation. Recordings were made using a Marantz PMD661 solid-state data recorder in digital format (.wav) at 44.1kHz 16 bit, leveraging Shure SM10A-CN head-worn dynamic cardioid microphones. To introduce the target words, a carrier phrase was employed, with all phrases translated into English as ‘write ___ one more time’. Each target dialect adhered to an informal orthographic norm in presenting the carrier phrase, encouraging participants to deliver their utterances in the colloquial register, rather than a formal one like Modern Standard Arabic. Table 3 outlines examples of these carrier phrases. Target utterances, interspersed with distractor utterances, were printed in pseudo-random order on paper sheets in Arabic script.

Dialect	**Target word**	**Carrier phrase in Arabic script**	**Target sound**
Egyptian	su:r “wall”	**اكتب سُوُر كمان مرة** “write **su:r** one more time”	/s/
Iraqi	$s^ʕ$add ” hold off”	**اكتب سَد بعد مرة** “write Sadd one more time”	/$s^ʕ$/

Table 3 Examples of carrier phrases used in the study

3.3.4 **Data analysis**

3.3.4.1 Segmentation and annotation

Orthographic transcriptions were aligned to the audio signal using the Prosody Lab Aligner tool in PRAAT, generating interval tiers at both the word and phone levels. Manual correction of alignment errors was performed based on the segmentation criteria outlined by Turk et al. (2006). Subsequently, PRAAT scripts were employed to identify the onset and offset of the target consonant, the offset of the target word, and the following vowel. The script also calculated the midpoint of the target fricative and multiple time points throughout the vocalic interval. The purpose of examining these multiple points in the vowel was to investigate potential acoustic differences in the vowel and to do this dynamically rather than only looking at the vowel midpoint.

3.3.4.1 Acoustic Analyses

3.3.4.1.1 Analysing Fricatives

All measurements were obtained following the segmentation criteria elucidated earlier; the duration of the fricative was calculated from the onset of frication noise to its offset, as illustrated in Figure 8 below. The same process was applied to the intensity measure. Regarding spectral analysis, fast Fourier transform (FFT) was utilised with a window length of 40-ms, double-Kaiser window being centred on the frication noise. Each spectrum was filtered using a pass-band Hann filter between two cut-off frequencies (7000Hz to 16000Hz), with a smooth value of 500Hz to remove low-frequency energy corresponding to background noise and/or voicing. To determine the spectral peak location measure, the window position was centred on the middle of the frication noise. This choice was informed by previous studies indicating that spectral peaks tend to occur at higher frequencies within the midpoint of frication noise (Al-Khairy, 2005; Behrens & Blumstein, 1988).

3.3.4.1.1 Analysing Vowels

3.3.4.1.1.1 Vowel Duration

Vowel duration was determined by measuring the time from the beginning of the vowel, marked by the appearance of the initial clear glottal pulse in both the periodic waveform and spectrogram, to the end of the vowel. The end of the vowel was identified as the point where the amplitude decreased, as illustrated in the Figure 8 below.

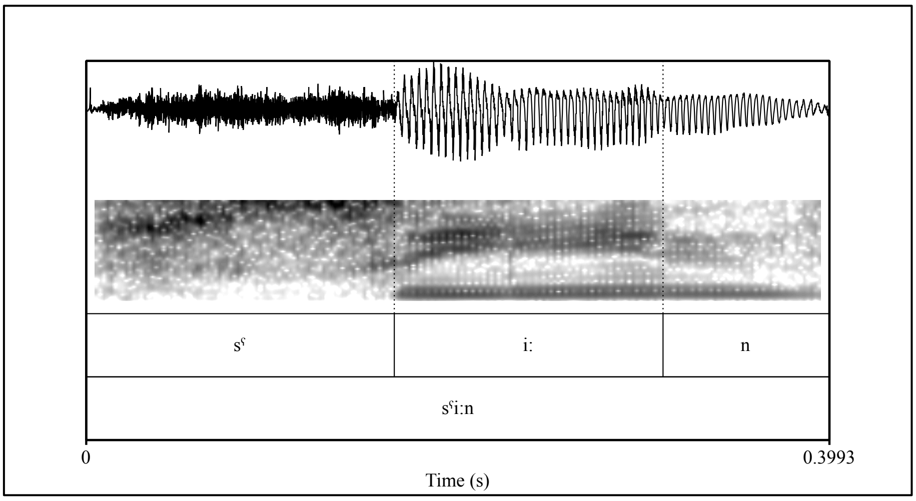


Figure 8 A sample showing how vowel duration was measured after an emphatic fricative

3.3.4.1.1.2 Vowel Formants

Formant frequencies (F1, F2, and F3) for each vowel (/a i u a: i: u:/) were obtained from the midpoint of each vowel, occurring after the target fricative. In addition, the first two formant frequency values (F1 and F2) were measured at ten equidistant points over the course of vowel's duration. The variation in formant frequency was analysed across these ten points to capture the trajectory of vowel patterns following the target fricatives. To observe the behaviour of the vowels, Generalised Additive Mixed Models (GAMMs) were employed at these ten points to quantify the time-varying formants, particularly F2. The objective was to explore the variation in formant patterns across all dialects, as vowels in an emphatic context typically exhibit a constriction resulting in a combination of high F1, low F2, and high F3 (Al-Tamimi, 2017; Kriba,2010; Yeou's, 2001; Jongman et al., 2007).

3.3.5 **Statistical Analysis**

3.3.5.1 Linear Mixed Effect Model

After obtaining the result, a PRAAT software was utilised for conducting the sample analysis (Boersma & Weenink, 2009), and the script was developed by the author. The results were modelled for the fricatives as well as the vowel formants (F1 and F2 at the mid-point) using a linear mixed-effects regression model (LMER) (Kuznetsova, Brockhoff, & Christensen, 2017) for each acoustic measure in turn as the dependent variable with this parallel model structure:

DV ~ condition * dialect * gender + vowel + (1 |speaker) + (1 | item)

importantly, all the dependant variables were tested against the same model structuring and aforementioned model was observed to be the best fit.

3.3.5.2 Fitting the Best Random Effects Model

In order to specify the random effects term that best fits the data, two different models were compared and tested using the "anova" function in R software (*R Core Team, 2020)* and the resulting output suggested that the model fit has been improved by adding both random effects, speaker, and item. Table 4 below shows that there is an improvement in between model 1 and model 2 and we can use model 2 as our random effects structure for the rest of the analysis.

nullmodel1: duration_fric ~ 1 + (1 | speaker)

nullmodel2: duration_fric ~ 1 + (1 | speaker) + (1 | item)

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
nullmodel1	3 -20860	20875	-10427	20854	-	-	-
nullmodel2	4 -20590	20610	-10291	20582	271.91	1	<2x10-^{16} ***

Table 4 ANOVA test results for comparing the random factors, speaker and item

3.3.5.3 Fitting the Best Fixed Effects Model

Here, we introduced specific predictors to the model and in order to reach for the fit model, multiple model structures were compared and tested by Likelihood Ratio Test (LRT). The first is the Null model that is constructed without a factor of the author interest and for the other four models, factors like, *dialect, condition, vowel type, and gender* were included one at a time. The "**condition**" predictor refers to the type of fricatives (emphatic or non-emphatic).

Nullmodel2: COG ~ 1 + (1 |speaker) + (1 | item)

1Predictor: COG ~ dialect + (1 |speaker) + (1 | item)

2 Predictor: COG ~ dialect + condition + (1 |speaker) + (1 | item)

3 Predictor: COG ~ dialect * condition + (1 |speaker) + (1 | item)

4 Predictor: COG ~ dialect * condition * gender + vowel + (1 |speaker) + (1 | item)

The resulting output suggested that adding the fixed effect "**dialect**" to the "**1Predictor**" model affected significantly the **COG** measurement of the fricatives ($\chi 2(1) = 16.541$, p= 0.02061). Thus, the two models are significantly different from each other (see Table 5). The model was also improved further by adding the predictor "condition" (see Table 6).

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
nullmodel2	4 -20590	20875	-10291	20582	-	-	-
1Predictor	11 -20587	20644	-10283	20565	6.541	7	0.02061 *

Table 5 ANOVA test results for comparing nullmodel2 and 1Predictor

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
1Predictor	11-20587	20643.6	-10283	20565.1	-	-	-
2Predictor	12 -20588	-5640.5	2863	-5726.1	26291	1	2.2e-16 ***

Table 6 ANOVA test results for comparing 1Predictor and 2Predictor

However, one of the aims of this study is to investigate the interactions between dialects, condition, and gender including to the vowel type so as to see if the condition is interdepended on any of the other factors or not. So, the model was improved by adding the rest of the factors of interest, leading to the best fit model "4 predictor". See Tables 7 and 8.

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
2Predictor	12-20588	20650	-10282	20564	-	-	-
3Predictor	19 -20584	20682	-10273	20546	18.069	7	0.01166 *

Table 7 ANOVA test results for comparing the interaction between 2Predictor and 3Predictor

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
3Predictor	19- 20584	20682	- 10273	20546	-	-	-
4Predictor	37 - 20521	20711	-10224	20447	99.085	18	3.257e-13 ***

Table 8 ANOVA test results for comparing the interaction between 3Predictor and 4Predictor

In order to account for physiological differences between male and female speakers, the formant extraction process in Praat was adjusted accordingly. Female speakers had their formants extracted with a maximum frequency ceiling of 5500 Hz, while male speakers had a ceiling of 5000 Hz. This adjustment ensures that the extracted formants accurately reflect the vocal characteristics of each gender, preventing any potential under- or over-estimation of formant frequencies. Vowel formant measurements in this study were not normalized because the dataset was largely balanced by gender, with a near equal number of male and female speakers. Additionally, male and female results were plotted separately and gender included as a fixed factor in all models. No normalization procedure was applied to the fricative measurements (e.g., COG, intensity, peak frequency), as there is currently no widely accepted method for normalizing fricatives.

3.4 Results

This section will present the model output for the predicted mean values of the fricative consonants, followed by the predicted mean values for short and long vowels. It is worth noting that all raw data figures, as well as tables detailing the mean and standard deviation for each acoustic measurement, are provided in the Appendices 9.3 – 9.11.

3.4.1 **Fricative Results**

3.4.1.1 Fricative Duration

The model fitted to the fricative duration midpoint values shows main effects of gender [β= -5.9, SE=134, p > 0.0001], dialect [β= -5.913, SE= 3.37, p > 0.001], and vowel type [β= -1.168, SE= 4.7, p > 0.02]. The results also indicate that distinguishing between emphatic contrasts using this metric is not effective, given that there is no significant effect of condition [β = -1.114e-05, SE = 3.483, p > 0.997]. See Figure 9 below, which illustrates the variation of this acoustic correlate across all dialects.

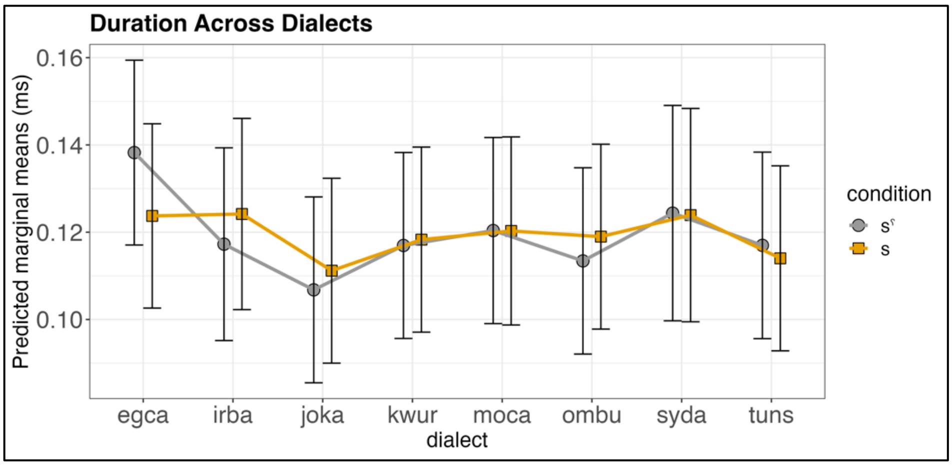


Figure 9 The 95% confidence intervals around the mean predicted values for fricative duration, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

Figure 10 shows that both male and female participants show a significant difference in the durations of the emphatic and the non-emphatic fricatives across all the dialects. Both egca and joka use different fricative durations more than the other dialects [β= 1.17, p > 0.0008], and [β= -1.033, p > 0.003]. The result also shows that there are two-way interactions between dialect and fricatives; egca and irba use greatly the duration measure to mark the s~sˤ contrast where the emphatic duration is higher than non-emphatic with egca [β= 7.269, SE= 1.634, p > 9.46e-06], but it is lower than non-emphatic with irba [β= -3.446, SE= 1.722, p > 0.05].

Syda dialect shows a three-way inter-dependence on fricative type, and gender. Thus, both male and female speakers use different fricative duration to differentiate between s~sˤ [β= 4.419, SE= 2.202, p > 0.05].

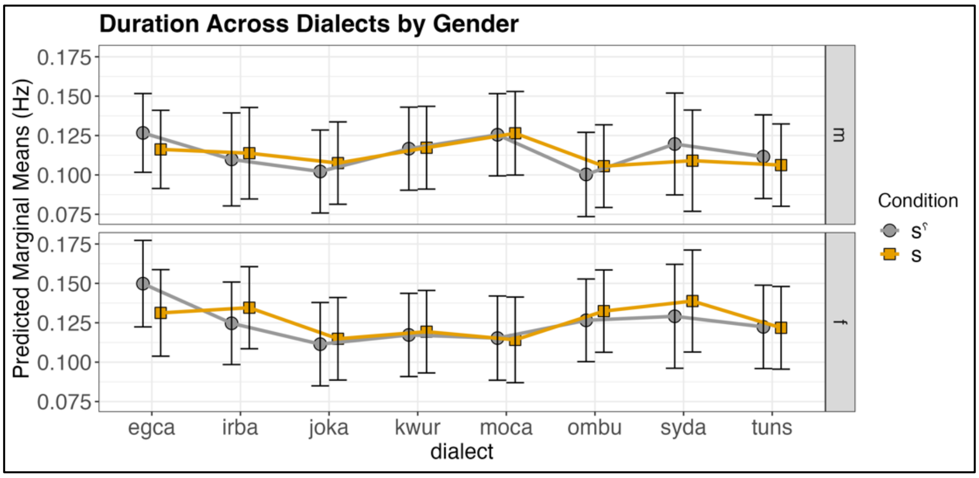


Figure 10 The 95% confidence intervals around the mean predicted values for fricative duration, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

3.4.1.2 Centre of Gravity (COG)

The LMER model indicates a main effect of gender, with male speakers across all dialects exhibiting lower COG values compared to females (β= -599.852, SE= 78.678, p > 2.82e-11). Additionally, there is another main effect observed in the egca dialect, where its speakers demonstrate the lowest COG values overall (β= -466.949, SE= 197.635, p > 0.0204). Conversely, Moroccan speakers exhibit the highest COG values across all dialects (β= 906.534, SE= 197.141, p > 1.39e-05). As the effect of vowel type, /a/ vowel shows higher COG values than vowels /i/ and /u/ [β= 425.416, SE= 74.795, p > 2.75e-05]. However, there is no effect of condition [β= -90.369, SE= 55.542, p > 0.1218]; this means that there is no significant difference between s~sˤ across all dialects. See Figure 11 below, which illustrates the variation of this acoustic correlate across all dialects.

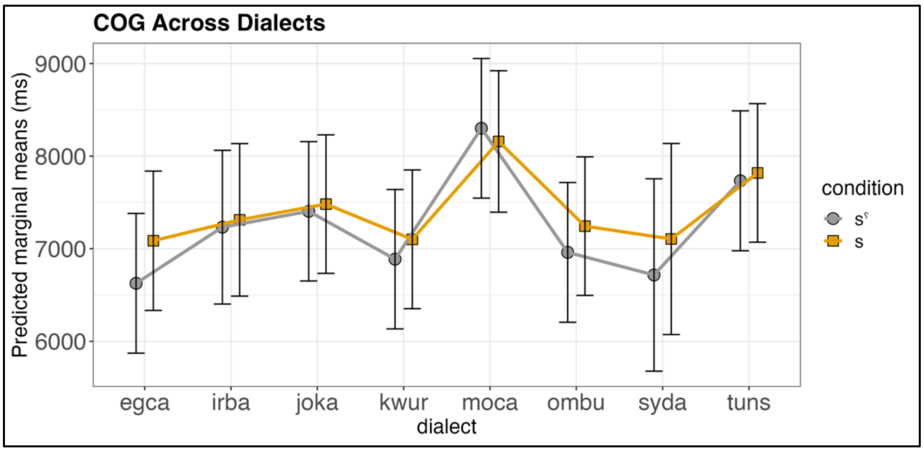


Figure 11 The 95% confidence intervals around the mean predicted values for COG, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

As depicted in Figure 11, the utilisation of this measure varies significantly between the egca and moca dialects compared to the others. There is a substantial effect observed in the moca dialect (β= 906.534, SE= 197.141, p > 0.0001), while a smaller effect is evident in the egca dialect (β= -466.949, SE= 197.635, p > 0.0204). Additionally, both dialects exhibit a two-way interaction with fricative type, indicating a significant utilisation of COG values to distinguish between the emphatic contrast (s~sˤ) (refer to Table 9). Another two-way interaction is also revealed by the prediction model results where Kuwaiti male and female speakers have different COG values of s~sˤ fricatives. Asfor the three-way interaction, egca male and female participants show different COG measures to distinguish between s~sˤ.

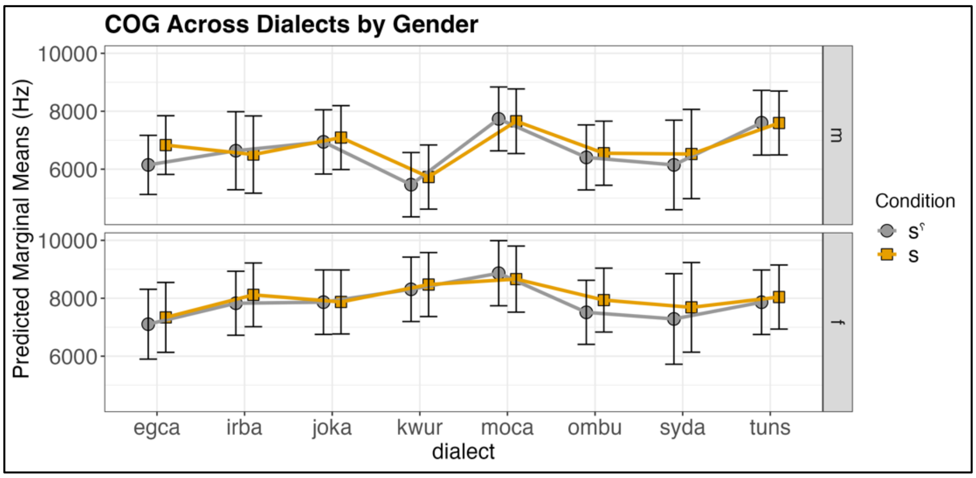


Figure 12 The 95% confidence intervals around the mean predicted values for COG, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

Interaction	Estimate	Std. Error	Pr(>|t|)
egca*fricative	-139.429	58.272	0.0169 *
moca*fricative	161.436	61.337	0.0086 **
Kwur*gender	-797.193	195.358	0.0001 ***
egca*fricative*gender	-115.496	56.785	0.0422 *

Table 9 LMER model output for the significant interactions among predictors

3.4.1.3 Intensity

According to the resulting output of the intensity measure, there are main effects of the gender [β= 1.45900, SE= 0.44451, p > 0.00144] and vowel type [β= 0.73556, SE= 0.20729, p > 0.00244]. The former indicates that male speakers of all dialects have higher intensity than females. Meanwhile, the latter shows that vowel /a/ is higher in intensity than other two vowels, /i/ and /u/. However, this measure does not distinguish between the emphatic contrast even though emphatic /sˤ/ show a slight lowering [β= -0.26293, SE= 0.15464, p > 0.10654]. See Figure 13 below, which illustrates the variation of this acoustic correlate across all dialects.

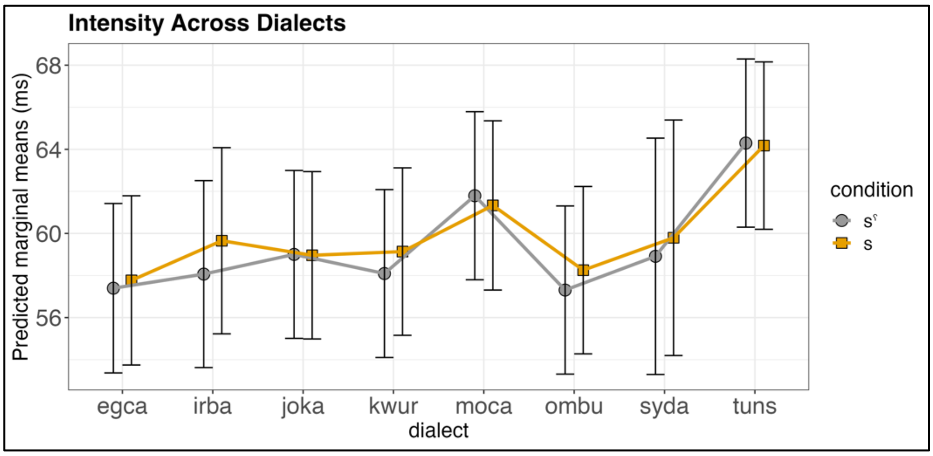


Figure 13 The 95% confidence intervals around the mean predicted values for intensity, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

Figure 14 reveals that the intensity mean values of the male and female participants in the dialects irba and joka are used differently than are observed in the other dialects [β= 2.53867, SE= 1.20972, p > 0.03872 and β= -3.46803, SE= 1.10125, p > 0.00223], respectively. Male speakers have higher intensity for the former while they have lower intensity for the latter. the result also shows that there are two-way interactions between fricative types and dialects and fricative types and gender. The former is seen in the dialects irba (where emphatic /sˤ/ is lower) and moca (where emphatic /sˤ/ is higher) since both of them use the value of the intensity to mark the contrast between s~sˤ [β= -0.53035, SE= 0.24261, p > 0.02902] and [β= 0.49319, SE= 0.23753, p > 0.03808], respectively. On the other hand, the latter shows that all male and female speakers use different intensity values to differentiate between the emphatic contrast, with the emphatic /sˤ/ being the lower [β= -0.22670, SE= 0.09133, p > 0.01320].

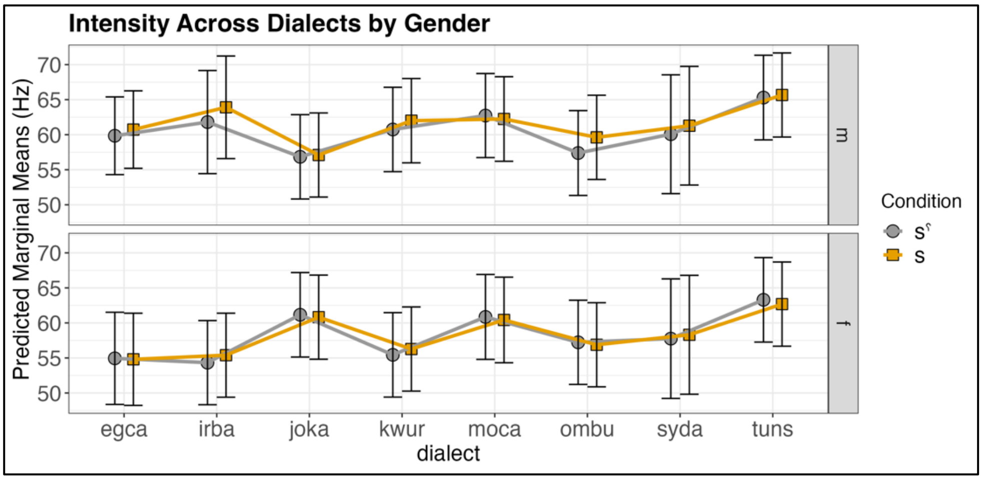


Figure 14 The 95% confidence intervals around the mean predicted values for intensity, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender

3.4.1.4 Peak Location (Hz)

This measure reveals a main effect of gender, with male speakers exhibiting lower energy compared to female speakers (β= -746.44, SE= 107.24, p > 0.0001). Additionally, there is a main effect observed in the joka and moca dialects, both of which demonstrate higher peak energy than other dialects (β= 576.50, SE= 266.31, p > 0.033104 and β= 1074.62, SE= 271.27, p > 0.01320, respectively). kwur dialect, however, show the lowest peak energy across the board [β= -711.67, SE= 266.22, p > 0.008955]. As for the effect of vowel type, /a/ vowel shows higher peak energy values than vowels /i/ and /u/ [β= 531.15, SE= 107.32, p > 0.0001]. However, there is no effect of condition which means this measure is not effective in distinguishing between the emphatic contrast [β= -0.26293, SE= 0.15464, p > 0.10654]. See Figure 15 below, which illustrates the variation of this acoustic correlate across all dialects.

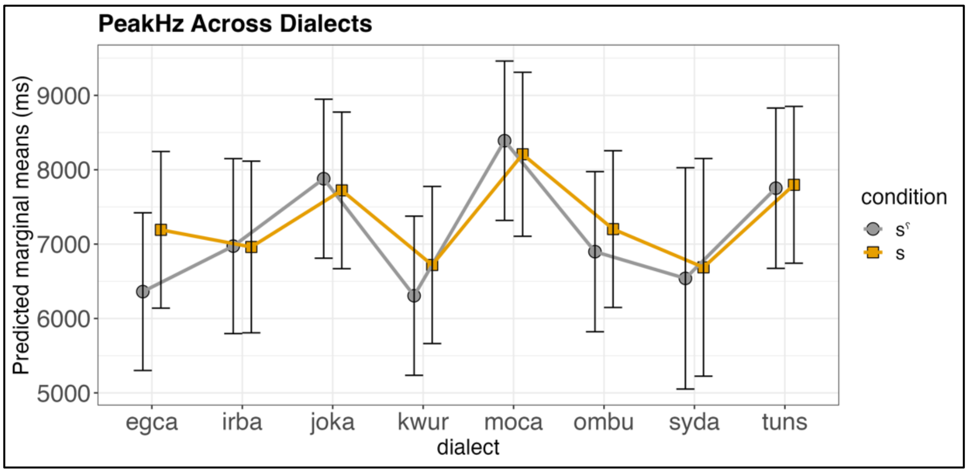


Figure 15 The 95% confidence intervals around the mean predicted values for PeakHz, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

As Figure 16 shows, kwur male versus female speakers use different peak location values of the emphatic contrast across the board, with male having lower peak energy [β= -1119.98, SE= 266.08, $p > 0.0001$]. Concerning the three-way interaction, irba male speakers are seen significantly higher in peak energy than female ones

[β= 284.35, SE= 137.19, $p > 0.04$].

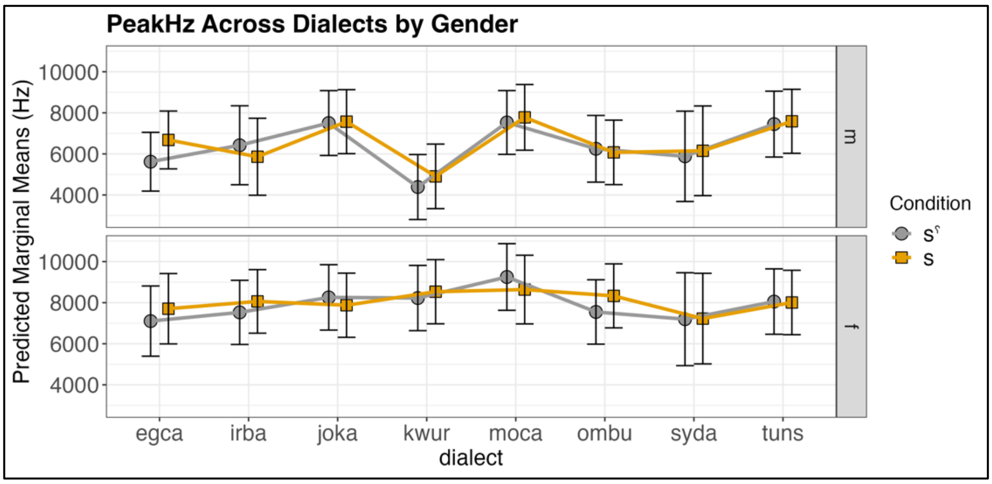


Figure 16 The 95% confidence intervals around the mean predicted values for PeakHz, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

3.4.1.5 Vowel quality effect on the fricative measures

Another point to highlight is that the vowel qualities /a/ and /i/, but not /u/, showed main effects in most measurements. Table 10 below presents an LMER output illustrating the vowel quality's primary effects for /a/ and /i/, with the former, demonstrating greater effects than the latter. Detailed information on the predicted marginal mean values for each measurement by vowel type is available in Appendix 9.12.

Measurements	Vowel Quality	Estimate	Std. Error	Pr(>|t|)
Fricative duration	a	-1.158	4.724	0.024 *
	i	6.050	5.052	0.25
COG	a	425.416	74.795	75e-05 ***
	i	272.611	79.300	0.0033 **
Intensity	a	0.73556	0.20729	0.0024 **
	i	0.52176	0.21848	0.03*
Peak location	a	531.15	107.32	8.74e-05 ***
	i	317.09	112.95	0.011422 *

Table 10 LMER output for the vowel quality main effects for /a/ and /i/

The overall findings indicate significant main effects of gender and vowel type across all measures, while the influence of dialect is observed primarily in specific measures such as fricative duration, COG, and peak location. However, the significance of these effects varies; gender exerts a considerable impact on measures like fricative duration, COG, and peak location, whereas intensity shows a lesser effect related to gender. Furthermore, the results reveal the presence of two-way and three-way interactions among the variables. In terms of two-way interactions, variations in the realisation of the fricatives ($s \sim s^{ʕ}$) are driven by gender in certain dialects, such as moca, irba, ombu, kwur, and joka. Additionally, there are dialects that exhibit distinct realisations of the emphatic contrast $s \sim s^{ʕ}$ irrespective of gender, including egca, irba, and moca. As for the three-way interaction, few dialects show interactions with gender; Syrian male and female speakers show different fricative durations in $s \sim s^{ʕ}$ relations. Egyptian male and female participants show different COG in $s \sim s^{ʕ}$. Iraqi male and female speakers show different COG in $s \sim s^{ʕ}$.

3.4.2 Short Vowel information Results

3.4.2.1 Vowel Duration in Short Vowels

Figure 17 below illustrates the vowel duration's average values following the emphatic contrast $s \sim s^{ʕ}$ across all dialects. The only effect that can be observed from the prediction model results is dialect effect on the vowel duration for Iraqi and Omani speakers; their duration for both fricatives is significantly longer than all other dialects, [β= 5.818, SE= 2.784, $p > 0.039$], and , [β= 8.190, SE= 2.621, $p > 0.002$], respectively. Another dialect effect is also seen among Moroccan speakers, which their duration for both fricatives is significantly shorter than all other dialect, [β= -1.101, SE= 2.707, $p > 0.0000$]. A vowel type effect is also observed for the vowel /i/, which has a significantly shorter duration compared to the vowels /a/ and /u/ [β = -1.626, SE = 4.575, $p < 0.006$]. However, the main effect of fricative type was observed not to contribute significantly to the emphatic contrast distinction although emphatic fricative is seen slightly higher than non-emphatics [β= 2.036, SE= 2.818, $p > 0.48741$].

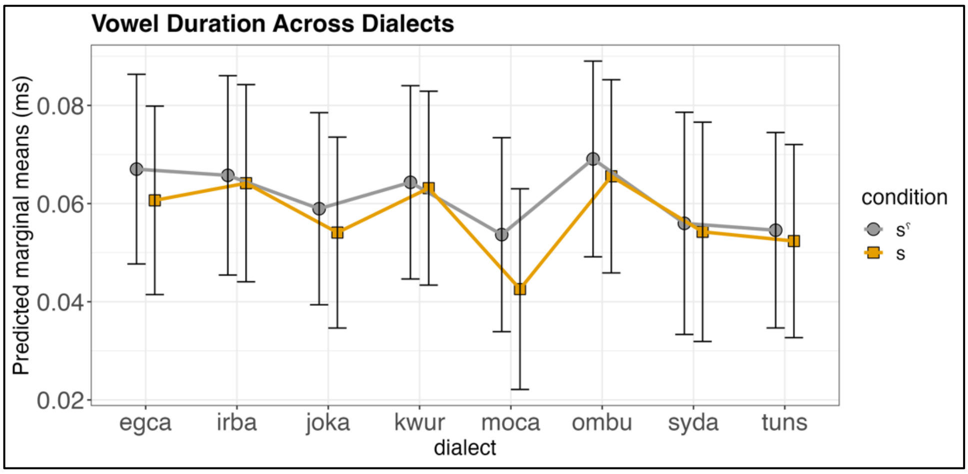


Figure 17 The 95% confidence intervals around the mean predicted values for short vowel duration, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

Regarding variable interaction, Iraqi and Syrian male and female speakers exhibit different vowel durations across all dialects. For the former, male speakers have significantly shorter durations than females [β= -6.287, SE= 2.783, $p > 0.0262$], while for the latter, male speakers have significantly longer durations than female [β= 1.165, SE= 3.628, $p > 0.001$. See Figure 18 below.

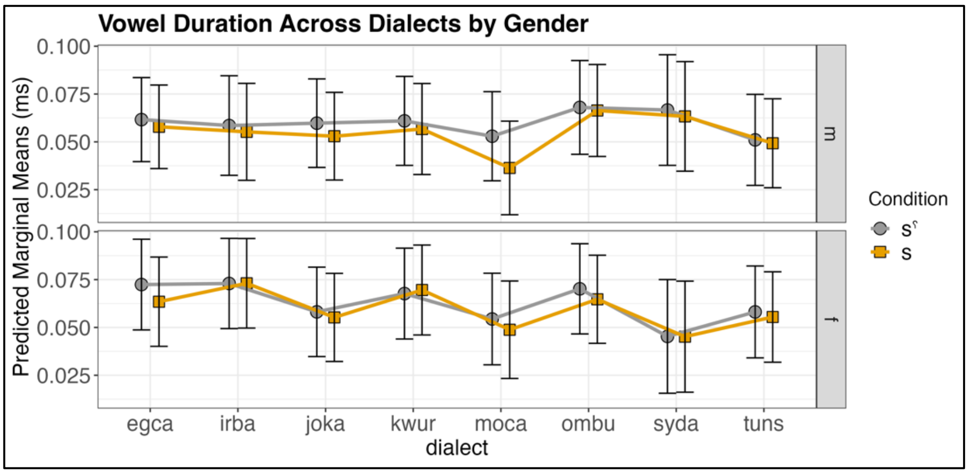


Figure 18 The 95% confidence intervals around the mean predicted values for short vowel duration, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

3.4.2.2 F1 Information in Short Vowels

Figure 19 displays the results of F1 measurements at the mid-point across all dialects. LMER results suggest that this measure is not effective in distinguishing between the emphatic contrast; thus, there is no effect of condition [β = 21.5110, SE = 10.5077, p = 0.81445]. The main effects are related to gender, where male speakers, across the board, have significantly lower F1 values than female speakers [β = -48.3015, SE = 4.6527, p < 0.000]. Additionally, the effect of dialect is observed in Moroccan and Syrian dialects, where all speakers have significantly lower F1 values compared to other dialects [β = -27.1513, SE = 12.4162, p = 0.03] and [β = -40.7449, SE = 16.6177, p = 0.01], respectively. The effect of dialect is also seen among Omani speakers, both male and female, who show significantly higher F1 values compared to all other dialects [β = 26.9421, SE = 11.8739, p = 0.02]. There is also an effect of vowel type on F1 values. The vowel /a/ shows a significantly higher F1 value compared to other vowels [β = 88.2356, SE = 14.1654, t = 6.1629, p < 0.000714], whereas the vowel /i/ exhibits a significantly lower F1 value [β = -44.4727, SE = 16.6210, t = -2.676, p = 0.038126].

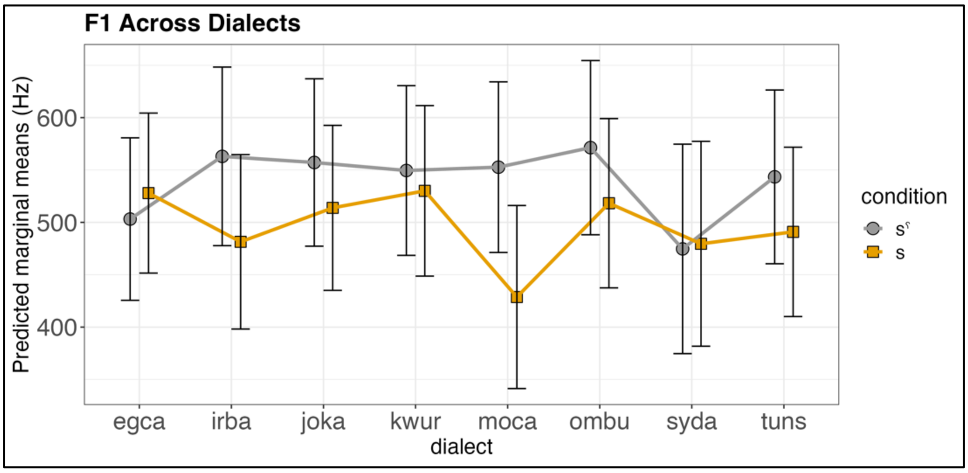


Figure 19 The 95% confidence intervals around the mean predicted values for F1in short vowels, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

As for the variable interaction, it is observed that Egyptian and Moroccan speakers use this measure to significantly differentiate between the emphatic contrast more than other dialects. For the former, the F1 value of the emphatic fricative is significantly lower than in other dialects [β = -33.9182, SE = 8.8731, p = 0.0001], while for the latter, it is significantly higher [β = 40.4798, SE = 10.5806, p = 0.0001]. Another interaction is also observed among male speakers of the Kuwaiti dialect, where the F1 value is significantly higher in males than in females [β = 31.3307, SE = 11.4561, p = 0.0071]. Refer to Figure 20 below.

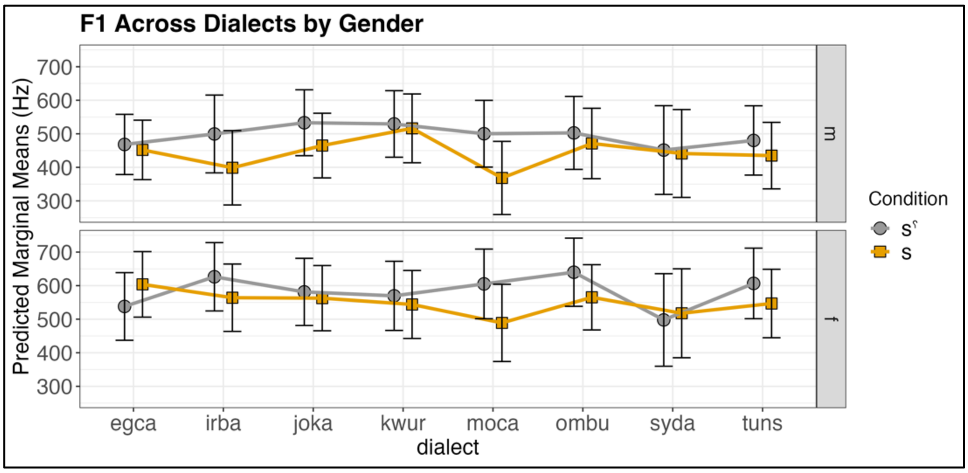


Figure 20 The 95% confidence intervals around the mean predicted values for F1 in short vowels, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

3.4.2.3 F2 Information in Short Vowels

Based on the predictive model outcomes, F2 value is influenced by fricative type, dialect, gender, and vowel type as indicated in Table 11 below. Examination of Figure 21 reveals that across all dialects, this measure is employed to distinguish the emphatic contrast, with the emphatic /sˤ/ consistently exhibiting a significantly lower F2 value compared to the non-emphatic /s/. Furthermore, a noticeable interaction is observed between dialect and gender, where male versus female speakers of the dialects egca, moca, and irba demonstrate a more pronounced utilization of F2 values compared to other dialects, see Figure 22. Similarly, another interaction between condition and dialect is observed, where kwur speakers have their F2 value for the emphatic fricative significantly higher than for the plain fricative across the board, while syda speakers have the lowest F2 value for the emphatic fricative compared to the plain counterpart among other dialects (refer to Table 11 below). Also, three-way interaction is observed between syda male speakers and fricative type, where their F2 value for the emphatic fricative is significantly different from that of female speakers, and this is only seen in this dialect (see Table 12 below).

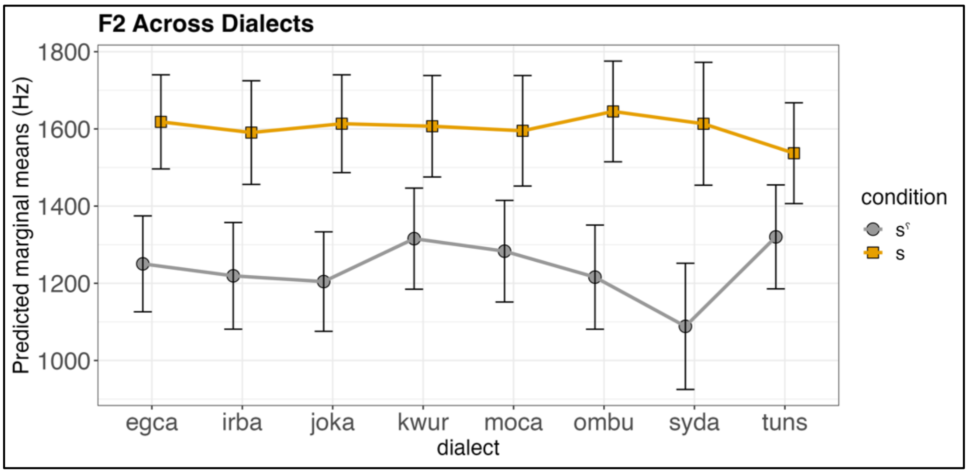


Figure 21 The 95% confidence intervals around the mean predicted values for F2 in short vowels, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

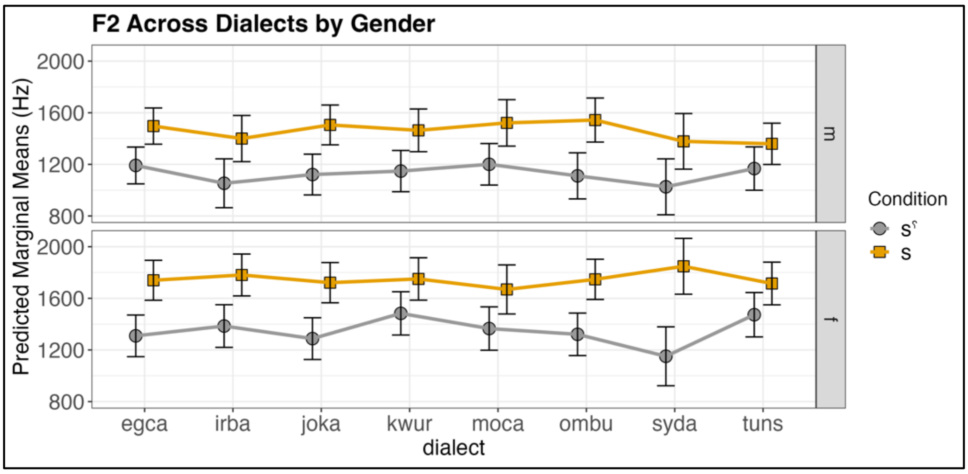


Figure 22 The 95% confidence intervals around the mean predicted values for F2 in short vowels, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

Main Effects of Measurement	**Estimate**	**Std. Error**	**Pr(>|t|)**
Fricative type	-182.645	18.164	0.0000 *
Dialect (kwur)	41.436	17.735	0.02 *
Dialect (syda)	-68.947	26.116	0.02 *
Gender (male)	-126.926	7.214	2e-16 ***
Vowel type /a/	78.720	24.425	0.008 **
Vowel type /i/	141.150	28.616	0.0006 ***

Table 11 LMER output for measurements main effects in F2 value

Interaction	Estimate	Std. Error	Pr(>|t|)
egca* gender	36.692	15.516	0.002 *
moca* gender	48.705	19.332	0.012 **
irba*gender	-51.302	19.380	0.009**
syda*fricative*gender	69.198	25.559	0.007**

Table 12 LMER output for the dialect*gender interaction

3.4.2.4 F3 Information in Short Vowels

The main effects of fricative type, dialect, and gender on F3 values are evident from the results of the predictive model, as shown in Table 13 and visualised in Figure 23 below. The F3 value for the emphatic /sˤ/ is marginally higher compared to the non-emphatic /s/, with its frequency found to be 55.970Hz higher than that of its plain counterpart. However, joka, kwur, moca speakers reveal a little to close disparity between /sˤ/ and /s/. With regard to irba speakers, the F3 frequency of both fricatives is observed to be lower across all dialects. Furthermore, male speakers differ than female speakers in which their F3 value for both fricatives is significantly lower than females'.

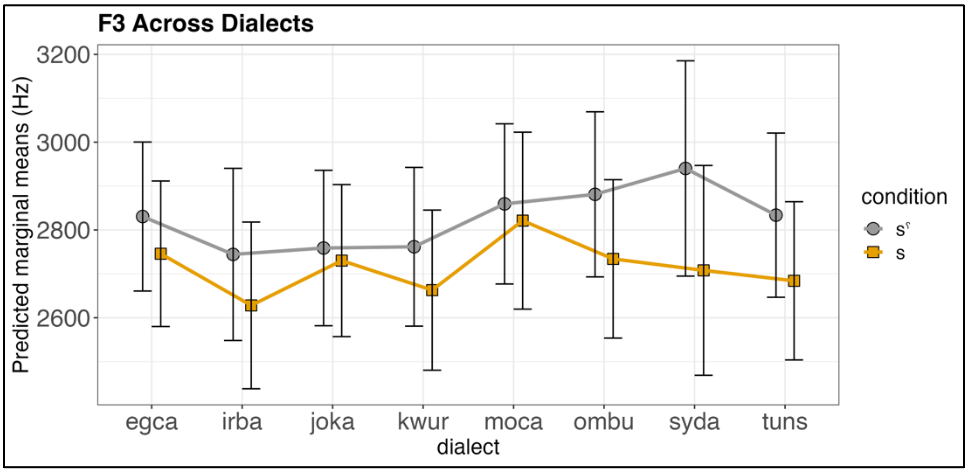


Figure 23 The 95% confidence intervals around the mean predicted values for F3 in short vowels, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

The interaction between fricative type and gender is evident in Figure 24 where male and female speakers across all dialects utilise different F3 values to distinguish the emphatic contrast [β= -27.009, SE= 10.236, p > 0.008]. Consequently, there is no dialect-condition interaction, indicating that all dialects employ F3 values to a similar extent after s~sˤ. However, there is a two-way interaction between dialect and gender among ombu and syda speakers, where male speakers have a significantly longer F3 value than females for the former, while the opposite is true for the latter. Additionally, a three-way interaction is observed between fricative type, dialect, and gender, where egca male speakers have a significantly longer F3 value for the emphatic fricative compared to females (see Table 13).

Main Effects of Measurement	Estimate	Std. Error	Pr(>|t|)
Fricative type /sˤ/	55.970	20.612	0.023
Dialect (irba)	-84.054	36.692	0.024
Gender (male)	-158.835	13.655	2e-16
Fricative type*gender	-27.009	10.236	0.008
Dialect*gender (ombu)	76.800	34.519	0.02
Dialect*gender (syda)	-101.791	48.302	0.03
Fricative*dialect*gender	51.865	21.826	0.01

Table 13 LMER output for measurements main effects in F3 value

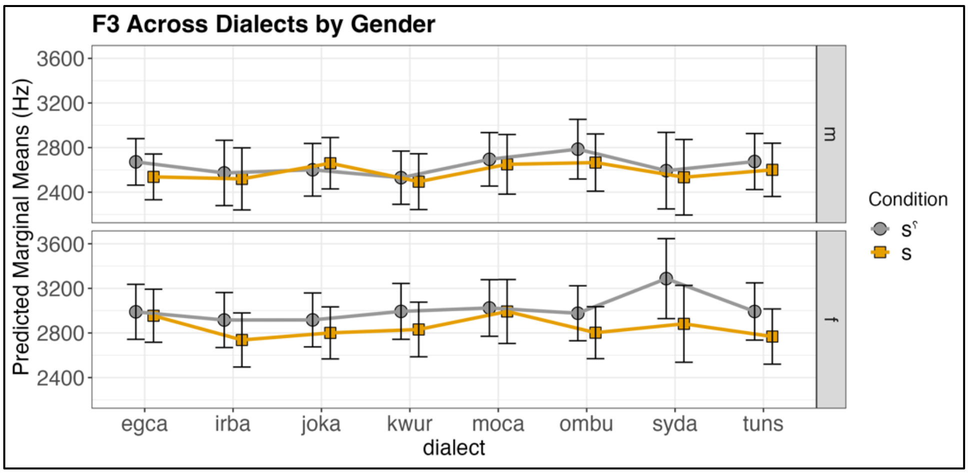


Figure 24 The 95% confidence intervals around the mean predicted values for F3 in short vowels, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

The comprehensive analysis of the short vowels following the fricatives /s/ and /sˤ/ revealed that the main effects of fricative type, dialect, gender, and vowel type were primarily prominent in F2, with the exception of vowel type affecting the F3 measure. In contrast, F1 was influenced by dialect, gender, and vowel type. The impact of dialect was evident only on vowel duration. However, the magnitude of these effects varied across different measures. For instance, gender displayed significant effects on F1, F2, and F3 values, whereas the effects of fricative type and dialect were more pronounced on F2 and F3 than on F1. Furthermore, the prediction model outcomes indicated the presence of two-way and three-way interactions among the variables under consideration. Regarding the two-way interaction, there is a gender-driven realization of the emphatic contrast /sˤ/ in some dialects, such as egca, kwur, syda, and irba. Additionally, some dialects differ from others in their realization of emphatic contrasts, regardless of gender (e.g., egca, kwur, moca, and syda). For the three-way interaction, two dialects showed interactions with fricative type and gender. Syrian male and female speakers demonstrated different emphatic contrast realizations regarding F2 values, while Egyptian male and female speakers exhibited different /sˤ/ realizations for F3 values, as observed in McCarthy (1994) and Watson (2007).

In terms of the effect of short vowel quality, the LMER model indicated that the vowels /a/ and /i/ exhibited more pronounced main effects across all formant values compared to the vowel /u/. However, regarding the interaction between fricative type and vowel type, the LMER outcomes suggested that the F2 value was the only formant affected by this interaction, specifically for the vowels /a/ and /i/, but not for /u/ (see Table 14 below). This finding aligns with previous research, which has noted that the F2 value is more influenced by /a/ and /i/ than by /u/, leading many studies to focus primarily on these two vowels (Jongman et al., 2007; Yeou, 1997; Card, 1983; Alhammad, 2014).

Measurements	**Vowel Quality**	**Estimate**	**Std. Error**	**Pr(>|t|)**
F1	a	0.6272	9.5239	0.94752
	i	-18.9711	38.7233	0.62442
F2	a	-38.248	16.3981	0.020112 *
	i	-174.400	66.586	0.009099 **
F3	a	-26.704	44.657	0.5640
	i	-148.877	113.857	0.1930

Table 14 LMER output for the interaction of vowel quality and fricative type effects for short /a/ and /i/.

3.4.3 **Long Vowel information Results**

3.4.3.1 Vowel Duration in Long vowels

Figure 25 below illustrates average vowel duration following the emphatic contrast s~sˤ across all dialects. The only effect that can be observed from the prediction model results is dialect effect on the vowel duration for Omani speakers, [β= 2.660, SE= 6.332, p > 0.000]. Their duration for both fricatives is significantly longer than all other dialects. However, the main effect of condition was observed not to contribute significantly to the emphatic contrast distinction although emphatic fricative is seen slightly higher than non-emphatics [β= 4.739, SE= 7.100, p > 0.4722].

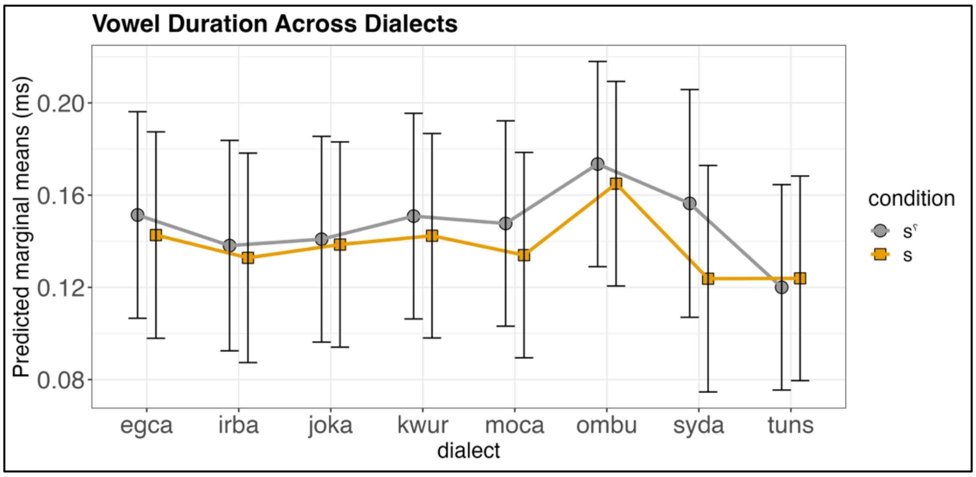


Figure 25 The 95% confidence intervals around the mean predicted values for long vowel duration, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

In terms of the two-way interaction, Irba and Syda male and female speakers exhibit different vowel durations across all dialects [β= -1.383, SE= 6.923, p > 0.04] and [β= 1.753, SE= 8.551, p > 0.04] respectively. Additionally, Syrian speakers utilise vowel duration to distinguish the s~sˤ contrast; specifically, the duration of the emphatic /sˤ/ is significantly longer than that of the plain /s/ [β= 1.159, SE= 3.221, p > 0.0003]. In terms of the three-way interaction, Egyptian male speakers demonstrate shorter vowel duration for the fricative /sˤ/ compared to female speakers [β= -7.650, SE= 2.616, p > 0.0003]. See Figure 26 below.

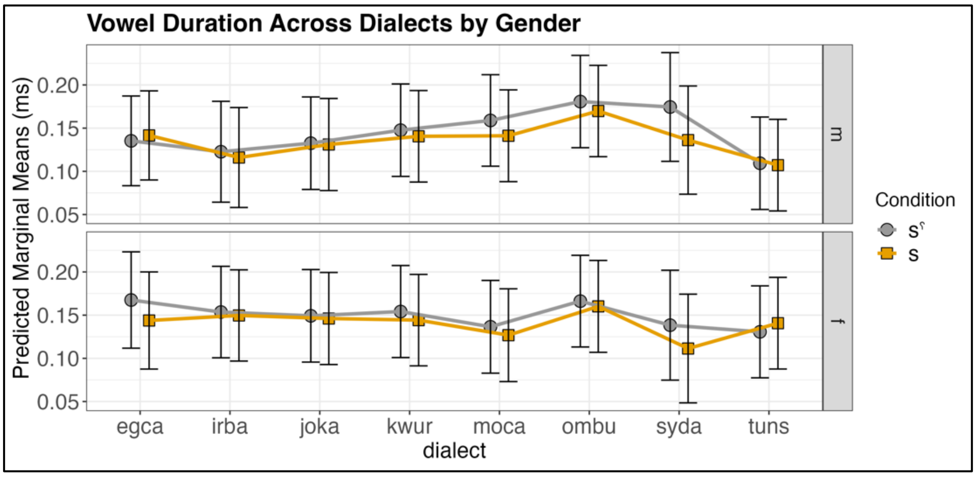


Figure 26 The 95% confidence intervals around the mean predicted values for long vowel duration, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

3.4.3.2 F1 Information in Long vowels

Figure 27 displays the results of F1 measurements at the mid-point across all dialects. The LMER results suggest that this measure is effective in distinguishing between the emphatic contrasts, showing a condition effect [$\beta = 15.1198$, SE = 6.0830, $p = 0.321$]. Additionally, significant main effects are observed for gender [$\beta = -55.2984$, SE = 3.4526, $p < 0.001$], and vowel type, with F1 values being significantly higher for /a/ [$\beta = 190.9272$, SE = 8.4663, $p < 0.001$] and significantly lower for /i/ [$\beta = 113.7000$, SE = 8.5011, $p < 0.001$].

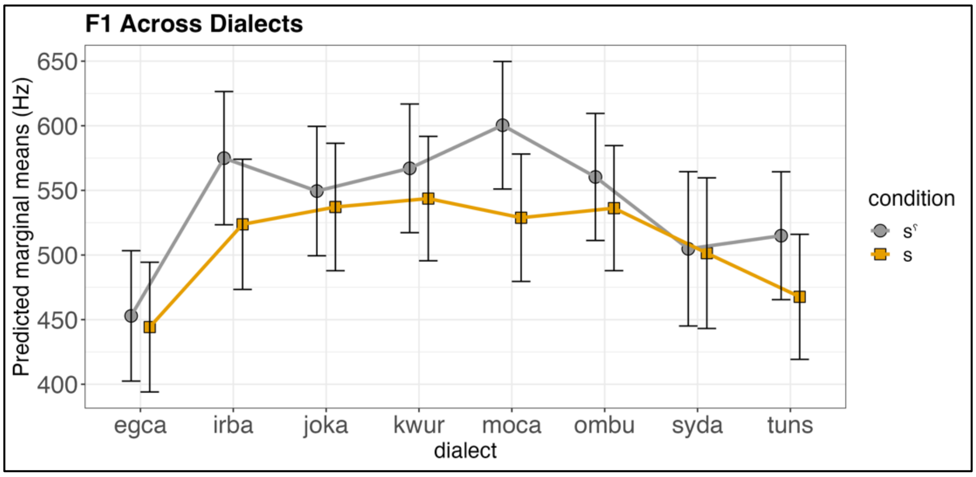


Figure 27 The 95% confidence intervals around the mean predicted values for F1 in long vowels, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

All participants, males and females in each dialect use the F1 values differently and they vary with regard to the degree of effects. So, egca and syada have F1 value significantly lower than other dialects [β= -76.9249 , SE= 8.9526 , p > 0.000, and β= -22.3801, SE= 11.4351, p > 0.053] respectively, while the speakers of the remaining dialects have their F1 value significantly higher see Table 15 below.

Main Effects of Diaelct	Estimate	Std. Error	Pr(>|t|)
irba	23.8497	9.2189	0.0114 *
joka	17.8131	8.8300	0.046*
kwur	29.8786	8.5545	0.0007 ***
moca	39.1310	8.6991	0.0000 ***
ombu	22.8342	8.5416	0.05

Table 15 LMER output for Dialect as a main effect in F1 value

Egyptian is the only dialect that has an interaction with fricative types where its speakers show different F1 values between /sˤ/ and /s/ [β= 20.6653, SE= 6.1107, p = 0.0007]. Refer to Figure 28 below.

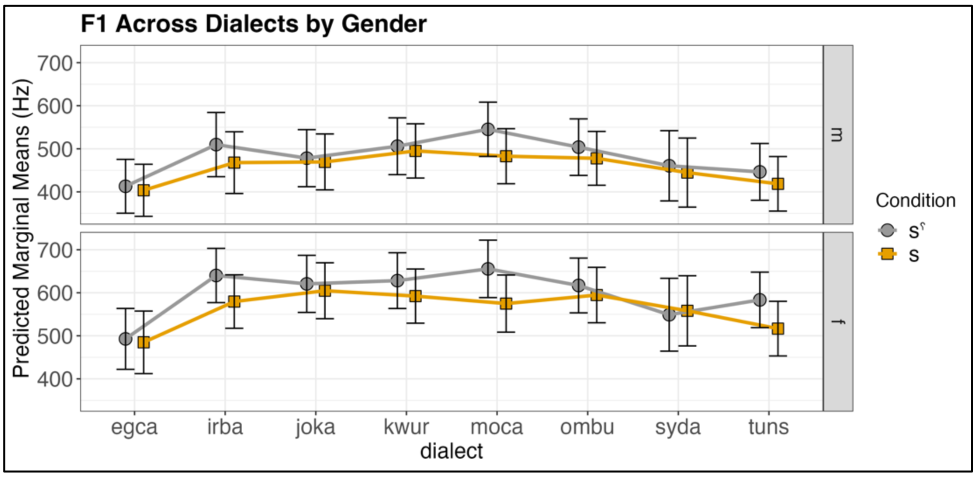


Figure 28 The 95% confidence intervals around the mean predicted values for F1 in long vowels, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender

3.4.3.3 F2 Information in Long vowels

Based on the predictive model outcomes, F2 value is influenced by fricative type, dialect, gender, and vowel type as indicated in Table 16 below. Examination of Figure 29 reveals that across all dialects, this measure is employed to distinguish the emphatic contrast, with the emphatic /sˤ/ consistently exhibiting a significantly lower F2 value compared to the non-emphatic /s/. Furthermore, the only noticeable interaction is observed between fricative type and gender, where male speakers, across the board, have significantly higher F1 values for the emphatic fricative /sˤ/ compared to female speakers. Additionally, there is an interaction between dialect and gender, where male speakers from the dialects egca, kwur, and syda show more noticeable utilization of F2 values compared to other dialects (refer to Table 17 and Figure 30).

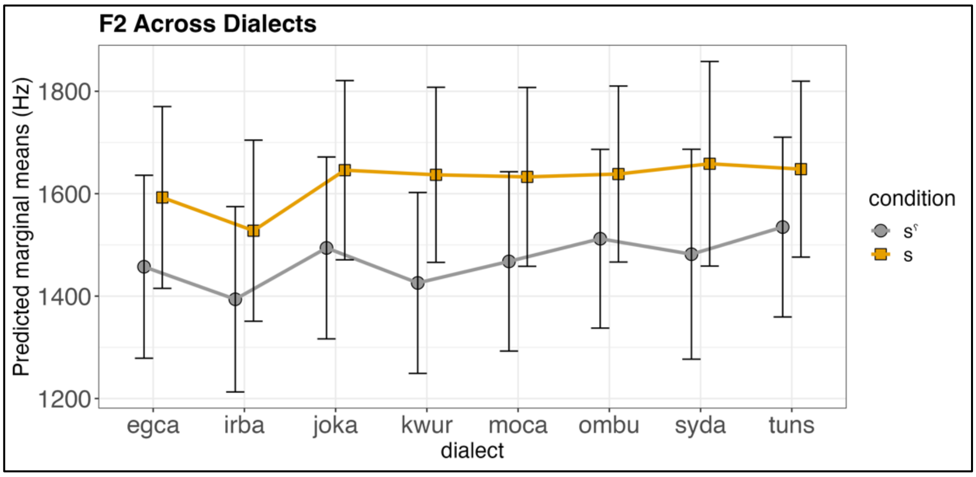


Figure 29 The 95% confidence intervals around the mean predicted values for F2 in long vowels, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

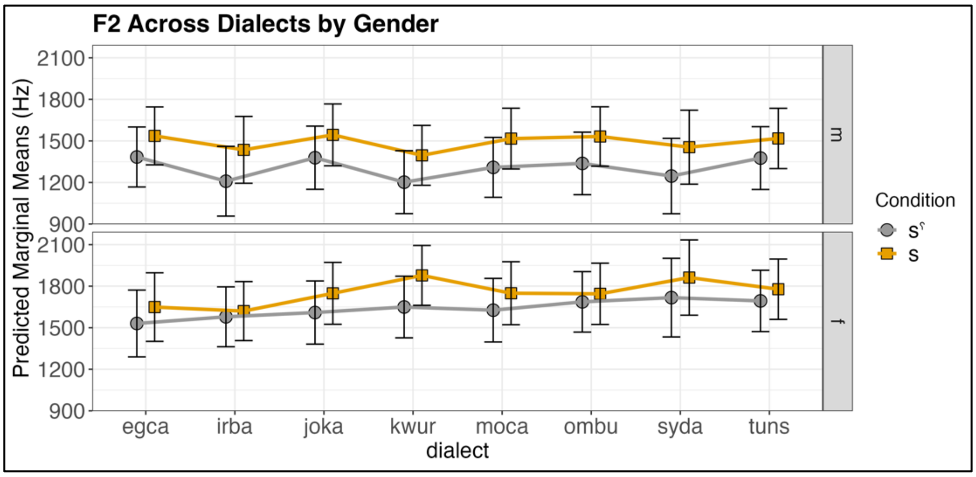


Figure 30 The 95% confidence intervals around the mean predicted values for F2 in long vowels, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

Main Effects of Measurement	Estimate	Std. Error	Pr(>|t|)
fricative type	-75.83844	23.27106	0.008 **
Dialect (irba)	-85.98482	28.50050	0.0033 **
gender	-148.39341	10.73930	2e-16 ***
vowel type /a/	-126.90199	32.39563	0.002 ***
vowel type /i/	723.81066	32.51077	3.44e-10 ***

Table 16 LMER output for measurements main effects in F2 value

Interaction	Estimate	Std. Error	Pr(>|t|)
egca* gender	83.47085	28.00322	0.003**
kwur* gender	-84.01470	26.53984	0.0029 **
syda*gender	-71.26983	35.39833	0.047 *
Fricative* gender	-17.29330	8.61840	0.004*

Table 17 LMER output for the dialect*gender interaction

3.4.3.4 F3 Information in Long vowels

The main effects of dialect, gender, and vowel type on F3 values are evident from the results of the predictive model, as shown in Table 18 and visualised in Figure 31 below. There is no main effect of fricative type in which this measure is not the one can differentiate between the emphatic contrast s~ $s^ʕ$/ as observed in the LMER model results [β= 16.946, SE= 12.568, p = 0.206118]. However, egca, irba, and syda speakers reveal only a slight, non-significant disparity between /$s^ʕ$/ and /s/. For Kuwaiti speakers, the F3 frequency of both fricatives is lower, and it is even lower across all dialects for Iraqi speakers. Furthermore, male speakers differ from female speakers, with their F3 values for both fricatives being significantly lower than those of females (see Table 18 and Figure 32 below).

Main Effects of Measurement	Estimate	Std. Error	Pr(>|t|)
Dialect (kwur)	-73.770	32.564	0.0026*
Dialect (irba)	-121.751	35.157	0.0008***
Gender (male)	-155.009	13.149	2e-16
vowel type /a/	-99.404	17.387	0.00019***
vowel type /i/	148.349	17.610	0.0000***

Table 18 LMER output for measurements main effects in F3 value

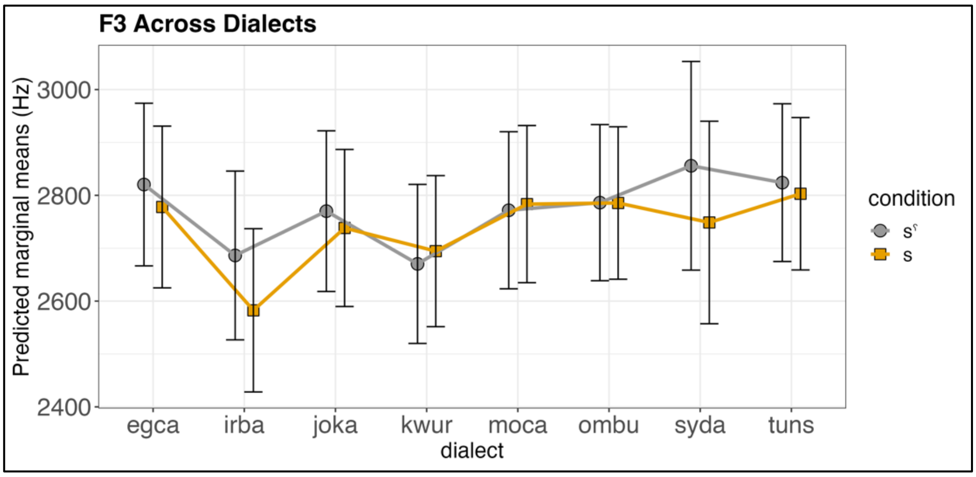


Figure 31 The 95% confidence intervals around the mean predicted values for F3 in long vowels, as estimated by the LMER model in /s/ and /sˤ/ across dialects.

The model results indicate that there is no interaction between variables, meaning there is no dialect-condition interaction. This suggests that all dialects employ F3 values to a similar extent after the contrast between /s/ and /sˤ/. Regarding vowel quality, the LMER analysis shows that the vowels /a/ and /i/ make more effective use of the F3 measure compared to the vowel /u/, with both fricative types showing significantly higher values for /i/ and lower values for /a/.

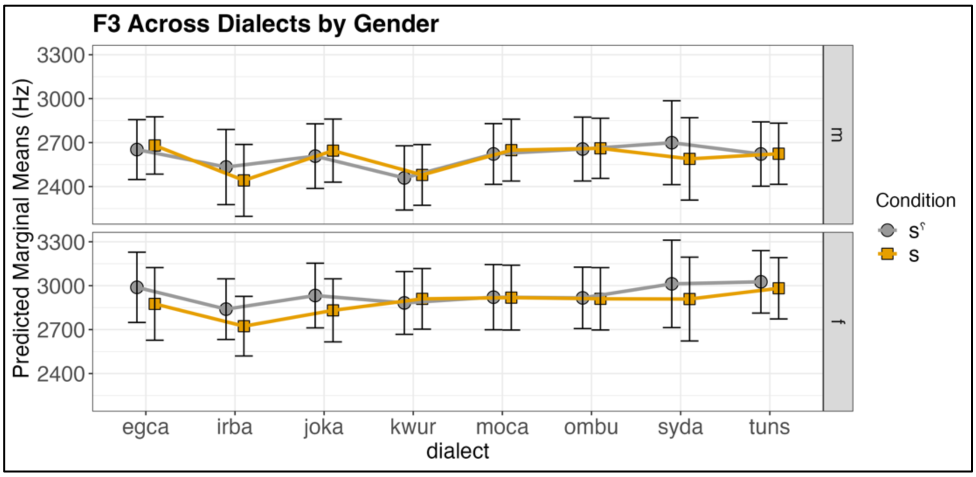


Figure 32 The 95% confidence intervals around the mean predicted values for F3 in long vowels, as estimated by the LMER model in /s/ and /sˤ/ by dialect and gender.

The comprehensive analysis of the long vowels following the fricatives /s/ and /sˤ/ revealed that the main effects of fricative type, dialect, gender, and vowel type were primarily prominent only in F2, with the exception of fricative type affecting the F3 measure. In contrast, F1 was influenced by fricative type, dialect, and vowel type. The impact of dialect was evident only on vowel duration. However, the magnitude of these effects varied across different measures. For instance, gender displayed significant effects on F2, and F3 values, whereas the effects of fricative type and dialect were more pronounced on F1 and F2 than on F3. Furthermore, the prediction model outcomes indicated the presence of two-way and three-way interactions among the variables under consideration. Regarding the two-way interaction, there is a gender-driven realization of the emphatic contrast /sˤ/ in some dialects, such as egca, kwur, syda, and irba. Additionally, some dialects differ from others in their realization of emphatic contrasts, regardless of gender (e.g., moca, and syda). For the three-way interaction, only one dialect showed interactions with fricative type and gender. Egyptian male and female speakers demonstrated different emphatic contrast realizations regarding the duration of the vowel.

In terms of the effect of long vowel quality, the LMER model indicated that the vowels /a/ and /i/ exhibited more pronounced main effects across all formant values compared to the vowel /u/. However, regarding the interaction between fricative type and vowel type, The LMER outcomes suggested that there is no significant interaction between specific formants and vowel type; thus, the effects across all vowels are relatively similar (see Table 19 below).

Measurements	Vowel Quality	Estimate	Std. Error	Pr(>|t|)
F1	a:	6.4549	15.3027	0.680539
	i:	13.8693	15.3375	0.383307
F2	a:	-81.756	53.873	0.153422
	i:	11.212	54.005	0.838763
F3	a:	-4.661	30.213	0.87862
	i:	-51.712	30.431	0.10143

Table 19 output for the interaction of the long vowel quality and fricative type effects for /a:/ and /i:/

3.4.4 Overview of Vowel Length Effects on Formant Measures

Our findings demonstrate that vowel length—whether short or long—is reflected in formant measurements in varied ways across all parameters. Notably, the F2 formant shows a clear influence of preceding fricative type in long vowels, with an even greater effect in short vowels. Generally, the effect of fricative type is most pronounced in F2 and F3 values when followed by short vowels, whereas it is more prominent in F1 and F2 values for long vowels. This distinction suggests that, for short vowels, F2 and F3 measures are useful for identifying the emphatic contrast (s~sˤ), while for long vowels, F1 and F2 are more effective. Additionally, factors such as dialect and gender, along with variable interactions, show different patterns depending on vowel length. In particular, short vowels displayed a greater number of main effects related to factors like gender, dialect, and vowel type across all formant measures compared to long vowels.

3.4.5 F2-F1 Difference for Short and Long Vowels

F2-F1 differences for short and long vowels exhibit distinct acoustic patterns across emphatic /sˤ/ and plain /s/ fricatives, as illustrated in Figure 33 below. For short vowels, /sˤ/ shows lower F2-F1 mean values, with /a/ ranging from 534.16 Hz to 896.74 Hz (mean: 715.45 Hz), /i/ ranging from 590.53 Hz to 1009.16 Hz (mean: 799.84 Hz), and /u/ ranging from 425.03 Hz to 704.16 Hz (mean: 564.59 Hz). In contrast, the plain fricative /s/ yields higher F2-F1 means: /a/ ranges from 936.95 Hz to 1254.66 Hz (mean: 1095.80 Hz), /i/ ranges from 1150.73

Hz to 1597.71 Hz (mean: 1374.22 Hz), and /u/ ranges from 654.38 Hz to 1033.94 Hz (mean: 844.16 Hz). These differences suggest that short vowels produced after /s/ exhibit a more expanded vowel space compared to those produced after /sˤ/, indicating a clearer separation between F1 and F2 for the plain fricative. For long vowels, the same trend persists but with varying degrees. The emphatic /sˤ/ has mean F2-F1 values of 571.92 Hz for /a:/ (range: 458.94 Hz to 684.89 Hz), 1601.23 Hz for /i:/ (range: 1312.13 Hz to 1890.33 Hz), and 530.54 Hz for /u:/ (range: 434.95 Hz to 626.13 Hz). Meanwhile, /s/ shows higher F2-F1 means: 848.36 Hz for /a:/ (range: 577.49 Hz to 1119.24 Hz), 1699.44 Hz for /i:/ (range: 1397.19 Hz to 2001.69 Hz), and 600.95 Hz for /u:/ (range: 423.69 Hz to 778.22 Hz). The more pronounced F2-F1 differences in the plain /s/ condition reflect a greater articulatory-acoustic distance, particularly for front vowels like /i:/.

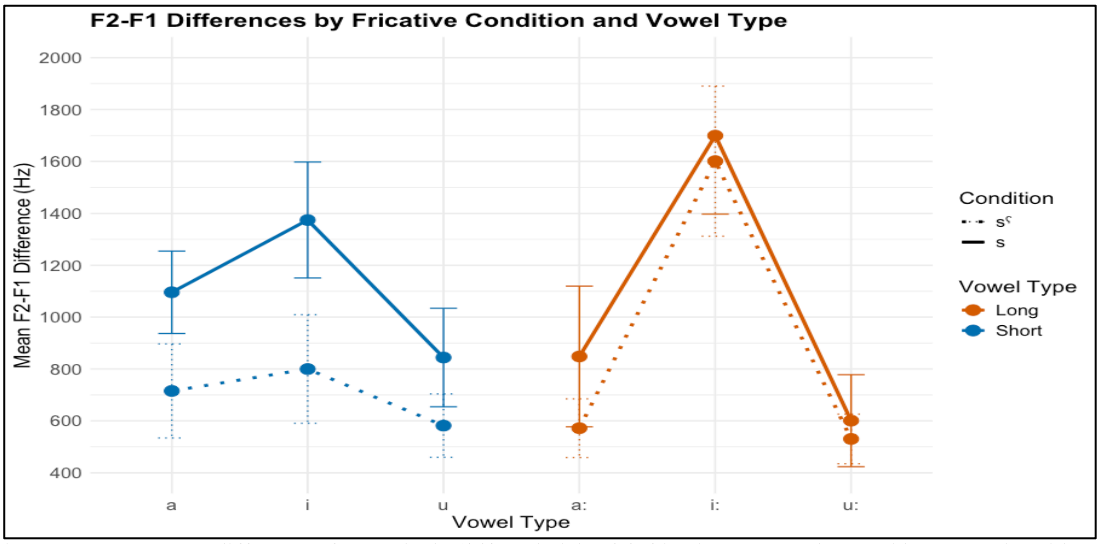


Figure 33 Mean F2-F1 differences for emphatic (/sˤ/) and plain (/s/) fricatives across short and long vowels, with raw means plotted and error bars representing ±1 standard deviation.

Building on these observations, the impact of fricative type extends beyond vowel quality distinctions in short vowels to influence long vowels even more substantially, as illustrated in Figures 34 and 35. Across both vowel lengths, the plain fricative /s/ results in a smaller F2-F1 difference than the emphatic fricative /sˤ/, indicating a more expanded vowel space under the plain condition. This effect is particularly notable in long vowels, where the difference between the plain and emphatic fricatives /s/ and sˤ/ is more marked. Thus, the F2-F1 difference analysis highlights that long vowels possess a larger, more spread vowel space compared to short vowels. Additionally, the effect of fricative type, while present in both vowel lengths, has a small impact on long vowels.

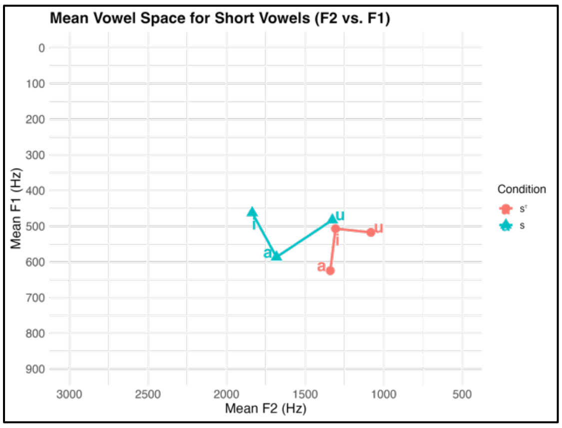


Figure 35 Mean Vowel Space for Short Vowels (F2 vs. F1)

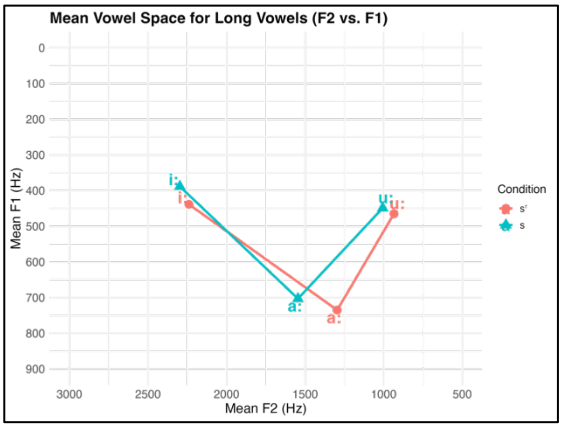


Figure 34 Mean Vowel Space for Long Vowels (F2 vs. F1)

3.4.6 **Vowel Trajectory Results**

Based on previous research, it is established that vowels occurring in the vicinity of emphatic fricatives consistently exhibit a distinct pattern characterised by high F1 and low F2 values. However, the behaviour of F3 has been found to vary across studies, with some reporting higher values and others lower (McCarthy, 1994; Watson, 2007). Therefore, the first two formant frequency values (F1 and F2) for each vowel (/a i u a: i: u:/) were measured at ten equidistant points throughout the duration of the vowel. Figure 36 illustrates the results of F1 and F2 trajectories concerning vowel quality in relation to fricative type and dialect. It is evident that the F2 value of the emphatic fricative /sˤ/ consistently registers lower across most vowel types. Additionally, long vowels exhibit more variation, with trajectories starting notably lower and gradually converging towards the end. Conversely, short vowels display greater stability and consistency throughout the contrast. Particularly for the vowel /a/, there is a discernible utilisation of the F2 measure to differentiate the emphatic contrast across all dialects. Similarly, vowel /i/ demonstrates behaviour similar to vowel /a/, while the vowel /u/ exhibits inconsistent F2 distinctions between fricative contrasts.

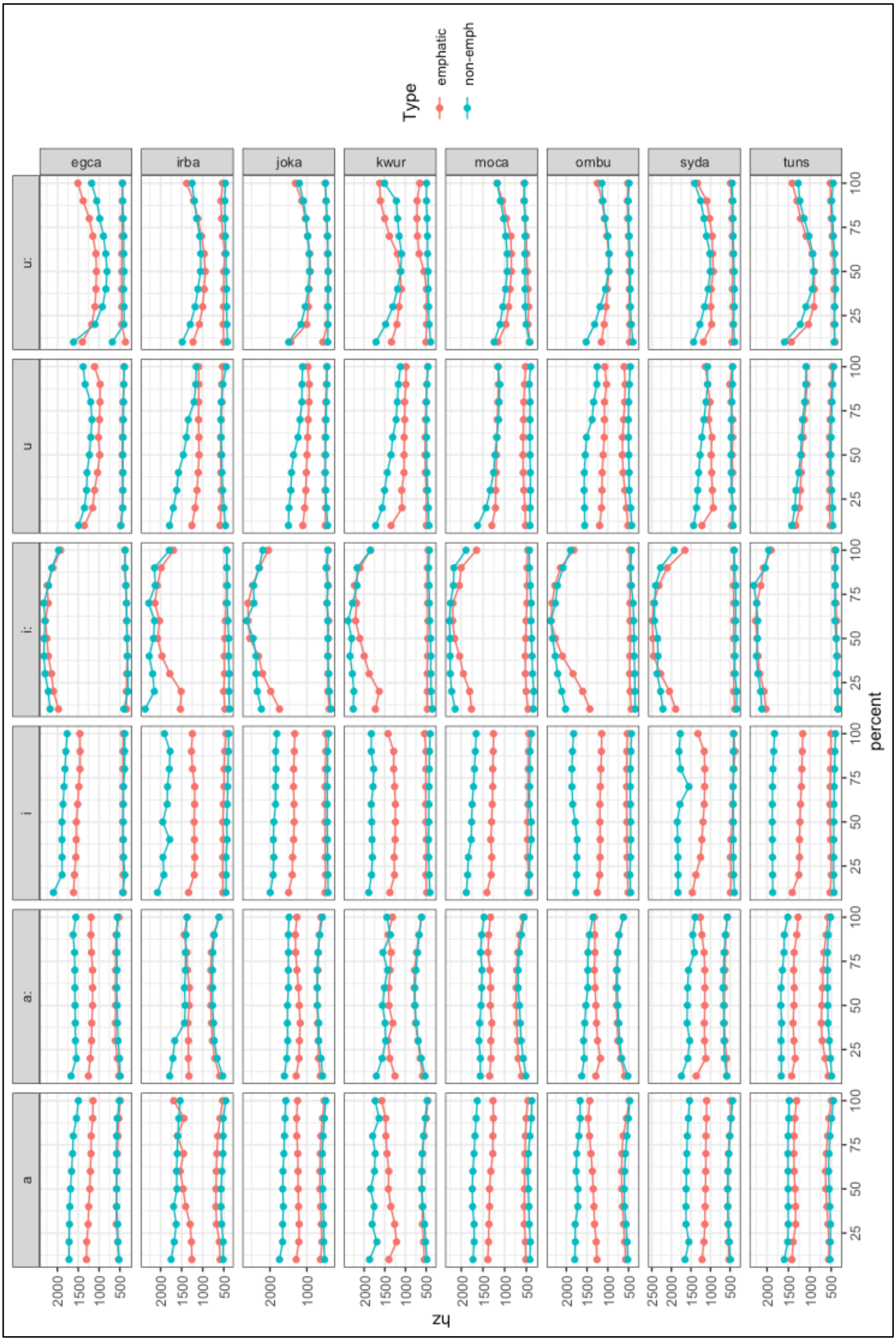


Figure 36 Vowel trajectory for F1 and F2 for (/a, i, u, a:, i:, u:/) across dialects for both conditions (emphatic and non-emphatic)

3.4.7 **GAMM Results for F2**

Based on the results obtained from the LMER model and the analysis of vowel-formant trajectories, the acoustic correlate that exhibited consistent and significant effects across all dialects was the F2 measure. Notably, F2 was consistently lowered in all dialects and effectively marked the emphatic contrast. Additionally, the long vowels /a:/ and /i:/ were found to exert greater influence on most of the acoustic correlates. Therefore, the F2 measure of the two long vowels /a:/ and /i:/ was subjected to time-varying analysis using Generalised Additive Mixed Models (GAMMs) (Wood, 2017). This analysis involved measuring ten equidistant points between the onset and offset of the vowel, and separate GAMMs were fitted to the time-varying F2 data for each dialect using the *MGCV::GAM* function in R (Wood, 2017). Predictor variables included a parametric term of fricative*vowel type and smooth terms of normalised percent and normalised percent-by-fricative*vowel type interaction. To simplify the model and enhance statistical power, gender was excluded as a predictor in the GAMMs. Consequently, the estimates of the model were obtained by collapsing over gender groups. Additionally, the GAMMs incorporated fitted random smooths of percent-by-speaker and percent-by-word to account for variability in the data. After comparing two models using the ITSADUG::COMPAREML function (Van Rij et al., 2015), it was determined that fitting random smooths of percent-by-speaker and percent-by-word significantly improved the model performance (see Table 20). Autocorrelation was addressed by implementing a first-order autoregressive (AR1) model.

Model	score	Edf	Difference	Df	P.Value
Predictor without random smooths	11507.63	36	-	-	-
Predictor with random smooths by speaker and word	11203.25	37	304.383	1.000	<2e-16***

Table 20 Comparing models using COMPAREML function.

The overall results for the F2 trajectory in /a:/ and /i:/ below show a significant difference between emphatic and non-emphatic fricatives, where all dialects use this measure to mark the emphatic contrast. However, the differences of height and shape in both vowel qualities vary across all dialects; the vowel quality /a:/ in Figure 37 below shows that the emphatic fricative /sˤ/ is lower at the beginning of the time-varying and tends to nearly converge at the end. However, the differences of the height between the emphatic contrast are observed to be significant throughout the time-intervals in dialects like egca, joka, moca, tuns, whereas some dialects exhibit significant differences in height at some particular parts of the curves like irba (0-30% interval), kwur (0-65% interval), ombu (0-90% interval), and syda (0-100% interval). This can clearly be visualised in Figure 37.

With similar manner, the emphatic and the non-emphatic fricatives /s~sˤ/ show a considerable difference in height at the beginning of the time intervals of the vowel /i:/ but they start to slightly converge at around the middle point. Figure 37 reveals that the trajectories of the emphatic contrast s~sˤ do form a little-to-overlap in confidence intervals in some dialects like irba, kwur, syda, while a complete converge can be viewed in the remainder of the dialects, egca, joka, ombu, and tuns. As for the significance of the height, Figure 38 illustrates that moca is the only dialect displaying a significant difference between the emphatic contrast throughout the trajectory. Also, the significant difference of height is evidenced at the first parts of the curves in some dialects. Specifically, egca, joka, and syda show significant difference in height between the intervals 0 – 30 %, while the significant parts of the dialects irba, kwur, and ombu extends between the intervals 0 – 50 %.

According to the overall shapes of the emphatic contrast in the vowel /a:/, the emphatic fricative /sˤ/ exhibits a more non-linear trajectory compared to the plain fricative /s/ in the dialects egca, moca, and syda, while the non-emphatic fricative /s/ shows a more non-linear trajectory in the dialects irba, kwur, ombu, tuns. However, the trajectories of the /i:/ data do reveal an absence of non-linear differences between the emphatic contrast.

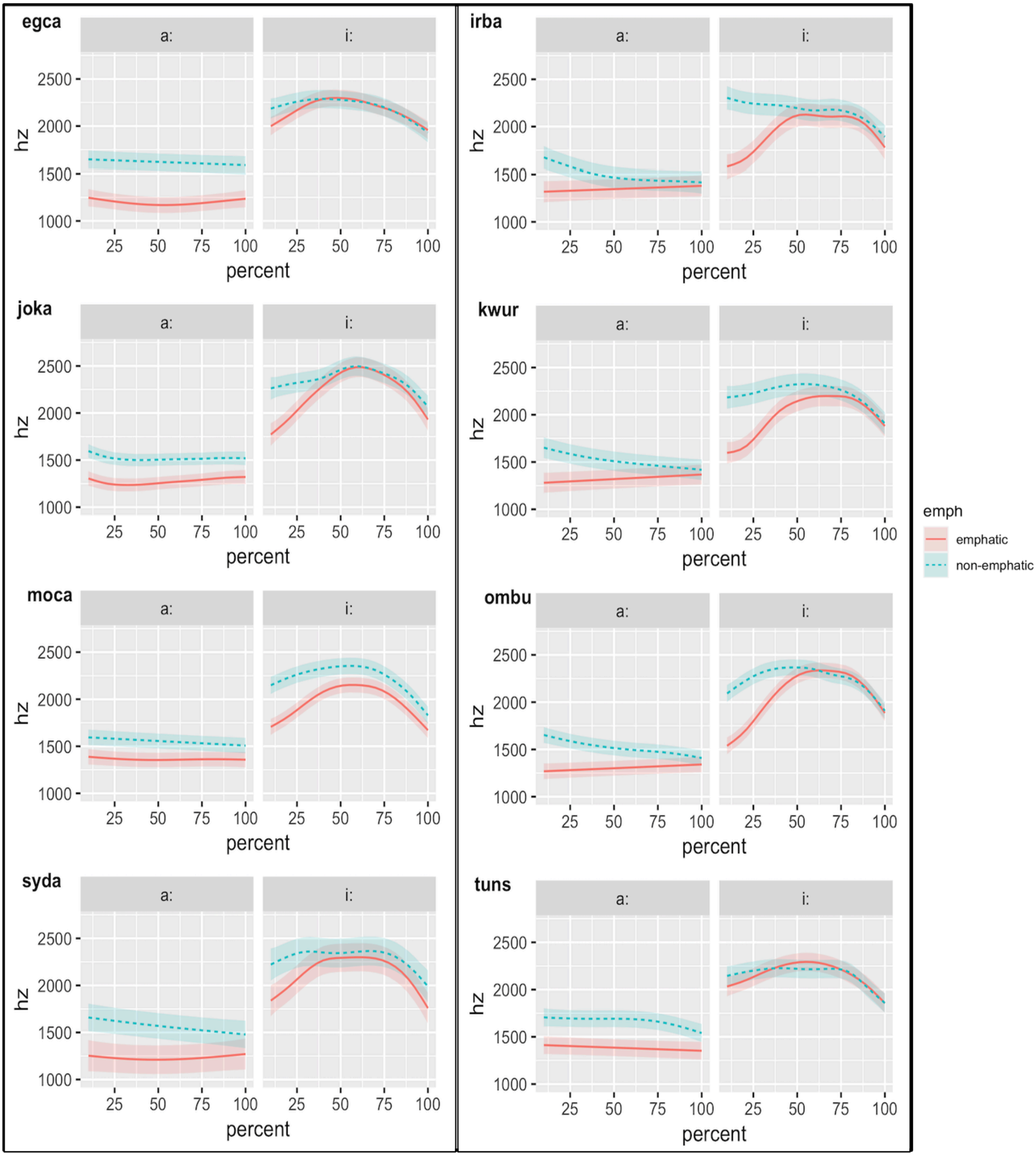


Figure 37 GAMM fits of the effects of normalised percent-by-speaker and percent-by-word

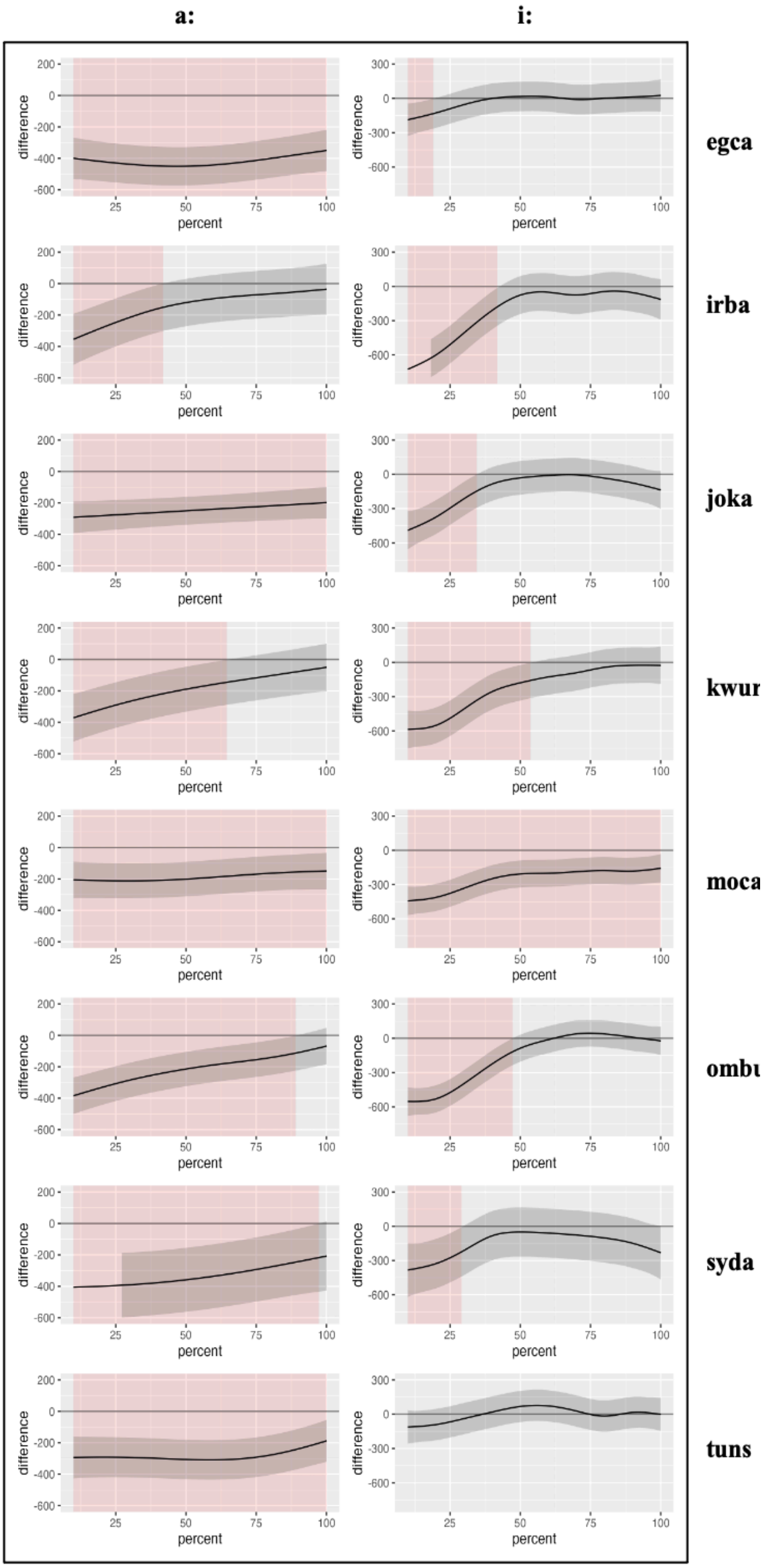


Figure 38 Plot-diff of the significant differences of F2 trajectories for /a:/ and /i:/ in all dialects

The overall findings from the F2 trajectories can be synthesised as follows: Firstly, all dialects utilised this measure to distinctly distinguish between s~sˤ; however, the extent of these differences varied across dialects. Concerning the significance of height, egca, joka, and moca dialects displayed significant variations in the time-varying emphatic contrast from the onset to the offset of the vowel /a:/, whereas significance in height was only observed in the initial parts of the trajectories for the irba, kwur, ombu, and syda dialects. For the vowel /i:/, egca, joka, and syda exhibited significant height differences between intervals 0 – 37%, while irba, kwur, and ombu displayed significance extending between intervals 0 - 52%. Secondly, there was complete overlap in the /i:/ data for the egca, joka, ombu, and tuns dialects, while partial convergence was observed between the emphatic contrast trajectories for the irba, kwur, and syda dialects. Thirdly, concerning the shape of the /a:/ trajectory, the emphatic fricative /sˤ/ demonstrated a more non-linear trajectory in the egca, moca, and syda dialects, while the plain fricative /s/ exhibited a more non-linear trajectory in the irba, kwur, ombu, and tuns dialects. However, there was an absence of non-linearity in the /i:/ data.

3.4.8 **Interim Discussion and Conclusion**

In summary, our analysis has revealed that much of the variation in this dataset stems from main effects related to dialects, genders, and vowel types. Although certain acoustic correlates intended to capture the fricative contrast between s and sˤ did not yield significant results, some variability was observed. This finding aligns with previous research, such as Sakr (2024), which similarly reported non-significant outcomes. Conversely, other studies have identified specific measures, including Peak location (Norlin, 1983), locus equations (Yeou, 1997), and intensity and duration for plosives (Kriba, 2009), as capable of distinguishing between the emphatic contrast, while measures like COG exhibited overlapping results (Norlin, 1983).

However, amidst this variability, the only consistent acoustic correlates emerging from this cross-dialectal dataset for the emphatic contrast are F2 and F3 measures in the case of short vowels and the F1 and F2 measures in the case of long vowels. It was initially anticipated that the second formant of the vowel following the emphatic /sˤ/ would be lowered due to the articulatory effect of the emphatic fricative. Instead, the vowel appears to be influenced by coarticulatory effects, as suggested by Yeou (1997). Some researchers attribute the lowering of F2 to the retraction of the tongue root [+RTR] (Davis, 1995; Watson, 1999; Al-Tamimi, 2017).

Hence, our findings indicate a distinct realisation between the emphatic and non-emphatic fricatives s and $s^ʕ$, where F2 values are lower for the former and higher for the latter. Moreover, the analysis of the first formant (F1) values in this dataset aligns with previous studies, indicating somewhat higher F1 values for the emphatic fricative /$s^ʕ$/ (Al-Masri & Jongman, 2004; Alosh, 1987; Al-Tamimi, 2017; Moisik, 2013; Stevens, 1968; Watson, 1999; Al-Khairy, 2005). Conversely, while F1 values for non-emphatics were generally lower across dialects, they did not exhibit significant variations to denote the emphatic contrast in the short vowel instances. This lack of significance was also reflected in the trajectory analysis of vowel formants, particularly with respect to F1.

In contrast, F2 values consistently showed lower results across most vowel types occurring in the emphatic context. As for F3, our findings suggest a marginal marking of the emphatic contrast, with higher frequencies observed in vowels following emphatic fricatives across all dialects except for Jordanian, Kuwaiti, and Moroccan speakers. However, this marginal effect aligns with previous research indicating inconsistent patterns for F3 measures (Al-Khairy, 2005; McCarthy, 1994).

Accounting for the different effect of short and long vowels on formants, our observations reveal that the impact of vowel length—whether short or long—varies across all formant measures. Notably, F2 measurements are affected for both short and long vowels, with a more pronounced influence observed in short vowels. This finding indicates that vowel length is reflected in formant characteristics differently. Generally, the fricative-type effect is prominent in F2 and F3 values when followed by short vowels, whereas it is more pronounced for F1 and F2 in long vowels. This suggests that, in the case of short vowels, F2 and F3 measurements effectively distinguish the emphatic contrast /s~$s^ʕ$/, while in long vowels, this contrast is more reliably distinguished by F1 and F2 values. Additionally, other factors, such as dialect and gender, and interactions among variables, vary between long and short vowels. Specifically, short vowels displayed more main effects of factors like gender, dialect, and vowel type across all formant measurements compared to long vowels. This vowel length effect aligns with predictions from previous studies, such as Al-Tamimi and Ferragne (2005), which suggest that vowel length influences both duration and mean values in Arabic.

The analysis of F2-F1 differences for short and long vowels provides significant insights into the impact of vowel length and fricative type on vowel space. The results indicate that long vowels exhibit higher mean F2-F1 differences than short vowels, suggesting a more distinct separation between F2 and F1 formants and, consequently, a more expanded vowel space. In contrast, short vowels have lower F2-F1 differences, leading to a more compact vowel space that may limit vowel quality distinction. This observation is consistent with the ranges reported, where long vowels display a broader range of F2-F1 differences compared to short vowels. The influence of fricative type (/s/ vs. /sˤ/) on F2-F1 differences is evident across both vowel lengths but is more observed in long vowels. Specifically, the plain fricative **/s/** consistently yields higher F2-F1 mean values compared to the emphatic /sˤ/, indicating a more expanded vowel space under the plain condition. This effect is particularly observed for the long vowel /i:/, where the articulatory-acoustic distance is greatest. These findings underscore the interaction between vowel length and fricative type, highlighting that long vowels are more susceptible to the influence of fricative type, which shapes their acoustic structure more extensively. Overall, the F2-F1 difference analysis illustrates how both vowel length and fricative type interact to affect vowel space configuration.

In section 3.4.4, it was observed that F2 measurements in most of the long vowels following the emphatic and non-emphatic fricatives exhibited greater movement in the trajectory compared to short vowels. To capture this non-linear variation, GAMMs were employed. The results of the GAMMs shed light on the distinct realisations of the emphatic contrast s~sˤ across dialects. F2 values were measured across all dialects, revealing a notable pattern: the tested dialects could be categorised into two groups based on their non-linear differences between the emphatic and non-emphatic fricatives.

The first group, comprising egca, moca, syda, joka, and tuns dialects, exhibited a pronounced degree of non-linear trajectory for the emphatic fricative /sˤ/, while lacking non-linearity in the non-emphatic counterpart. Conversely, the second group, including irba, kwur, and ombu syda dialects, displayed a non-linear trajectory for the non-emphatic fricative /s/, but lacked non-linearity in the emphatic context. Significant differences in height further differentiated the emphatic contrast for /a:/ data, indicating continuous significant differences between /sˤ/ and /s/ trajectories in egca, joka, moca, and tuns dialects, whereas in irba, kwur, ombu, and syda dialects, significant height differences were confined to the initial intervals of the trajectory.

Regarding acoustic measurements, although no correlates were identified within the fricatives themselves to mark the emphatic contrast, these findings corroborate previous literature. However, there was evidence of dialectal variation in the realisation of the s~sˤ contrast, particularly notable in egca, irba, moca, and syda dialects across various correlates such as fricative duration, COG, peak location (Hz), F1, and F3. Nevertheless, the variation in fricatives' realisation of the emphatic contrast was limited compared to the typology observed for plain~emphatic plosives contrast. It should also be noted that [s] spectra are known to be highly speaker-specific, likely due to unique anatomical characteristics, particularly dentition (Kavanagh, 2012; Smorenburg & Heeren, 2020). Such speaker-specific differences may act as a confounding factor in the interpretation of COG and peak frequency results, potentially accounting for much of the observed variation. This inherent variability underscores the complexity of interpreting these acoustic correlates, as some of the differences observed may be as much due to individual speaker characteristics as to dialectal variation. Future studies might consider controlling for or quantifying this factor, particularly by examining the influence of speaker-specific anatomy on [s] spectra.

From the acoustic results of this chapter, it was evident that the vowel formant information provided significant cues to mark the emphatic contrast s~sˤ. Consequently, the subsequent perceptual study in Chapter 4 will utilise resynthesized stimuli to ascertain whether the findings of the perceptual experiment align with acoustic. Additionally, it will investigate whether the information conveyed by vowels in the emphatic context is utilised in recognising emphasis. Furthermore, the study will explore the F2 continuum, manipulating its size difference and trajectory slope to determine whether listeners across all dialects attend differently to manipulated and genuine signals.

3.5 Conclusion

The chapter has examined the emphatic contrast between the fricative consonants, /s/ and /sˤ/ across eight Arabic dialects. Employing statistical methodologies including Linear Mixed Effects Models (LMER) and Generalised Additive Mixed Models (GAMMs), the study inspects a range of acoustic parameters such as Centre of Gravity (COG), Intensity, Peak Location, and Vowel Quality. Through a rigorous analysis, the research uncovers several noteworthy findings. Notably, significant main effects of gender, vowel type, and dialect on acoustic measures are revealed, with particular emphasis on the role of formant frequencies

F2 and F3 in consistently differentiating the emphatic contrast. While some measures such as COG and Intensity do not yield statistically significant results, the trajectories of F2, especially in long vowels following fricatives, exhibit distinctive patterns across dialects Moreover, the study highlights the essential role of vowel formants as significant cues in distinguishing the emphatic contrast, shedding light on their contribution to acoustic variability. Additionally, the analysis brings to the fore the considerable dialectal variation in the realisation of the s~sˤ contrast, particularly evident in measures like fricative duration and F1. Nevertheless, among this variability, F2 and F3 consistently emerge as robust markers of the emphatic contrast across all dialectal variants.

4 Chapter 4: The Perception Study

4.1 Introduction

According to the previous results from the acoustic chapter, the measures that exhibited significant information to mark the emphatic contrast s~sˤ were observed on the vowel formant information; vowel F1 tended to rise, whilst F2 was consistently lower in emphatic fricatives compared to non-emphatics. More specifically, the GAMM results revealed that the majority of the emphatic contrast effect was observable at different points along the trajectory of the vowel following the emphatic contrast /s/ ~ /sˤ/. The extent of this effect varied across dialects, both in the magnitude of the F2 difference and in the portion of the vowel's F2 trajectory that was affected. In some cases, the impact was observed only at the beginning of the trajectory, persisting for 20% to 50% of the F2 trajectory, while in others, the effect extended throughout the entire trajectory, *refer to section 4.5 in the acoustic study*. The findings also revealed that other variables, such as COG and fricative duration, could display some difference in the emphatic contrast, though these differences were statistically insignificant. This may suggest that the acoustic correlates of the fricative consonants do not play a significant role in identifying the distinction for the emphatic contrast. As a result, this perceptual investigation employing resynthesed stimuli will ascertain whether or not the perceptual experiment's findings map onto the observed acoustic properties. It will also determine whether the information offered by the vowels (in the vicinity of emphatics) is used to recognise emphasis. Specifically, the effect of F2 will be studied, and its size difference and trajectory slope will be manipulated, in order to determine whether listeners across dialects will attend differently to the manipulated and real signals.

4.2 Background

While the acoustic analysis of emphasis in Arabic has received considerable attention in the literature, very few studies have focussed on perceptual cues of emphasis. Jongman et al., (2011) for example, investigated the perception of emphasis (T & S) along with their plain-counterparts (t & s) using crossed-spliced CVC syllables in which emphatic/plain consonants were spliced onto plain or emphatic syllable portions, respectively, to create hybrid stimuli. Specifically, the study examined patterns such as C̣VC (where a subscript dot indicates emphasis) and hybrid patterns like C̣ṾC and ṾC̣, where emphasis is distributed across the syllable. The findings revealed that words with emphatic target consonants received significantly more emphatic responses than words with plain consonants. It was also found that words with emphatic C̣Ṿ or ṾC̣ portions received significantly more emphatic responses than words with plain CV or VC portions. In addition, the presence of an emphatic ṾC̣ or C̣Ṿ portion resulted in a dramatically greater percentage of emphatic responses (around 93%) than the presence of an emphatic consonant. Altogether, the effect of the vowel contributed more to the perception of the emphatic versus plain distinction than the target consonant. Additionally, the study reported that the findings on perception somehow correlated with their acoustic results; acoustically, emphasis is observed during the vowel only, whereas perceptually, listeners identify emphasis by relying on the entire portions, either ṾC̣ or C̣Ṿ but not as much with the emphatic consonant itself. It is worth mentioning that the researcher did not specifically explore the effect of the formant information of the vowel (itself) that occurs in the vicinity of the emphasis; instead, they relied on the entire word or syllable, containing both consonants and vowels, to check listeners' judgments of emphasis identification. Thus, the assumption that perception aligns with acoustic is not accurate. However, this current study will focus primarily on the vowels' formant information, particularly the F2 cue, to further investigate this phenomenon.

Using synthetic stimuli, Obrecht (1968), investigated the role of the F2 transition from an initial consonant to the following vowel, in the Lebanese dialect. He used a Haskins pattern playback synthesizer to create the stimuli modelled after a single Lebanese male speaker. He analysed different segments in a monosyllabic word context. The value of F2 was varied systematically in 120 Hz steps ranging from 1080 to 1800 Hz. The total duration of the syllable ranged from 300 to 400ms while the F2 transition duration was fixed at 70ms. F3 was flat and fixed at 3000Hz, whereas F1 varied depending on the vowel used. The results from the /tˤi: - ti:/ continuum suggested that perception of the emphatic versus non-emphatic distinction was categorical; the continuum endpoints were perceived as /tˤi:/ with low F2 and as /ti:/ with high F2 values. The turning point between the emphatic contrast was identified at around 1560 Hz of the F2 value. In other words, when F2 values were about 1560Hz, listeners perceived the stimuli as plain; but, when F2 values were lowered, they categorised the stimuli as emphatic. Yeou (1995) replicated Obrecht’s experiments for the Moroccan dialect of Arabic with the emphatic fricative contrast, /sˤi: ~ si:/. The results were similar to Obrecht’s. Yeou confirmed the role of F2 in a synthetic /sˤi: ~ si:/ continuum, and further showed that F1 alone was not a perceptual cue to the emphatic/plain distinction.

Using naturally produced utterances, Ali and Daniloff (1974) compared minimal pair emphatic/non-emphatic Iraqi words in which the target consonant occurred in either word-initial or word-final position and found that listeners correctly identified the missing consonant from truncated stimuli with about 68 per cent accuracy (73% for emphatic stems, 62% for plain stems). The results suggested that the emphatic consonant itself is not required for reasonably accurate perception of the emphatic and non-emphatic distinction. It was reported that neither vowel quality nor consonant type had an effect, nor did position of the target consonant. This suggests that the vocalic portion of the plain/emphatic words included sufficient cues for listeners to detect the distinction; so, the formant information influenced the listeners' judgment on which category of emphatic contrast to select.

Although there were some other perceptual studies on emphasis their focus was primarily to examine sociolinguistic impact where they looked at the social factor, such as gender. Their aim was to see how Arabic male and female listeners perceive the emphatic contrast (Kahn 1975; Alwan 1986; Wahba 1993; and Al-Wer 2000). One of the findings on Cairene dialect, reported by Kahn (1975), suggested that both male and female participants perceived the emphatic contrast equally and there was no main effect of gender.

The above reviewed studies are among the very few that considered the perceptual correlates of emphasis in Arabic. Some studies suggested that F2 transition is a robust perceptual cue where listeners can distinguish between the emphatic contrast with high accuracy. Other studies used crossed-spliced syllables to investigate the relative contribution of different cues to the emphatic contrast, but without stimuli (re)synthesis. Even though some studies reported that the findings of their perception results closely match their acoustic ones (e.g. Jongman et al., 2011) some perceptual findings were not consistent with the acoustic studies results. This may be due to different sounds (consonants and vowels) included in each study.

All in all, the studies summarized above show that prior research studies have typically concentrated on comparing emphasis perception in only one Arabic dialect each, and that most have focused on examining the emphatic plosives. Similarly, the main focus of prior research has typically been on the consonant itself (or the whole word as a perception unit) and on the F2 lowering of the following vowel. So far, no study has investigated the detailed characteristics of the vowel trajectory (in terms of different height and slope).

The current study will provide unique new data on the relative contribution of consonantal and vocalic information to the emphatic and non-emphatic distinction. The aim is to examine the emphatic fricative /sˤ/ and its plain counterpart /s/ perceptually across four Arabic dialects. It will use resynthetised stimuli, created by the author to explore participants' responses toward the cues typical of an emphatic and non-emphatic fricative in their own dialect, as well as the cues typical in other dialects. Two vowel qualities (/a:/ and /i:/) will be used in the present experiment and in the main control condition their F2 trajectories (only) will be manipulated. F1 and F3 will not be manipulated due to time constraints and the fact that they are not the primary focus of this study.

4.3 Rationale

Based on the acoustic results from the previous chapter, the measures which provided significant information to mark the emphatic contrast s~sˤ were shown to be the vowel formant information; vowel F1 tended to rise whereas F2 was observed to be consistently lower, in emphatic fricatives as compared to the non-emphatics. The GAMM results revealed that the majority of the emphatic contrast was observable at different points in the trajectory of the vowel following the emphatic contrast s~sˤ, and dialects differ in how large the F2 difference is and how much of the vowel is affected. More specifically, The GAMM results revealed that the majority of the emphatic contrast effect was observable at different points along the trajectory of the vowel following the emphatic contrast /s/ ~ /sˤ/. The extent of this effect varied across dialects, both in the magnitude of the F2 difference and in the portion of the vowel's F2 trajectory that was affected. In some cases, the impact was observed only at the beginning of the trajectory, persisting for 20% to 50% of the F2 trajectory, while in others, the effect extended throughout the entire trajectory. The results also suggested that the other measures such as COG and fricative duration exhibited some different realisations in the emphatic contrast, although these differences remain statistically insignificant. This suggests that the fricative consonants themselves did not play a major role in identifying the distinction for the emphatic contrast. Thus, this perceptual study involving resynthesised stimuli will confirm whether or not the findings of the acoustic experiment map onto perception. In particular, it will determine whether emphasis recognition is attributed to the information provided in emphatic consonants or the vowels following the emphatics. Moreover, the F2 size difference and trajectory slope will be manipulated with a view to ascertaining whether the listeners across all dialects will attend to the manipulated cues differently. The overall prediction is that listeners will attend to their own native cues with greater accuracy than listeners who are not used to hearing them in their dialects. This prediction was motivated by the overall results of the Acoustic study. This experiment aims to address the following questions.

4.3.1 **Research questions:**

1- Are listeners more 'accurate' at identifying emphatic words when given the cues to emphasis in their own dialect versus other dialects, when presented with real stimuli?

- Does this vary by vowel [a:~i:]?

2- Are listeners more 'accurate' at identifying emphatic words when given the cues to emphasis in their own dialect versus other dialects, when presented with stimuli in which only F2 is manipulated?

- Does this vary by vowel [a:~i:]?

In order to answer these questions, the experiment's design aims to discover what vowel characteristics are crucial to emphasis perception and in what conditions.

4.4 Method

4.4.1 **Participants**

This experiment recruited 119 participants from three Arabic dialects, Moroccan (24 participants), Syrian (39 participants), and Iraqi (56 participants). All participants were native listeners/speakers of their own dialects. The approach of recruitment was by sending emails via academic organisations; the author asked specific contacts in various countries to send out emails to their students. All participants were in the 18 to 45 age range. It was ensured that none of these participants suffer from any hearing or speech impairment. The three dialects were selected because for each, the F2 trajectory in the acoustic study was different to other dialects; each represents a distinct pattern of vowel F2 height differences between /s/ and /sˤ/, making it interesting to explore how listeners attend to these varying acoustic signals. This provides distinct comparisons to explore in the perception study, as illustrated in Figure 39, which displays plots for only the three selected dialects plots extracted from Figure 37 in the acoustic study. As we can see, these plots display different F2 trajectory patterns for each vowel, /a:/ and /i:/. Specifically, in some dialects the trajectory of the emphatic vowel */a:/* starts low at around 1250 Hz and never becomes high (e.g., Syrian), while in other dialects F2 starts at a medium height (1380 Hz) and the difference with the plain counterpart trajectory remains medium throughout (e.g., Moroccan). In contrast, in other dialects the effect of a different F2 height can be seen only in the first 30% of the */a:/* trajectory, as in the Iraqi. As for the vowel */i:/* trajectory, the affected proportion of the F2 trajectory after the emphatic fricative is seen only in the first 30% of the vowel in the Syrian dialect (with a small F2 difference) but in the first 50% of the vowel in Iraqi (and with a bigger F2 difference), while the F2 difference is maintained throughout the whole vowel in

Moroccan (albeit with a small F2 difference). In Appendices 10.1, 10.2, and 10.3, you can find raw (unedited) spectrograms from specific speakers of each target dialect, illustrating the F2 trajectories for the vowels /a:/ and /i:/ in both emphatic and plain conditions.

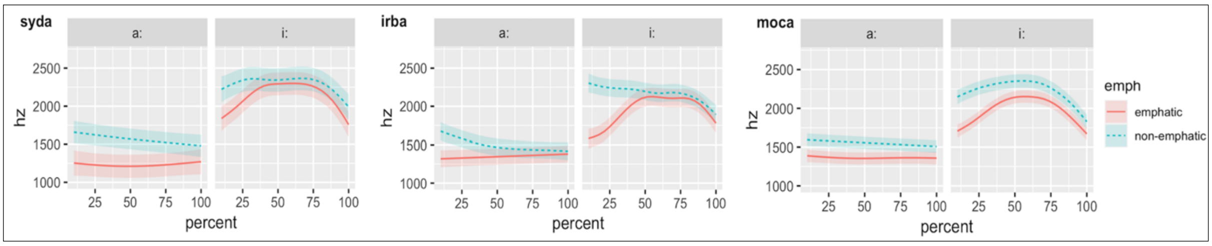


Figure 39 GAMM fits of the effects of normalized percent-by-speaker and percent-by-word. These are the trajectories for the dialects: (syda = Syrian), (moca = Moroccan), and (irba = Iraqi).

4.4.2 **Stimuli**

This experiment used a subset of the recordings from the acoustic study as base stimuli. The monosyllabic pairs /siin ~ sˤiin/ and /saar ~ sˤaar/ were used in this experiment since the aim is to examine the characteristics of the F2 trajectory of the long vowels /a:/ and /i:/ in the emphatic and non-emphatic contexts. According to the Zipfian distribution, these two pairs are among the words that are most frequently occurring in Arabic language (refer to Table 21 below). The stimuli also included another two pairs as distractors, namely, /tiin ~ tˤiin/ and /taab ~ tˤaab/.

The Plain consonant				***The Emphatic consonant***				***Position & Vowel***	
Arabic	***item***	***type***	***zipf***	***Arabic***	***item***	***type***	***zipf***	***position***	***Vowel***
سين	**siin**	**target**	**3**	صين	**sˤiin**	**target**	**4**	**initial**	**i:**
سار	**saar**	**target**	**4**	صار	**sˤaar**	**target**	**4**	**initial**	**a:**
تين	**tiin**	**filler**	**4**	طين	**tˤiin**	**filler**	**4**	**initial**	**i:**
تاب	**taab**	**filler**	**4**	طاب	**tˤaab**	**filler**	**4**	**initial**	**a:**

Table 21: Stimuli that are used for the perception experiment.

Each target word was first extracted from the acoustic study carrier phrase. Both waveforms and spectrograms were examined in order to determine the boundary between the carrier phrase and the target word. In the case of both word- initial stops and fricatives, the onset of the first consonant was extracted after the visible burst in the preceding stop sound /b/ and at a start of a visible energy frication seen in the waveform. The word-final boundary was placed at the end of the visible burst for word-final stops, at the end of visible frication for word-final fricatives, and at the end of the fainter formants for the word-final nasals. Since these stimuli were followed by the closure and burst of the initial stop in the carrier phrase, the boundary was always clearly visible in the waveform.

The initial consonant C_1 in the target monosyllable C_1VC_2 refers to the /s/ sound: in all manipulated stimuli C_1 was extracted from an Omani dialect recording, since that dialect displayed values closest to the average COG value across all dialects. The target consonant C_2 refers to the final consonants (/r/, or /n/) occurring in /saar/ and /siin/ which were extracted from the same as for C1' i.e. from Omani dialect. As for the base stimulus vowel sound, this was also obtained from one male speaker from the Omani sound files. This particular speaker was selected because the author (as a native Arab listener) judged him as highly intelligible. Then, the vowels (only) within the four target base stimuli C_1VC_2 syllables were resynthesized and manipulated by the author to create four conditions for each vowel /a:/ and /i:/.

Each condition has its vowel trajectory manipulated based on the values found in the acoustic study results. Specifically, the F2 values were taken from the GAMM output, reproduced in Figure 39 above. Before creating the stimuli, the author re-ran a GAMM model on the plain fricative results only, without including fixed factors for variables like dialect, or random factors using the following model structure.

- *Model = gam (hz ~ Vowel + s (percent) + s (percent, by = Vowel, bs = 'cr')*

The aim was to obtain F2 formant values to use in the plain base-stimuli that are averaged across all dialects' realization of each vowel. We treat the plain fricative as the 'default' vowel so we can see how much the manipulated conditions vary from the base (e.g., creating more emphatic-like sound). The results of this approach are illustrated in Figure 40, and the corresponding manipulated spectrograms and plots for these two stimuli are provided in Appendices 10.4 and 10.5 (for /a:/) and Appendices 10.6 and 10.7 (for /i:/). Using the vowel from the non-emphatic context (siin/saar) to create the base stimuli is crucial to remove any potential confounds such as F3 differences and to ensure that listeners are exposed to differences in F2 only. As for F1 and F3, neither were manipulated but rather normalised by obtaining their average values across all dialects for both /a:/ and /i:/[6]; values for these can be seen in Appendices, 10.8 and 10.9.

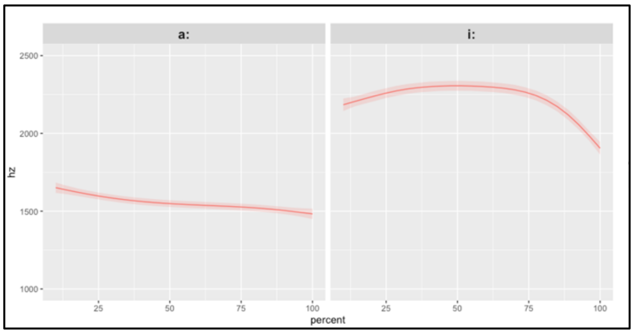


Figure 40 Base stimuli values for F2 in vowels a: and i:

4.4.3 **Vowel re-synthesis**

After creating the base stimuli, the author created the four conditions for each vowel /a:/ and /i:/ using the values that were observed from the acoustic results. As explained before, the sound files that were utilised for extracting target sounds were from a speaker of the Omani dialect. The Omani dialect is not among the ones that were tested, and its trajectory in plain condition was closest to the overall shape of the base one. Also, relative to the overall mean value for fricative COG that was found across all dialects (μ = 7108 Hz) the Omani mean value was set in the middle compared to the other dialects, as shown in Table 22 and Figure 41 below.

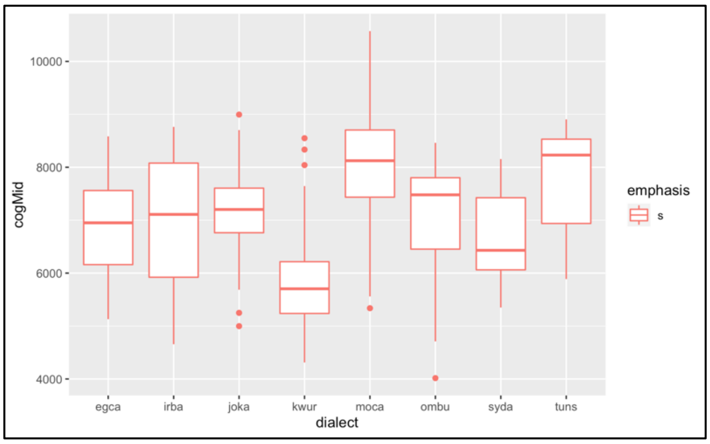


Figure 41 Mean value of COG in for each dialect

[6] As shown in Figure 40, the observed shift or drop in the F2 frequency of the vowel may be attributed to the influence of the following nasal sound. Carignan and Zellou (2023) explain that vowel nasalisation causes several acoustic modifications, including a decrease in formant amplitudes, an increase in formant bandwidths, changes in formant frequencies, a shift of spectral energy toward lower frequencies, and the introduction of poles (formants) and zeros (anti-formants) into the acoustic spectrum (citing Fujimura & Lindqvist, 1971; Maeda, 1993; Feng & Castelli, 1996; Chen, 1997; Styler, 2017; Carignan, 2018). Additionally, studies such as Serrurier and Badin (2008) have shown that F2 lowering in nasalised vowels is often a result of velopharyngeal coupling (or nasalisation in general), where the oral and nasal cavities are connected. Moreover, Delvaux (2009), as cited in Carignan (2018), demonstrated that F2 lowering alone is sufficient to trigger the perception of nasality in synthesised French vowels. Therefore, it can be speculated that the F2 drop in the vowel might be influenced by the presence of the following nasal sound.

Egyptian	Iraqi	Jordanian	Kuwaiti
6938.480	6937.051	7177.347	5963.466
Moroccan	Omani	Syrian	Tunisian
8009.763	7060.559	6682.580	7832.638

Table 22 Mean COG for each dialect (in the acoustic study)

As discussed earlier in the previous sections, this experiment aims to manipulate the F2 trajectory height and its duration through the vowel, for each vowel. The manipulated values were derived from the acoustic output as was seen in Figure 39 above. Thus, the vowels /a:/ and /i:/ were extracted from the base stimuli, after moving the start/end to a zero-crossing point. The duration of the vowel was then manipulated (lengthened to 2 seconds) using the command *convert* where the standard pitch was set at 180 Hz and the factor value increased up to 2 seconds. This command includes a function named *overlap-add* or *Time-Domain Pitch-synchronous Overlap-And-Add* (Moulines & Charpentier, 1990). This function creates new points on the Pitch Tier and copies the period information pulses from the source sound to the manipulated one.

After that, the vowel sound was converted to LPC so as to extract the source (glottal pulses) from the original sound and is achieved by inverse filtering. Since the aim is to manipulate the F2 information, the main Sound object was converted to a Formant object so as to perform a short-term spectral analysis on the F2 information. Within this step, the author used a time step of 25 percent of the analysis window length and extracted five formants per frame; this is ideal for any human speech analysis. The maximum frequency of the formant search range was set at an average celling of 5000 Hz since the manipulated vowels are all calibrated for an adult male speaker. For more details about the relations between the vocal tract size and the average ceiling frequency for male and female speakers, see Escudero et al. (2009). Since the vowel intensity increased to 77dB as a result of the manipulation, intensity was rescaled and normalised to 64dB, which was its original value.

After having manipulated the vowel, the initial fricative and the final non-target consonant were combined with the vowel with smooth cross-fading between the sounds; the function *concatenate* with 0.01*ms* overlap was used, to remove any potential gaps between the spliced sounds during which the earlier sound fades out and the later sound fades in.

For the F2 trajectory manipulation, the following lines of code were used to decrease the F2 values of the trajectory and to manipulate the size of the F2 effect according to the required condition. This contrasts with other steps of the manipulations (e.g., the duration effect) which were manipulated manually by the author.

a) if row = 2 then self - 500 else self fi

b) if row = 2 and col = 5 then self - 500 else self fi

It was observed that the overall F2 frequency fall ranged from 200 Hz to 500 Hz but that the extent of the initial F2 effect varies across all dialects; some show an effect of F2 height for the first 30%, while the effect is seen in the first 50% in some others. The F2 height is also seen throughout the whole trajectory in some other dialects. Thus, the following vowel trajectory were manipulated according to the following values.

4.4.4 Vowel /a:/

There are four conditions created for this vowel. For the first condition (Iraqi), the F2 trajectory (cf. the plain trajectory) is lowered by around 500 Hz from the onset and starts to rise after 30% of the vowel duration, merging to the base stimuli (*Low_30*). The trajectory of the second condition (Syrian) is manipulated to start low with a 500 Hz difference (from the plain trajectory) and never becomes high; the large F2 difference is seen throughout the whole trajectory (*Low_100*). The third condition (Moroccan) has the F2 trajectory dropped by around 200 Hz compared to the plain base stimuli and it remains at a medium height difference for the entire vowel (*Medium_100*). Finally, the fourth condition is the manipulated plain version with values set to resemble the plain base stimulus trajectory. Figure 42 provides an example of stylised representations of F2 trajectories for the four conditions (represented by dotted lines) in comparison to the base stimulus (shown as a solid line). This figure is manually drawn to illustrate how the conditions would differ relative to the base stimuli. In Appendices 10.4 and 10.5, you can find spectrograms and plots for the manipulated F2 trajectories of the vowel /a:/. Notably, the spectrogram for each manipulated condition is displayed separately in Appendix 10.4, while all manipulated conditions, along with the base stimuli, are plotted together in Appendix 10.5.

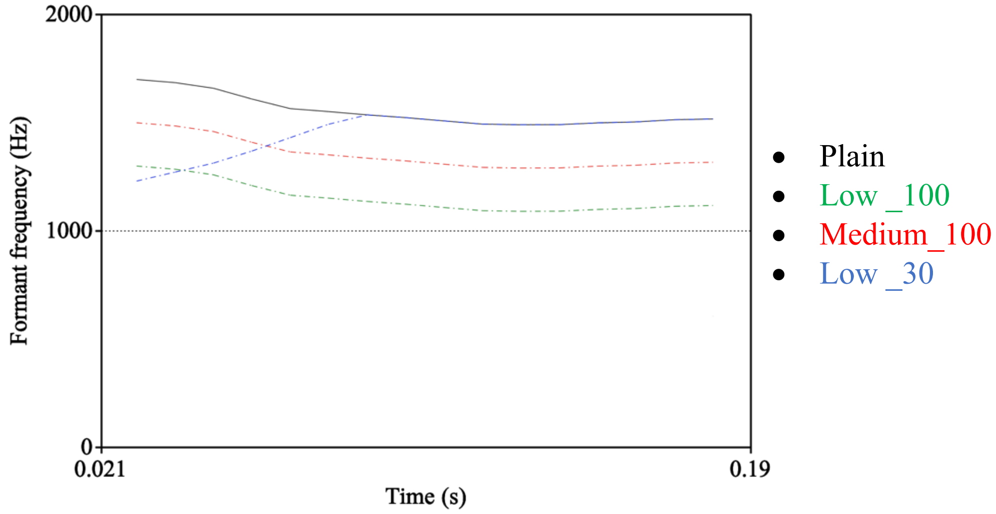


Figure 42 A stylised representations of F2 trajectories for the vowel /a:/ in the four conditions: Low_100 = (Syrian), Medium_100 = (Moroccan), and Low_30 = (Iraqi)

4.4.5 **Vowel /i:/**

Another four conditions were also created for the vowel /i:/. The trajectory of the first condition (Syrian) is dropped by 200 Hz compared to the plain base stimulus and merges with the base stimulus trajectory after 30% of the vowel duration (*Medium_30*). The second condition (Iraqi) is manipulated to have a 500 Hz difference compared to the plain trajectory and starts to become high at around 50% through the vowel (*Low_50*). As for the third condition (Moroccan), the trajectory is 200 Hz lower than the base stimuli and remains at this medium height difference from the plain trajectory for the entire vowel (*Meduim_100*). The final condition is the manipulated plain stimulus that has its trajectory set with the values of the base one. Figure 43 presents a manually drawn example of stylized representations of how the conditions compare to the base stimuli. The four contours represent the vowel /i:/, with the three emphatic conditions depicted in dotted lines and the base stimulus shown as a solid line. In Appendices 10.6 and 10.7, you can find spectrograms and plots for the manipulated F2 trajectories of the vowel /i:/. Notably, the spectrogram for each manipulated condition is displayed separately in Appendix 10.6, while all manipulated conditions, along with the base stimuli, are plotted together in Appendix 10.7.

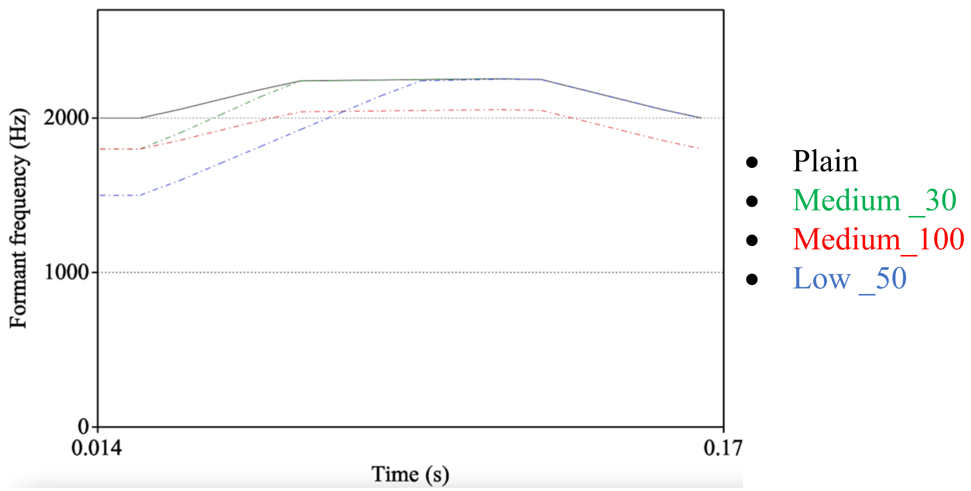


Figure 43 A stylised representations of F2 trajectories for the vowel /i:/ in the four conditions: Meduim_30= (Syrian), Medium_100 = (Moroccan), and Low_50 = (Iraqi)

Because the trajectory of each condition is similar to dialects such as Syrian, Iraqi, and Moroccan, and because there is an attempt to avoid any potential confounds created by sequence and order effects, the experimental procedure used these stimuli across three separate blocks, namely, Syrian-like, Moroccan-like, and Iraqi-like that each include two target vowels /a: and i:/, and eight trajectories (2 target items x 2 repetitions, plus 4 distractors), yielding a total of 16 stimuli for each block. As for the distractors, they were included solely in the manipulated stimuli, including two natural stimuli and two with their vowel trajectories spliced from the main stimuli conditions. Thus, the manipulated vowel trajectory specific to each block—Syrian-like, Moroccan-like, and Iraqi-like—was used for the distractors occurring within the same block.

The experiment also included three additional 'mini-blocks', each containing four 'real' (i.e., unedited and unmanipulated) tokens for the target contrasts /siin ~ sˤiin/ and /saar ~ sˤaar/. These mini-blocks were designed to serve a specific purpose: to compare participants' responses to the resynthesized stimuli with responses to authentic, ecologically valid stimuli, thus creating a baseline for assessing the perceptual realism of the manipulated stimuli. Also, the rationale behind including these real stimuli was to evaluate how much the manipulated stimuli differ from or align with naturally produced tokens in terms of listener perception. The process of preparing the real stimuli for both vowels (/a:/ and /i:/) involved selecting audio recordings from the acoustic dataset. The recordings were chosen based on their quality (e.g., minimal background noise), as determined by the author. Once the eligible speaker's tokens were identified, the target words were extracted from the carrier phrases.

This extraction was executed after moving the start and end points of each word to a zero-crossing to ensure natural sound transitions. Importantly, no further manipulation was conducted on these real stimuli to preserve their natural acoustic characteristics. It is crucial to clarify that no distractors were included in these real stimulus blocks. The decision to not include distractors was intentional, as in this part of the experiment the objective was to maintain focus on the target stimuli and maximize participant exposure to the relevant phonetic contrasts. For more details, refer to Table 23.

Overall, therefore the main part of the experiment includes 3 block types x 2 vowels x 8 trajectories for a total of 48 manipulated stimuli, then another three mini blocks each containing 12 real stimuli, yielding a total of 60 stimuli. Table 23 below shows a summary of all the block types along with the items and conditions that are used to create the experiment. Letters in upper case in the item column represent the segments that are emphatic; letters in lower case represent the segments that are plain. The condition column represents the trajectory types, manipulated or real.

No.	Block-type	Item	Consonant	Vowel	Condition	Item Type	Stimuli Type
1	Syrian-like	si:n	fricative	i:	plain	target item	manipulated
2	Syrian-like	sˤi:n	fricative	i:	medium_30	target item	manipulated
3	Syrian-like	ti:n	plosive	i:	longlag	natural distractor	manipulated
4	Syrian-like	tˤi:n	plosive	i:	shortlag	natural distractor	manipulated
5	Syrian-like	ti:n	plosive	i:	plain	spliced distractor	manipulated
6	Syrian-like	tˤi:n	plosive	i:	medium_30	spliced distractor	manipulated
7	Syrian-like	sa:r	fricative	a:	plain	target item	manipulated
8	Syrian-like	sˤa:r	fricative	a:	low_100	target item	manipulated
9	Syrian-like	ta:bit	plosive	a:	longlag	natural distractor	manipulated
10	Syrian-like	tˤa:bit	plosive	a:	shortlag	natural distractor	manipulated
11	Syrian-like	ta:bit	plosive	a:	plain	spliced distractor	manipulated
12	Syrian-like	tˤa:bit	plosive	a:	low_100	spliced distractor	manipulated
13	Moroccan -like	si:n	fricative	i:	plain	target item	manipulated
14	Moroccan -like	sˤi:n	fricative	i:	medium_100	target item	manipulated
15	Moroccan -like	ti:n	plosive	i:	longlag	natural distractor	manipulated
16	Moroccan -like	tˤi:n	plosive	i:	shortlag	natural distractor	manipulated
17	Moroccan -like	ti:n	plosive	i:	plain	spliced distractor	manipulated
18	Moroccan -like	tˤi:n	plosive	i:	medium_100	spliced distractor	manipulated
19	Moroccan -like	sa:r	fricative	a:	plain	target item	manipulated
20	Moroccan -like	sˤa:r	fricative	a:	medium_100	target item	manipulated
21	Moroccan -like	ta:bit	plosive	a:	longlag	natural distractor	manipulated
22	Moroccan -like	tˤa:bit	plosive	a:	shortlag	natural distractor	manipulated
23	Moroccan -like	ta:bit	plosive	a:	plain	spliced distractor	manipulated
24	Moroccan -like	tˤa:bit	plosive	a:	medium_100	spliced distractor	manipulated
25	Iraqi-like	si:n	fricative	i:	plain	target item	manipulated
26	Iraqi-like	sˤi:n	fricative	i:	low_50	target item	manipulated
27	Iraqi-like	ti:n	plosive	i:	longlag	natural distractor	manipulated
28	Iraqi-like	tˤi:n	plosive	i:	shortlag	natural distractor	manipulated
29	Iraqi-like	ti:n	plosive	i:	plain	spliced distractor	manipulated
30	Iraqi-like	tˤi:n	plosive	i:	low_50	spliced distractor	manipulated
31	Iraqi-like	sa:r	fricative	a:	plain	target item	manipulated
32	Iraqi-like	sˤa:r	fricative	a:	low_30	target item	manipulated
33	Iraqi-like	ta:bit	plosive	a:	longlag	natural distractor	manipulated
34	Iraqi-like	tˤa:bit	plosive	a:	shortlag	natural distractor	manipulated
35	Iraqi-like	ta:bit	plosive	a:	plain	spliced distractor	manipulated
36	Iraqi-like	tˤa:bit	plosive	a:	low_30	spliced distractor	manipulated
37	Syrian-real	si:n	fricative	i:	plain	target item	real
38	Syrian-real	sˤi:n	fricative	i:	emphatic	target item	real
39	Syrian-real	sa:r	fricative	a:	plain	target item	real
40	Syrian-real	sˤa:r	fricative	a:	emphatic	target item	real
41	Moroccan -real	si:n	fricative	i:	plain	target item	real
42	Moroccan -real	sˤi:n	fricative	i:	emphatic	target item	real
43	Moroccan -real	sa:r	fricative	a:	plain	target item	real
44	Moroccan -real	sˤa:r	fricative	a:	emphatic	target item	real
45	Iraqi-like	si:n	fricative	i:	plain	target item	real
46	Iraqi-like	sˤi:n	fricative	i:	emphatic	target item	real
47	Iraqi-like	sa:r	fricative	a:	plain	target item	real
48	Iraqi-like	sˤa:r	fricative	a:	emphatic	target item	real

Table 23 summary of all the block types, items and conditions that are used to create the experiment

4.4.6 **Experiment implementation using Gorilla Platform**

The Gorilla Experiment Builder was used to implement the perception study (Anwyl-Irvine et al., 2019). Three main steps were undergone throughout the process of building the experiment, creating the consent form as well as the participants questionnaire, building the main task that includes the stimuli, and finally using the graphical drag-and-drop interface to link between the tasks. Since the current study concerns examining listeners of Arabic dialects the study was built using Arabic orthography and all the instructions were provided in Arabic text. An English translated version of the consent and questionnaire forms are included in the Appendices, 10.10, and 10.11.

The first step was to create a consent form, which appears as the first page on the screen. After participants confirm their consent by clicking the "Agree" button, the participant questionnaire form is presented. The aim of this form is to collect essential information about each participant's place of residence, and the dialect the participant speaks. The form also included two (closed-ended) questions to measure participants' confidence in identifying and exposure to Arabic dialects. To obtain a numeric rating, the participant was presented with a 'Text entry box' and requested to simply write whatever rating most fits their opinion, but it was made sure that only numeric values range from 1 – 100 could be inputted. The aim here is to determine the level of confidence of participants identifying other Arabic dialects; the responses are a potential explanatory variable accounting for some of the variation that might be observed in the responses. The form also included a question related to the degree of exposure to each other dialects; this will also help the researcher to interpret the results. For more details, the questionnaire form was included in the Appendices, 10.10 and 10.11.

The second step in creating the experiment was to build the task section which is the most important part where participants will have the chance to listen to the stimuli. The task included a series of displays, the instruction, the trials, and the debrief. In the first display, there were two screens: the instructions which participant should read and then click on the start button, then the timeline screen where participants were given 3 seconds before moving to the next display to start listening to the stimuli. In the second display, there is a fixation screen created to show for 250 milliseconds followed by a trial screen which has three zones: the stimuli, and two other buttons to click on. The trial screen was driven by the spreadsheet where all the names of sound files and stimuli were included in spreadsheet rows. The procedure here was to instruct the spreadsheet to apply the same structure of the two screens and loop through all the stimuli allowing participants to listen to them all. The randomised-trials column within the spreadsheet was used to randomise all trials, including for the main stimuli and the fillers.

4.4.7 Procedure

After designing the experiment using Gorilla software, the URL of the experiment was sent to the assigned participants along with instructions for performing the test. Participants were instructed to use headphones or earphones and to take the test in a quiet room. This test is an Identification task which is oriented to make the participant recover a very robust internal representation (e.g., a phoneme). Participants listen to a certain sound and choose a label from a range of two possibilities. The study included two parts: three main blocks of manipulated sounds, and another three mini blocks of real sounds. Participants were presented first with three main blocks, each has 16 trials. The blocks as well as the trials within them were presented randomly. Within each trial, orthographic representations of the two members of the relevant minimal pair were shown in Arabic script on the screen. The location of the response alternatives was randomised across trials. After listening to the audio stimulus, participants would indicate the word heard by them (word containing an emphatic or non-emphatic consonant). They would do that by clicking on the appropriate word displayed on their screen. The next trial would then commence automatically following a gap of 250 milliseconds. When participants finished responding to the first three blocks, they were instructed that those first three blocks represented the main stimuli where the sounds were manipulated, and that they would now hear another three mini blocks with real words, thus, they would hear to a total of 60 trials. The design of the mini blocks had the same settings as the main one but with no randomisation of the orders of the blocks; the order is fixed, starting with the Syrian block, followed by the Moroccan block, and then the Iraqi block. Participants were started off with five practice trials (which include manipulated stimuli) to familiarize themselves with the nature of the stimuli and the presentation rate. The practice stimuli were not used in the main experiment.

4.4.8 Data Analysis

Participant responses were inputted into R software for statistical analysis using multiple R packages. The data was modelled using a Generalised Logistic Mixed-Effects Regression Model (GLMER) for each dialect block, (Moroccan, Syrian, and Iraqi). Recall that each block represents two vowel trajectory manipulation types that corresponds to use of stimuli in a particular dialect, or real stimuli from a particular dialect for the mini blocks. The GLMER model was selected because it is well-suited for a binary categorical dependent variable and can account for both fixed effects (predictors) and random effects. Also, similar perceptual studies (Ryu, 2018; Seo et al., 2022; Zhou, 2022) have employed this type of model and highlighted its effectiveness for their data analysis.

4.4.8.1 Fitting the Best Fixed Effects Model

We introduced specific predictors to the model and in order to reach for the model fit, multiple structures were compared and tested by Likelihood Ratio Test (LRT). The first presented model is the Null model that is constructed without factors of the author interest, and it contains only the participant answer response, as a binary categorical dependant variable as well as the random intercepts for item and subject. The predictors of interest were then included one at time to test if they improved the model further; the predictors are dialect blocks where each represents a dialect (Moroccan, Syrian, and Iraqi), the listeners' dialect (Moroccan, Syrian, and Iraqi), the trajectory type (emphatic or plain), and finally the vowel type (/a/ and /i/). The structure of the second model includes the interaction between the predictor's dialect blocks and trajectory. The resulting output suggests that adding this interaction improved the fit of the model significantly ($\chi^2(5) = 56.658$, $p < 5.948e-11$). This indicates that there is a considerable interaction between trajectory and each dialect block, see Table 24 below.

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
nullmodel	2810.4	2828.0	-1402.2	2804.4	-	-	-
Model: two interactions	2763.8	2810.6	-1373.9	2747.8	56.658	5	5.948e-11 ***

Table 24 ANOVA test results for comparing the nullmodel, and model with two-way interaction.

The predictor " listeners' dialect" was then added to the model, creating interaction among the three variables. This addition aimed to explore whether there exists a relationship between listeners' accuracy of response and the dialects they speak. Notably, this inclusion resulted in an enhancement of the model's fit ($\chi^2(12) = 112.3$, $p < 2.2e-16$), as presented in Table 25 below.

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
Model: two interactions	2763.8	2810.6	-1373.9	2747.8	-	-	-
Model: three interactions	2655.4	2772.4	-1307.7	2615.4	112.3	12	2.2e-16 ***

Table 25 ANOVA test results for comparing the two-way interaction, and model with three-way interaction.

Finally, the vowel type variable was added to the model and the resulting output in Table 26 suggests that adding this predictor further improved the model ($\chi^2(1) = 8.9524$, $p < 0.002771$), and thus the best-fitting model for the response was determined by considering the interaction between 'Dialect Block', 'trajectory', and 'Listeners' Dialect', while also accounting for the 'Vowel' predictor. Additionally, we included random intercepts for both 'item' and 'subject'. This model [7]was fitted using family = binomial argument to fit a GLMM with a binomial error distribution, which is appropriate for modelling binary outcomes. This allows the model to estimate the probabilities of success (correct responses) as a function of the predictor variables and random effects.

Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
Model: three interactions	2655.4	2772.4	-1307.7	2615.4	-	-	-
Best Model Fit	2648.5	2771.3	-1303.2	2606.5	8.9524	1	0.002771 **

Table 26 ANOVA test results for seeing model improvement by adding the vowel type as predictor.

It is worth noting that another model was fitted to examine the interaction between the factor of Exposure, the Dialect block, and Listeners' dialects, in order to see the effects of this factor on listeners' responses. The same approach was also applied to the other factors: Attitude and Confidence.

After fitting the Generalized Linear Mixed Effects Regression (GLMER) model, pairwise comparisons were conducted to evaluate differences between specific levels of the fixed effects, particularly Dialect block, trajectory, Dialect's listeners, and Vowel type, as well as

[7] Best Model Fit <- glmer(response ~ Dialect Block'* listeners' dialect * trajectory + Vowel type + (1|item) + (1|subject), data = data, family = binomial)

their interactions. These comparisons were generated using the **emmeans** package in R, which calculates estimated marginal means (EMMs)—predicted probabilities of the dependent variable (participant responses) averaged across the levels of other variables in the model. Following model fitting, the EMMs were computed using the emmeans() function as follows:

```
emm <- emmeans(m1, ~ Dialect block * trajectory * Dialect's listeners + Vowel type)
```

This computation produced the EMMs for all combinations of the levels of Dialect block, trajectory, Dialect's listeners, and Vowel type. Pairwise comparisons were then performed using the pairs() function:

```
pairs(emm, by = c("trajectory", "Dialect"))
```

The by argument specified that comparisons were conducted within each level of trajectory and Dialect block, enabling a detailed examination of the differences between the levels of Dialect's listeners and Vowel type within these contexts. This approach allowed for the assessment of the statistical significance of these differences while accounting for the complex structure of the GLMER model, which included random effects for item and subject. The integration of EMMs and pairwise comparisons provided a robust framework for analyzing and interpreting the intricate interactions between the fixed effects and their implications for the dependent variable.

4.5 Results

The results are presented in the following structure: section 1 will present the results of the real trajectory for each dialect block for the vowel /a/ followed by the results of the real trajectory for the vowel /i/. Section 2 will then present the results of the manipulated trajectory for each vowel /a/ and /i/ across the three dialect blocks. A summary and an interim discussion follow each results section, setting the stage for the subsequent main discussion section. The aim here is to see how listeners who are speakers of Moroccan, Syrian, and Iraqi Arabic, perceive the F2 trajectories of their own and other dialects in both stimulus types (real and manipulated). The mean and standard deviation (SD) values for the Correct responses of each dialect's listeners, categorized by vowel type and trajectory, for both real and manipulated stimuli, can be found in Appendices 10.6 and 10.7.

4.5.1 **Results of Real stimuli**

4.5.2 **Results: Real stimuli /a:/**

4.5.2.1 Dialect Block: Moroccan real /a:/

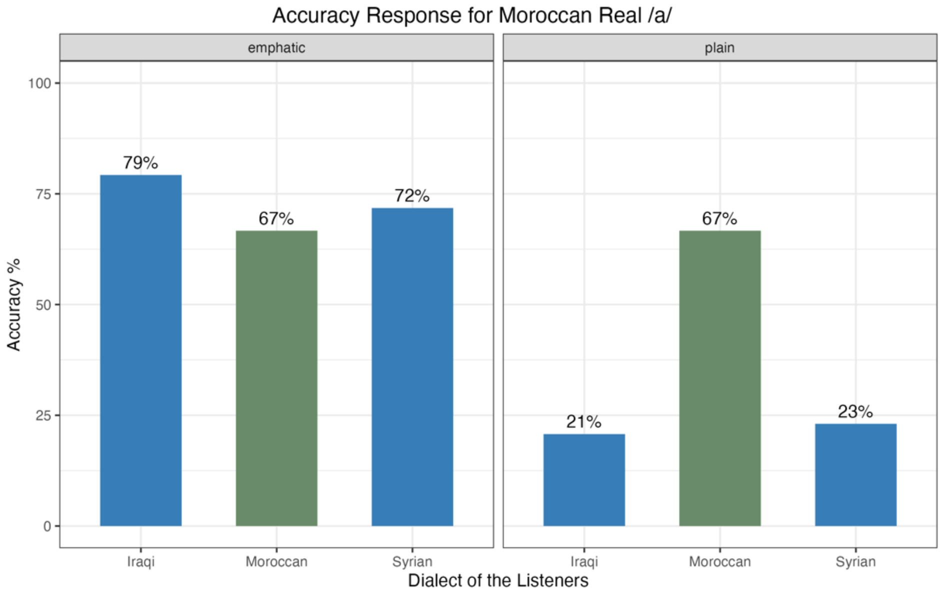


Figure 44 Accuracy Response for each dialect listeners for the Moroccan vowel /a/

Figure 44 above shows that for the emphatic condition, the Iraqi listeners achieved the highest percentage of accuracy at 79%, followed closely by Syrian listeners with 72%. Moroccan listeners had slightly lower percentages of 67%. These results indicate a relatively good overall accuracy in perceiving the Moroccan emphatic vowel with Iraqi standing the best chance. It is also interesting that Moroccan listeners did not have the highest correct responses even though this trajectory is their own native one. However, the pairwise results in Table 27 indicate that the Moroccan-Iraqi contrast is not significant ($\beta = 0.6466$, $SE = 0.644$, $p < 0.747$). Similarly, the Moroccan-Syrian contrast is not significant ($\beta = -0.2412$, $SE = 0.653$, $p < 0.9828$). This indicates that all listeners perform equally well at identifying the Moroccan emphatic vowel.

On the other hand, the performance of the listeners in perceiving the plain trajectory shows variability. Moroccan listeners achieved the highest score of accuracy at 67%. In contrast, listeners of the dialects Syrian and Iraqi had lower percentages of 23%, and 21% respectively. These findings suggest that the ability to accurately perceive the plain trajectory differs among dialect listeners, with only Moroccan listeners who demonstrated higher proficiency than others. This is confirmed by the pairwise results; the Moroccan-Iraqi results show a significant difference (β = -2.0329, *SE* = 0.644, *p* < 0.0045) and the Moroccan Syrian contrast results do also (β = 1.8971, *SE* = 0.667, *p* < 0.0123).

Overall, the results indicate that the Moroccan listeners generally exhibit a higher level of accuracy than other dialect listeners in perceiving the plain Moroccan /a:/ trajectory but have no advantage with the emphatic. The other dialect listeners show high accuracy responses for the emphatic trajectory but not the plain one. This means that all dialect listeners can identify the emphatic trajectory but erroneously identify the plain Moroccan /a:/ as emphatic unlike the Moroccan listeners. Figures 45 and 46 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Moroccan dialect and each dialect of the speaker.

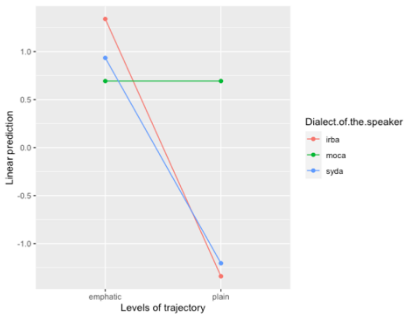


Figure 45 Predicted Probability: Moroccan by each Listener Group Dialect

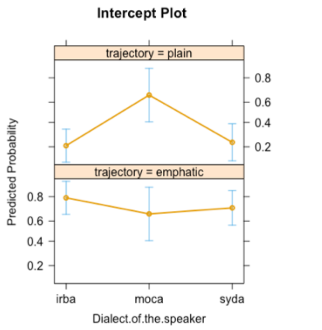


Figure 46 Linear Predicted: Moroccan by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Moroccan	Iraqi - Moroccan	0.647	0.644	1.004	0.5743
Emphatic	Moroccan	Moroccan - Syrian	-0.241	0.653	-0.369	0.9276
Emphatic	Moroccan	Iraqi - Syrian	0.405	0.491	0.825	0.6873
Plain	Moroccan	Iraqi - Moroccan	-2.033	0.644	-3.157	0.0045
Plain	Moroccan	Moroccan- Syrian	1.897	0.667	2.846	0.0123
Plain	Moroccan	Iraqi - Syrian	-0.136	0.509	-0.267	0.9615

Table 27 Pairwise comparisons Results: Moroccan real /a:/ stimuli by each dialect Listener Group

4.5.2.2 *Dialect Block: Syrian real /a:/*

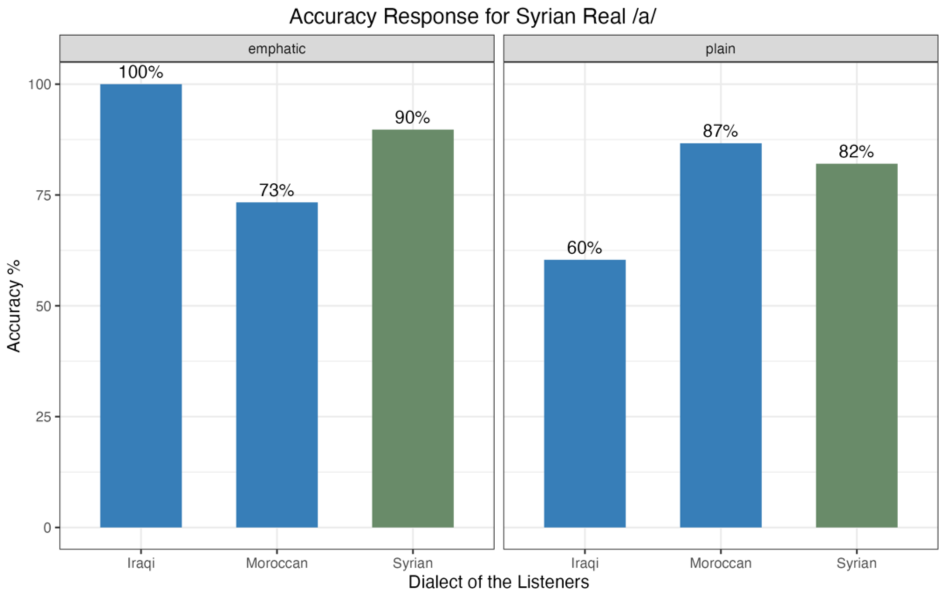


Figure 47 Percentage of the responses for each dialect listeners for the Syrian trajectories

Figure 47 shows that for the emphatic condition, the Moroccan listeners demonstrated an accuracy rate of 73%, indicating a relatively high level of accuracy in identifying the emphatic /a/. However, the Iraqi listeners achieved a perfect accuracy rate of 100% and the accuracy rate for the Syrian listeners was 90%. The pairwise results in Table 28 shows that the Syrian-Iraqi and Syrian-Moroccan listener group contrasts show no significant differences, indicating that all listeners are attending to the emphatic Syrian cues equally well.

In the plain condition, the Iraqi listeners exhibited an accuracy rate of 60%, indicating a moderate level of correct responses. In contrast, the dialects Moroccan and Syrian demonstrated higher accuracy rates of 87% and 82% respectively, suggesting a better perception of the vowel /a:/ in the plain condition for these dialects. As for the emphatic condition, the contrast between the listener groups show no significant differences.

All in all, the high accuracy for all listeners suggests that, with all available cues, all listeners were able to attend to these two conditions above chance and equally well, and thus could recognise the Syrian vowel /a:/. Figures 48 and 49 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Syrian dialect and each dialect of the listener.

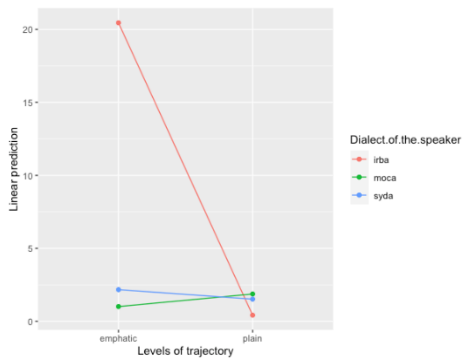


Figure 48 Linear Predicted: Syrian by each Listener Group Dialect

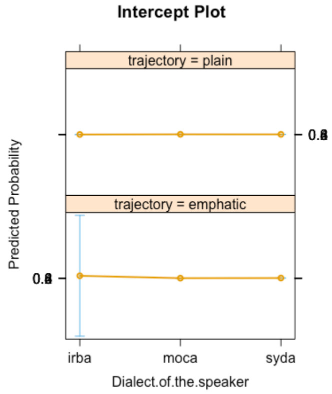


Figure 49 Predicted Probability: Syrian by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Syrian	Iraqi - Syrian	17.205	2212.410	0.008	1.0000
Emphatic	Syrian	Moroccan - Syrian	-1.157	0.787	-1.471	0.3051
Emphatic	Syrian	Iraqi - Moroccan	18.362	2212.410	0.008	1.0000
Plain	Syrian	Iraqi - Syrian	-1.099	0.503	-2.184	0.0739
Plain	Syrian	Moroccan - Syrian	0.352	0.867	0.406	0.9131
Plain	Syrian	Iraqi - Moroccan	-1.451	0.810	-1.791	0.1725

Table 28 Pairwise comparisons Results: Syrian real /a:/ stimuli by each dialect Listener Group

4.5.2.3 Dialect Block: Iraqi real /a:/

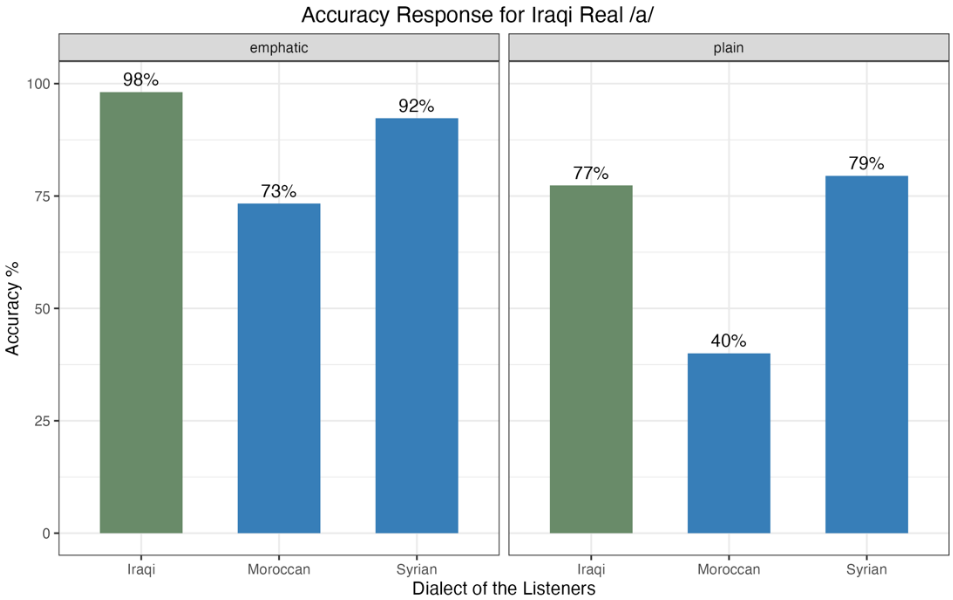


Figure 50 Percentage of the responses for each dialect listeners for the Iraqi trajectories

Figure 50 shows that for the emphatic trajectory, Iraqi listeners exhibited a notable accuracy rate of 98%. Syrian listeners follow with an equally high accuracy rate of 92%, but Moroccan listeners demonstrated a somewhat lower raw accuracy rate of 73%. This indicates that Iraqi and Syrian listeners have do better in attending to real Iraqi cues. The pairwise comparison result in Table 29 shows that the Iraqi-Moroccan listener group contrast is significant (β = 2.940, *SE* = 1.166, $p < 0.0314$), indicating that Iraqi listeners are attending to the cues better than the Moroccan listeners. However, the Iraqi-Syrian contrast is not significant (β = 1.4663, *SE* = 1.175, $p < 0.4248$) and this suggests that both dialects' listeners are attending to the cues to the same extent.

In the plain condition, Iraqi and Syrian listeners exhibited an accuracy rate of 77% and 79% respectively. However, Moroccan listeners displayed a much lower (below chance) accuracy rate of 40%. The pairwise comparison suggest the Iraqi-Moroccan listener group contrast is significant (β = 1.6341, *SE* = 0.621, $p < 0.0231$), suggesting that Moroccan listeners, with all given real cues, were unable to accurately perceive the Iraqi plain vowel and instead heard it as more like an emphatic /a:/ vowel. However, the Iraqi-Syrian listener group contrast is not significant which means that both dialects attend to the real Iraqi plain cues with a high degree of accuracy (β = -0.126, *SE* = 0.515, $p < 0.9676$). Regarding the Moroccan-Syrian listener group contrast, the pairwise comparison is significant (β = -1.760, *SE* = 0.660, *p* <

0.0208). This suggests that the Syrians do better at identifying Iraqi plain /a:/ than the Moroccan do.

These findings suggest that Iraqi and Syrian listeners were attending to the emphatic and plain cues of the Iraqi dialect with high accuracy, but Moroccan listeners recorded a very low accuracy of perceiving these conditions. Figures 51 and 52 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Iraqi dialect and each dialect of the listener.

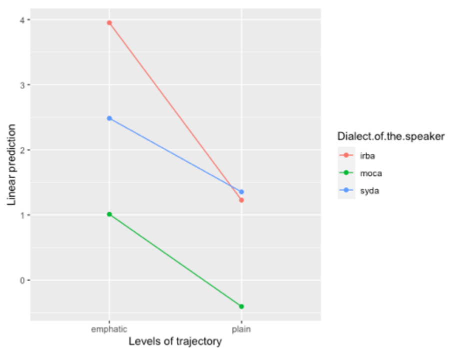


Figure 51 Linear Predicted: Iraqi by each Listener Group Dialect

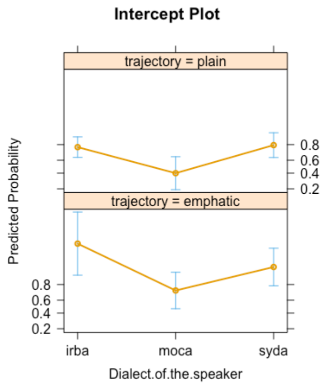


Figure 52 Predicted Probability: Iraqi by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Iraqi	Iraqi - Moroccan	2.940	1.166	2.521	0.0314
Emphatic	Iraqi	Iraqi - Syrian	1.466	1.175	1.248	0.4248
Emphatic	Iraqi	Moroccan - Syrian	-1.473	0.838	-1.758	0.1838
Plain	Iraqi	Iraqi - Moroccan	1.634	0.621	2.632	0.0231
Plain	Iraqi	Iraqi - Syrian	-0.126	0.515	-0.245	0.9676
Plain	Iraqi	Moroccan - Syrian	-1.760	0.660	-2.668	0.0208

Table 29 Pairwise comparisons Results: Iraqi real /a:/ stimuli by each dialect Listener Group

4.5.3 **Results: Real stimuli /i:/**

4.5.3.1 Dialect Block: Moroccan real /i:/

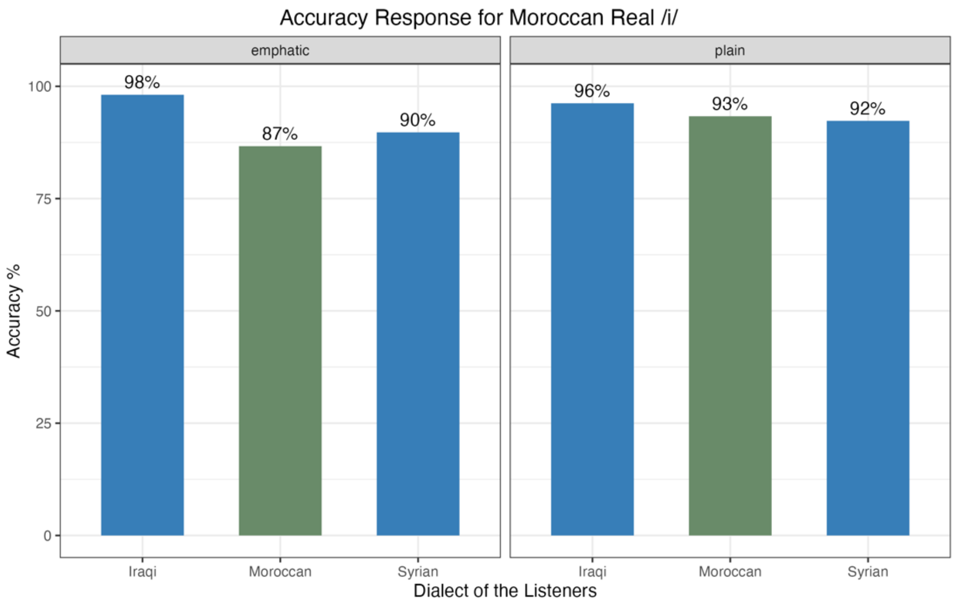


Figure 53 Percentage of the responses for each dialect listeners for the Moroccan trajectories

Figure 53 shows that for the emphatic condition, a high accuracy rate of 98% was recorded among Iraqi listeners. The Moroccan listeners exhibited a slightly lower percentage, with 87% of listeners correctly perceiving the vowel /i/. The Syrian listeners showed a comparable level of accuracy, with 90% of listeners perceiving the vowel correctly. The pairwise comparison test results in Table 30 show that when either Syrian or Iraqi listeners are contrasted with Moroccan listeners the differences are not significant, indicating that all dialects can attend to the cues of the Moroccan /i/ equally well and there are no challenges facing listeners in identifying the Moroccan cues of an emphatic /i/ vowel.

In the plain condition, the accuracy rate of responses remained consistently high among the different dialects. The Iraqi listeners also exhibited a high accuracy rate of 96%. The Moroccan listeners recorded a slightly lower but still substantial accuracy rate of 93%. The Syrian listeners maintained a similar level of accuracy, with 92% of listeners perceiving the vowel correctly. Also, the pairwise comparison suggests that the contrast between Moroccan and Iraqi listeners shows no significant difference which is the same with the Moroccan-Syrian contrast. This means that, like the emphatic condition, all listeners can attend to the plain condition /i:/ in real Moroccan stimuli with no difficulties.

These findings suggest that all dialects can recognise the Moroccan cues to both conditions with high accuracy. This can clearly be noted from the pairwise test; none of the contrast differences among Moroccan and Syrian and Iraqi reached significance at the specified significance level (e.g., $\alpha = 0.05$). Therefore, all dialects are perceiving the Moroccan target vowel accurately. Figures 54 and 55 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Moroccan dialect and each dialect of the listener.

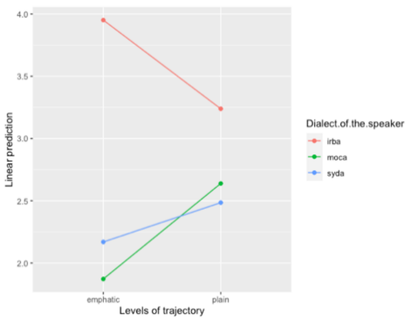


Figure 54 Linear Predicted: Moroccan by each Listener Group Dialect

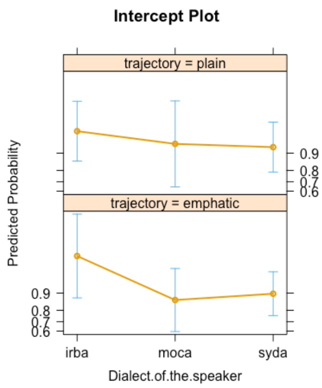


Figure 55 Predicted Probability: Moroccan by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Moroccan	Iraqi - Moroccan	2.496	1.513	1.650	0.2247
Emphatic	Moroccan	Moroccan - Syrian	-0.339	1.183	-0.286	0.9558
Emphatic	Moroccan	Iraqi - Syrian	2.157	1.314	1.642	0.2279
Plain	Moroccan	Iraqi - Moroccan	0.870	1.450	0.600	0.8203
Plain	Moroccan	Moroccan - Syrian	0.115	1.383	0.083	0.9962
Plain	Moroccan	Iraqi - Syrian	0.985	1.085	0.907	0.6357

Table 30 Pairwise comparisons Results: Moroccan real /i:/ stimuli by each dialect Listener Group

4.5.3.2 Dialect Block: Syrian real /i:/

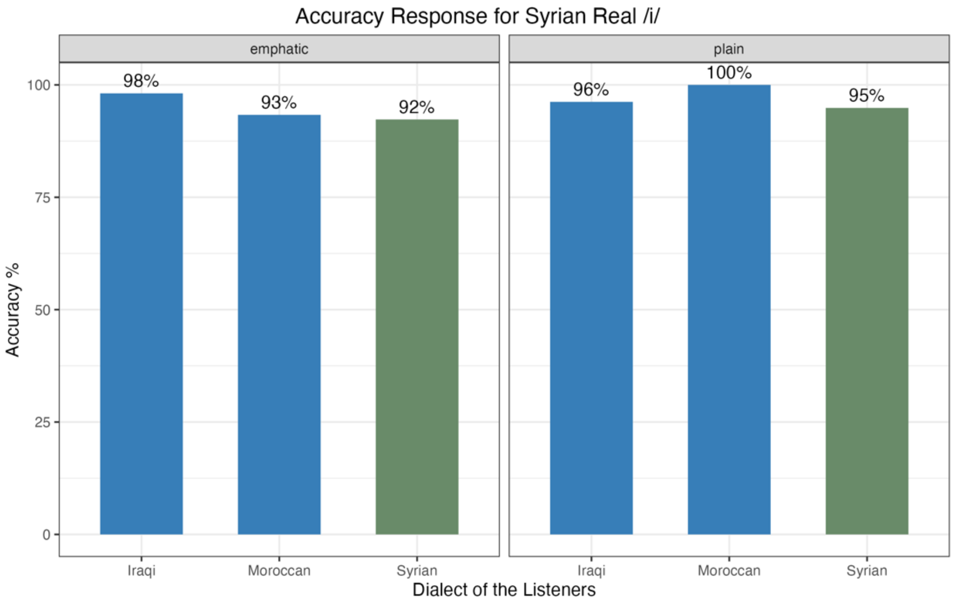


Figure 56 Percentage of the responses for each dialect listeners for the Syrian trajectories

Figure 56 shows that in emphatic condition, the Iraqi listeners demonstrate exceptional performance, with 98% accuracy in perceiving the Syrian vowel /i/. The Moroccan listeners exhibit a slightly lower accuracy rate of 93%. The Syrian listeners follow closely with 92% accuracy.

As for the plain condition, Moroccan listeners achieve a perfect score of 100% accuracy in perceiving the vowel /i/ in this condition. The Iraqi listeners maintain a high level of accuracy at 96%and the Syrian listeners show 95% accuracy. From Table 31 below it can be seen that there is no significant difference between any of the dialect listener groups, confirming that all dialects have robust perception performance in response to real Syrian stimuli.

In summary, the high accuracy of correct responses for all dialects reflects a robust perception of Syrian cues to the emphatic and plain conditions. Figures 57 and 58 show the Linear predicted mean and the intercept plot for the interaction between the Syrian listeners and each dialect of the speaker.

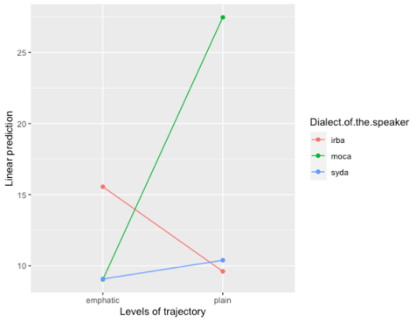


Figure 57 Linear Predicted: Syrian by each Listener Group Dialect

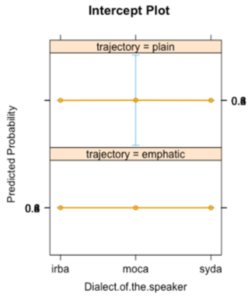


Figure 58 Predicted Probability: Syrian by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Syrian	Iraqi - Syrian	1.620	1.300	1.246	0.4262
Emphatic	Syrian	Moroccan - Syrian	0.148	1.432	0.103	0.9941
Emphatic	Syrian	Iraqi - Moroccan	1.472	1.643	0.896	0.6429
Plain	Syrian	Iraqi - Syrian	0.363	1.151	0.316	0.9465
Plain	Syrian	Moroccan - Syrian	3.994	7.799	0.512	0.8654
Plain	Syrian	Iraqi - Moroccan	-3.631	7.796	-0.466	0.8873

Table 31 Pairwise comparisons Results: Syrian real /i:/ stimuli by each dialect Listener Group

4.5.3.3 *Dialect Block: Iraqi real /i:/*

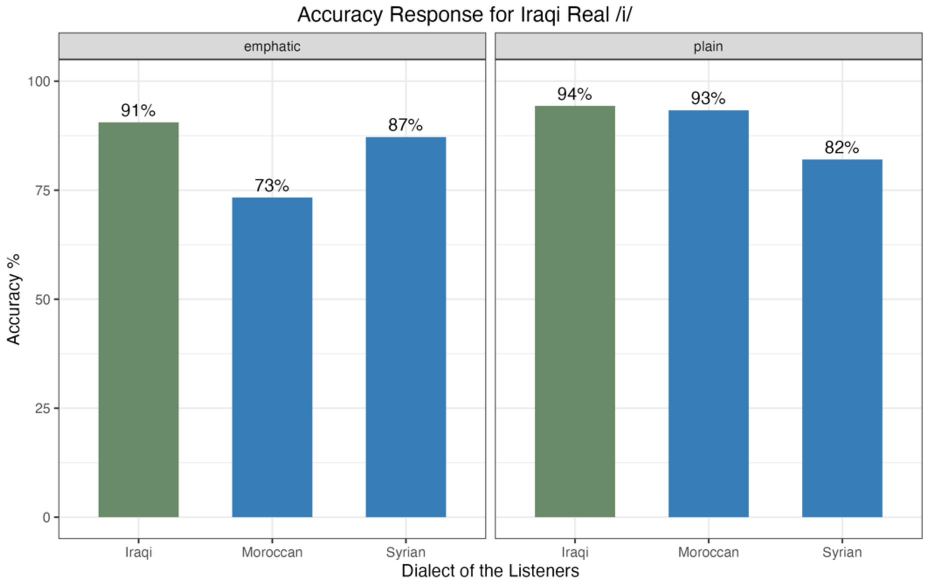


Figure 59 Percentage of the responses for each dialect listeners for the Iraqi trajectories

Figure 59 shows that in emphatic condition, the Iraqi listeners achieved a percentage of 91% correct responses. Moroccan listeners achieved a slightly lower percentage of 73% correct responses, and Syrian listeners achieved a percentage of 87% correct responses. The Iraqi-Moroccan listener group contrast appears different but the difference is not significant (β = 1.724, *SE* = 0.999, $p < 0.1951$). The Iraqi - Syrian listener group contrast is also not significant (β = 0.396, *SE* = 0.826, $p < 0.8809$), suggesting that both dialects, Syrian and Moroccan can attend to the given vowel cues in the Iraqi emphatic condition as well as the Iraqi listeners.

Moving on to the plain condition, the Iraqi listeners displayed 94% correct responses, Moroccan listeners 93% correct responses and Syrian listeners achieved 82% correct responses. The pairwise comparisons result in Table 32 show that the Iraqi-Moroccan listener group contrast is not significant (β 0.205, *SE* = 1.431, $p < 0.9888$). Furthermore, the contrast Iraqi-Syrian listener group contrast is not significant difference either (β 1.574, *SE* = 0.898, $p < 0.1858$).

The findings suggest that all listener groups are accurately attending to real cues in both condition types in Iraqi dialect with high accuracy. Figures 60 and 61 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Iraqi dialect and each dialects of the listeners.

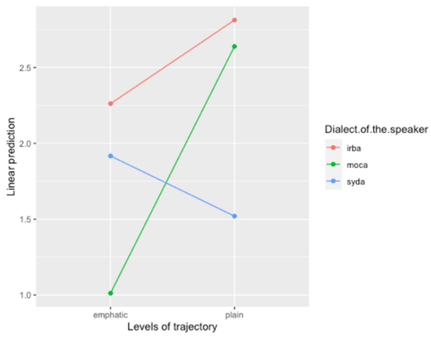


Figure 60 Linear Predicted: Iraqi by each Listener Group Dialect

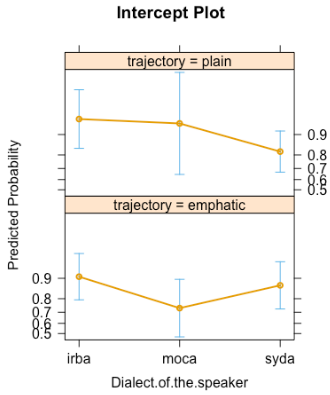


Figure 61 Predicted Probability: Iraqi by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Iraqi	Iraqi - Moroccan	1.724	0.999	1.727	0.1951
Emphatic	Iraqi	Iraqi - Syrian	0.396	0.826	0.480	0.8809
Emphatic	Iraqi	Moroccan-Syrian	-1.328	1.015	-1.309	0.3902
Plain	Iraqi	Iraqi - Moroccan	0.205	1.431	0.143	0.9888
Plain	Iraqi	Iraqi - Syrian	1.574	0.898	1.753	0.1858
Plain	Iraqi	Moroccan-Syrian	1.370	1.367	1.002	0.5757

Table 32 Pairwise comparisons Results: Iraqi real stimuli /i:/ by each dialect Listener Group

4.5.3.4 Summary and Interim Discussion (Real Stimuli)

From the results above we can conclude that it is fairly obvious that all listeners from all tested dialects can identify everyone else's cues of the vowel /i:/ for both conditions, plain and emphatic. However, the only issues that listeners faced is when listening to the cues of the vowel /a:/ for the plain condition, where they found it difficult to identify it. The difference between two vowels was seen to be significant according to LMER model, (β

1.708444, SE = 0.535970, $p < 0.001435$), see Figure 62 below. The only significant result when listening to real emphatic /a:/ is that Moroccan listeners do somewhat less well than Iraqi listeners in identifying an Iraqi emphatic /a:/. The Moroccans are still performing well above chance (73% accuracy) - but they find it harder than the native Iraqi listeners (and Syrian listeners don't find it harder than Iraqis). Thus, the main finding from this section is that everyone can identify a target word containing an emphatic when presented with real stimuli. The issue of why listeners so often misidentify a plain /a:/ as emphatic, across dialects, in real stimuli, is interesting and it may indicate there is something about variation across dialects in the size and shape of the overall vowel triangle.

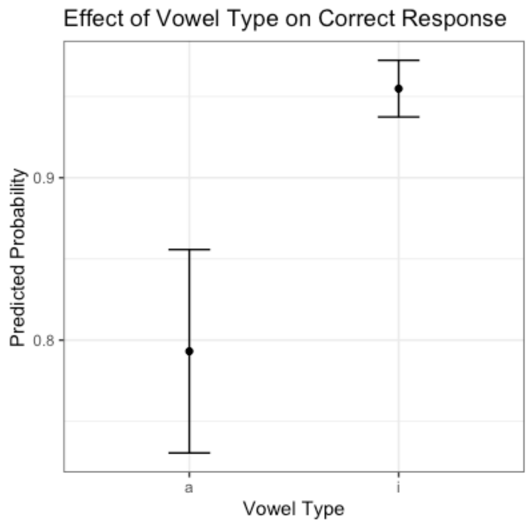


Figure 62 Predicted probability for Vowel type for real stimuli

The prediction was that listeners of such a dialect are better in attending to their own vowels and can easily identify the cues of their own trajectory types. Listeners were all proficient with identifying emphatics; the differences lie only in the relative accuracy of identifying the plain /a:/. This suggests that there is something related to the overall shape and size of the vowel trajectory across dialects, rather than cues to emphasis. Because there was no consistency in the results this prediction was rejected.

4.5.4 Results of Manipulated stimuli

4.5.5 Results: Manipulated stimuli /a:/

4.5.5.1 Dialect Block: Moroccan manipulated /a:/

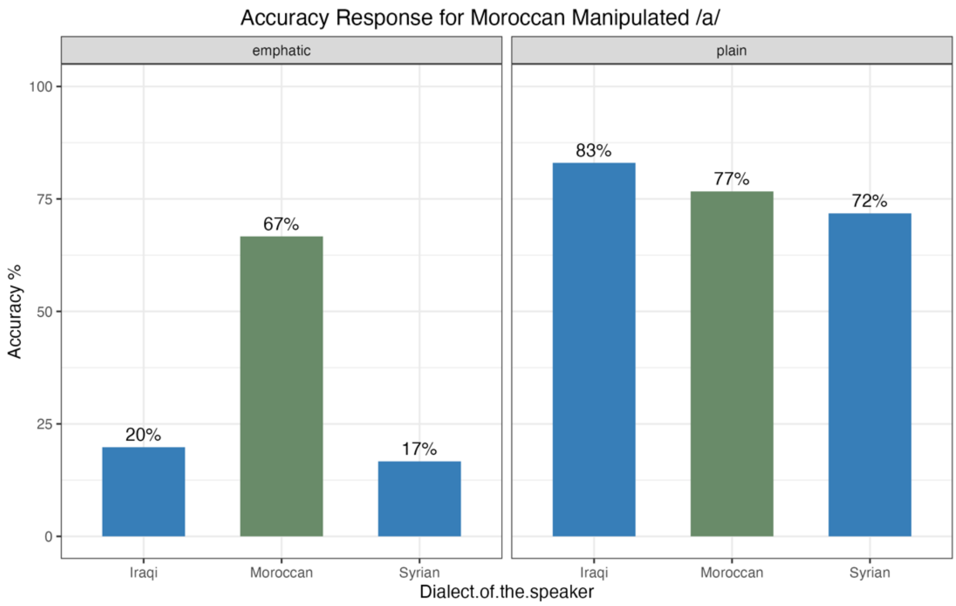


Figure 63 Percentage of the responses for each dialect listeners for the Moroccan trajectories

Figure 63 shows that for the Moroccan-like emphatic condition, the highest percentage of correct responses is observed for the listeners of Moroccan dialect with 67% accuracy. Nevertheless, Syrian and Iraqi listeners exhibit lower percentages of correct responses, with 17% and 20% respectively, indicating potential challenges in accurately perceiving the Moroccan cues of the emphatic condition, unlike Moroccan listeners. Table 33 below suggests that for the emphatic condition, Moroccan listeners are significantly different than both Iraqi listeners (β -2.091, *SE* = 0.458, *p* < .0001), and Syrian listeners (β 2.303, *SE* = 0.492, *p* < .0001), respectively. This indicates that Iraqi and Syrian listeners, with only F2 cues presented, cannot identify the emphatic Moroccan vowel /a:/.

In the Moroccan plain condition, the results display higher levels of accuracy overall. Iraqi listeners demonstrate the highest accuracy, with 83% of the responses correct. Syrian listeners also show a notably high percentage of correct responses, with 72% accuracy. In contrast, Moroccan dialect listeners, who are native to the dialect, demonstrated 77%

accuracy for the plain condition. The pairwise comparison in Table 28 show that in plain condition, the contrasts between the Iraqi-Moroccan (β 0.397, *SE* = 0.503, $p < 0.8592$) and Moroccan-Syrian listener groups (β 2.303, *SE* = 0.5, $p < 0.9565$) are not significant. Thus, Moroccan listeners are the only ones who can identify their Moroccan-like emphatic F2 cues well but perform equally as well as all dialect listeners in recognising the plain condition. Figures 64 and 65 below show the Linear predicted mean and the intercept plot.

Overall, for the listeners who are hearing their own dialect (Moroccan) in the emphatic condition, they recorded a high accuracy while other dialect listeners were not able to accurately identify this condition. However, in the plain condition, all listeners performed equally well and found no difficulties identifying the plain vowel.

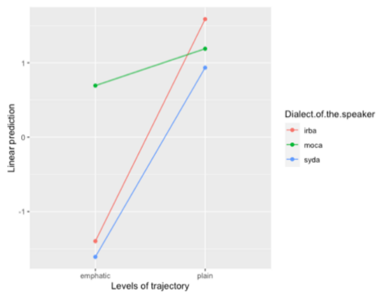


Figure 65 Linear Predicted: Moroccan by each speaker's Dialect`

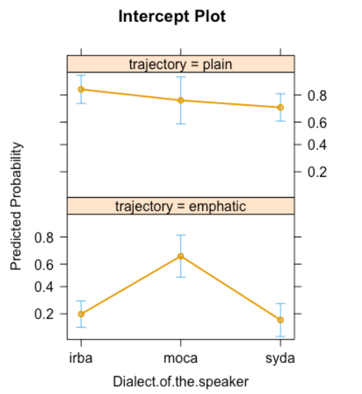


Figure 64 Predicted Probability: Moroccan by each speaker's Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Moroccan	Iraqi- Moroccan	-2.091	0.458	-4.57	<.0001
Emphatic	Moroccan	Moroccan - Syrian	2.303	0.492	4.678	<.0001
Emphatic	Moroccan	Iraqi - Syrian	0.211	0.389	0.543	0.8503
Plain	Moroccan	Iraqi - Moroccan	0.397	0.503	0.79	0.8592
Plain	Moroccan	Moroccan - Syrian	0.255	0.5	0.511	0.9565
Plain	Moroccan	Iraqi - Syrian	0.653	0.361	1.809	0.1667

Table 33 Pairwise comparisons Results: Moroccan manipulated /a:/ by each dialect Listener Group

4.5.5.2 Dialect Block: Syrian manipulated /a:/

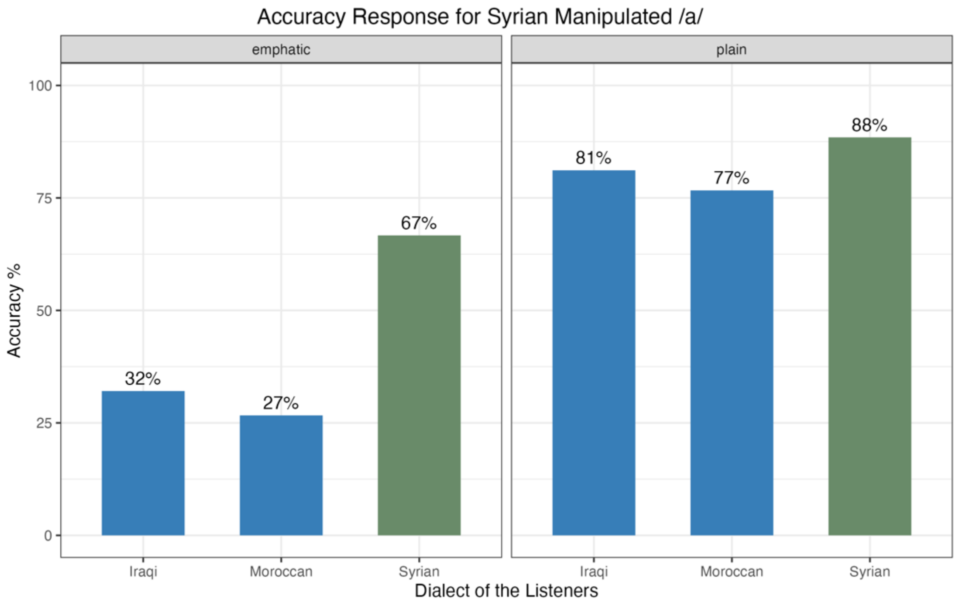


Figure 66 Percentage of the responses for each dialect listeners for the Syrian trajectories

Figure 66 shows that for the Syrian-like emphatic F2 trajectory condition, the highest percentage of correct responses is observed for the listeners of Syrian dialect with 67% accuracy. However, Iraqi and Moroccan listeners display lower percentages of correct responses, with 32% and 27% respectively, indicating potential challenges in accurately perceiving the Syrian emphatic condition F2 trajectory cues, unlike Syrian listeners. The pairwise comparison results in Table 34 confirm that only Syrian listeners can recognise their own F2 trajectory cues for the emphatic vowel and other dialect listeners, Moroccan and Iraqi, are unable to attend to these cues; the contrast between Iraqi and Syrian listeners is significant (β -1.443, $SE = 0.318$, $p < 0.0001$)Similarly, the contrast between Moroccan and Syrian is also significant (β -1.705, $SE = 0.478$, $p < 0.002$).

Moving on to the Syrian plain condition, the results show higher levels of accuracy overall. Syrian listeners demonstrate the highest accuracy, with 88% of the responses correct. Iraqi listeners also show a notably high percentage of correct responses, with 81% accuracy. On the other hand, Moroccan listeners achieve 77% accuracy, suggesting a moderate to high level of proficiency in perceiving the Syrian plain condition. The pairwise comparison in Table 34 show that the contrasts between Syrian and Iraqi (β -0.578, $SE = 0.433$, $p < 0.5397$) and between Syrian and Moroccan (β -0.847, $SE = 0.559$, $p < 0.4272$) do not show any significant differences.

In summary, only Syrian listeners can reliably identify their own emphatic F2 trajectory cues while other dialect listeners demonstrate lower accuracy for the Syrian emphatic condition; they display higher levels of accuracy in the plain condition, suggesting that listeners of Iraqi and Moroccan hear the manipulated 'Syrian-like' emphatic F2 trajectory as plain, but they do attend to the presented 'plain' F2 trajectory cues mostly accurately. Figures 67 and 68 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Syrian dialect and each dialect of the listeners.

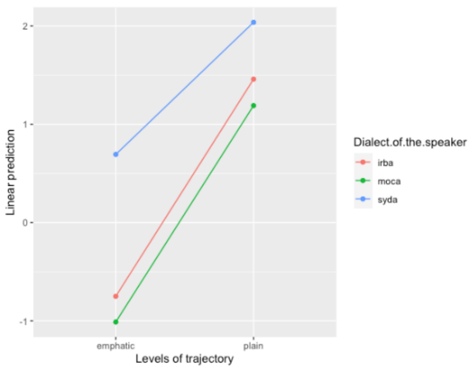


Figure 68 Linear Predicted: Syrian by each speaker's Listener Group

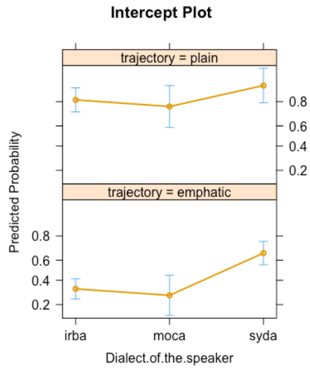


Figure 67 Predicted Probability: Syrian by each speaker's Listener Group

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Syrian	Iraqi - Syrian	-1.443	0.318	-4.542	<.0001
Emphatic	Syrian	Moroccan - Syrian	-1.705	0.478	-3.569	0.002
Emphatic	Syrian	Iraqi - Moroccan	0.261	0.462	0.565	0.8386
Plain	Syrian	Iraqi - Syrian	-0.578	0.433	-1.336	0.5397
Plain	Syrian	Moroccan - Syrian	-0.847	0.559	-1.517	0.4272
Plain	Syrian	Iraqi - Moroccan	0.269	0.498	0.540	0.8514

Table 34 Pairwise comparisons Results: Syrian manipulated /a:/ stimuli by each dialect Listener Group

4.5.5.3 Dialect Block: Iraqi manipulated /a:/

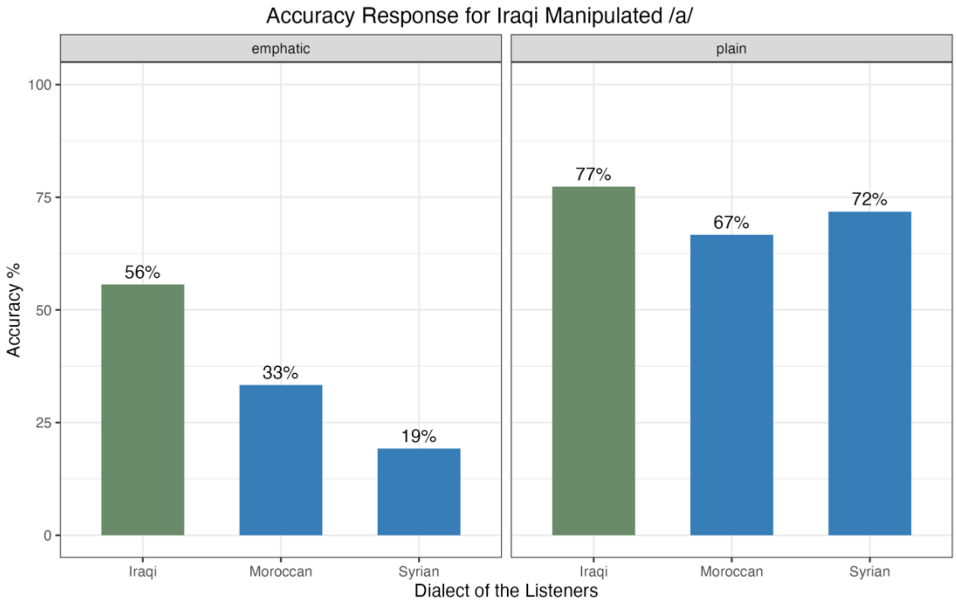


Figure 69 Percentage of the responses for each dialect listeners for the Iraqi trajectories

In the Iraqi-like emphatic manipulated F2 trajectory condition, Iraqi listeners exhibit the highest percentage of correct responses at 56%, but it is notable that this is only just above chance. On the other hand, Moroccan listeners show a lower accuracy rate still of 33%, while the Syrian listeners have the lowest percentage of correct responses at 19%, suggesting potential challenges in perceiving the Iraqi emphatic F2 trajectory for both listener groups. When looking at the pairwise comparison results in Table 35, it can be seen that the contrast difference between Iraqi and Moroccan approaches but is not significant (β 0.921, SE = 0.434, $p < 0.0854$) indicating that the low performance of the Moroccan listeners is in fact no worse than performance of the Iraqi listeners. The contrast between the Iraqi and Syrian listener groups is significant (β 1.662, SE = 0.348, $p < 0.0001$) which means that Syrian listeners perform worse at perceiving the emphatic cues for the Iraqi-like F2 trajectory of an /a/ vowel, with only Iraqi listeners achieving an above chance performance when attending to these cues.

In plain condition, the Iraqi listeners show high accuracy with 77% correct responses; the Syrian listeners follow with 72%, suggesting a relatively better ability to perceive the plain F2 trajectory compared to the emphatic one. Moroccan listeners show 67% accuracy, demonstrating a moderate level of performance. The pairwise results suggest that the contrast between Iraqi and Moroccan listeners is not significant (β 0.536, SE = 0.452, $p < 0.6357$), and nor is the contrast between Iraqi and Syrian listeners (β 0.294, *SE* = 0.342, *p* < 0.8255), indicating that both dialect listeners, Moroccan and Syrian are able to attend to the plain condition F2 trajectory equally well as Iraqi listeners.

In summary, these findings suggest that Iraqi listeners stand out as the most proficient in both conditions but with significant different contrasts with the other listener groups in response to the emphatic Iraqi-like F2 trajectory. Syrian listeners show a notable difference in performance, with much better accuracy in response to the plain condition compared to the emphatic one. Unlike the Moroccan listeners, who perform above chance in identifying the plain condition but below chance in the emphatic condition. Figures 70 and 71 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Iraqi dialect and each dialect of the speaker.

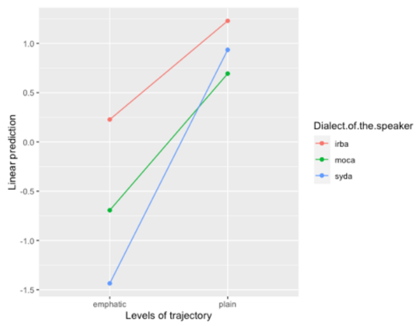


Figure 70 Linear Predicted: Iraqi by each speaker's Listener Group

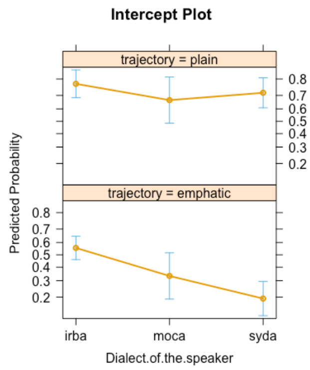


Figure 71 Predicted Probability: Iraqi by each speaker's Listener Group

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Iraqi	Iraqi - Moroccan	0.921	0.434	2.122	0.0854
Emphatic	Iraqi	Iraqi - Syrian	1.662	0.348	4.783	<.0001
Emphatic	Iraqi	Moroccan - Syrian	0.742	0.482	1.539	0.2728
Plain	Iraqi	Iraqi - Moroccan	0.536	0.452	1.186	0.4615
Plain	Iraqi	Iraqi - Syrian	0.294	0.342	0.86	0.6656
Plain	Iraqi	Moroccan - Syrian	-0.241	0.462	-0.522	0.8605

Table 35 Pairwise comparisons Results: Iraqi manipulated /a:/ stimuli by each dialect Listener Group

4.5.6 **Results: Manipulated stimuli /i:/**

4.5.6.1 Dialect Block: Moroccan manipulated /i:/

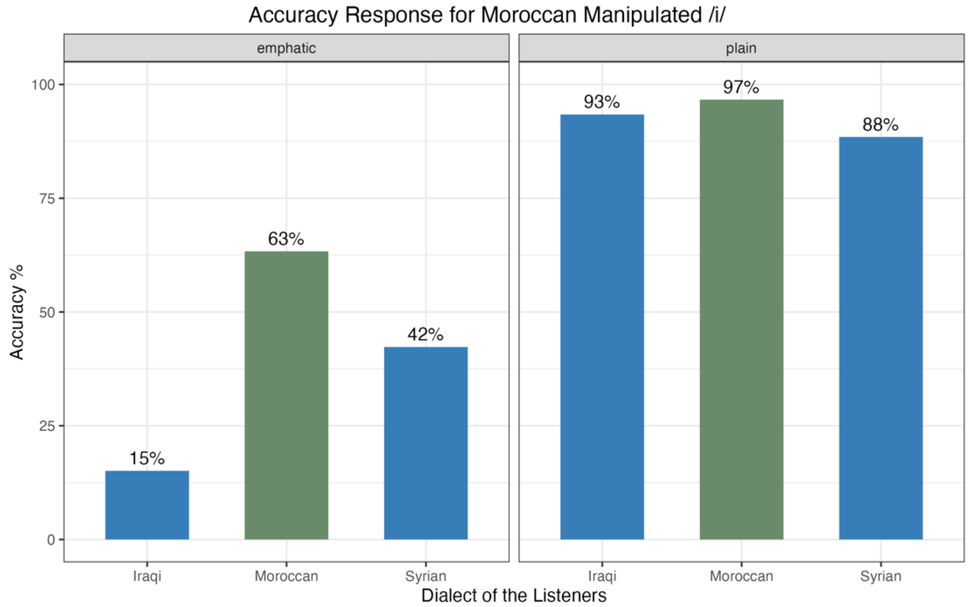


Figure 72 Percentage of the responses for each dialect listeners for the Moroccan trajectories

Figure 72 shows that for the emphatic condition, listeners of the dialect Iraqi exhibited a very low accuracy rate of 15%. Also, the dialect Syrian achieved an accuracy rate just below chance of 42%. On the other hand, Moroccan listeners displayed a somewhat higher accuracy rate of 63%. These findings suggest that listeners of the Moroccan dialect have better perception of the Moroccan emphatic F2 trajectory cues and compared to Iraqi and Syrian listeners. The pairwise comparison results in Table 36 show that the contrasts between the Moroccan and Iraqi groups (β -2.2120, SE = 0.334, p < .0001) and Moroccan and Syrian groups (β 1.5202, SE = 0.331, p < 0.0001) both significantly differ in discriminatory accuracy of the Moroccan-like emphatic F2 trajectory condition for the vowel /i/. It is also interesting to find that the contrast between Iraqi and Syrian is significantly different with Syrian better able to recognise the emphatic (β -1.4583, SE = 0.371, p < 0.0002). This suggests that Moroccan listeners perform better than all other groups at attending to the emphatic cues of their native condition, followed by Syrian listeners who perform just below chance at perceiving this condition, but who are significantly better than the Iraqi listeners in accurately perceiving the Moroccan emphatic vowel /i/.

Regarding the plain condition, all dialects showed relatively high accuracy rates. Iraqi listeners maintained a high accuracy rate of 93%. Moroccan listeners exhibited the highest accuracy rate of 97%. Syrian listeners achieved an accuracy rate of 88%. Looking at the pairwise results, it can be seen that the contrast between Moroccan and Iraqi listeners is not significant (β -0.7135, *SE* = 1.094, $p < 0.7912$). Similarly, the contrast between Moroccan and Syrian listeners is not significant (β 1.3354, *SE* = 1.082, $p < 0.4333$), suggesting that all dialect listeners can attend to the plain condition accurately.

In summary, the Moroccan listeners demonstrate the highest accuracy rate perceiving their own emphatic F2 trajectory condition significantly better than other listener groups, while the dialects Syrian and Iraqi are able to recognise the plain condition but not the emphatic one. Also, there was a significant contrast between Iraqi and Syrian in perceiving the emphatic condition. Figures 73 and 74 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Moroccan dialect and each dialect of the listeners.

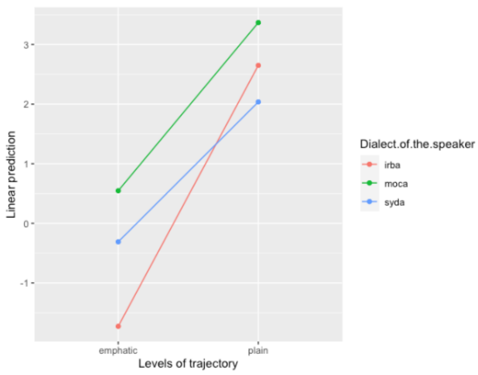


Figure 73 Linear Predicted: Moroccan by each Listener Group Dialect

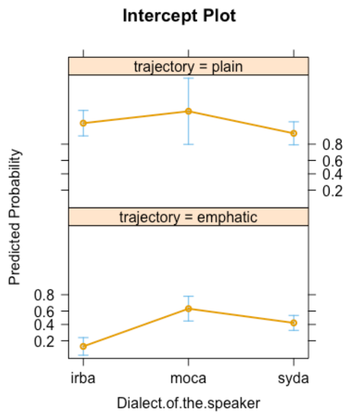


Figure 74 Predicted Probability: Moroccan by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Moroccan	Iraqi - Moroccan	-2.2120	0.334	-4.798	<.0001
Emphatic	Moroccan	Moroccan - Syrian	1.5202	0.331	1.908	0.0001
Emphatic	Moroccan	Iraqi - Syrian	-1.4583	0.371	-3.934	0.0002
Plain	Moroccan	Iraqi - Moroccan	-0.7135	1.094	-0.652	0.7912
Plain	Moroccan	Moroccan - Syrian	1.3354	1.082	1.234	0.4333
Plain	Moroccan	Iraqi - Syrian	0.6219	0.537	1.157	0.4790

Table 36 Pairwise comparisons Results: Moroccan manipulated /i:/ stimuli by each dialect Listener Group

4.5.6.2 Dialect Block: Syrian manipulated /i:/

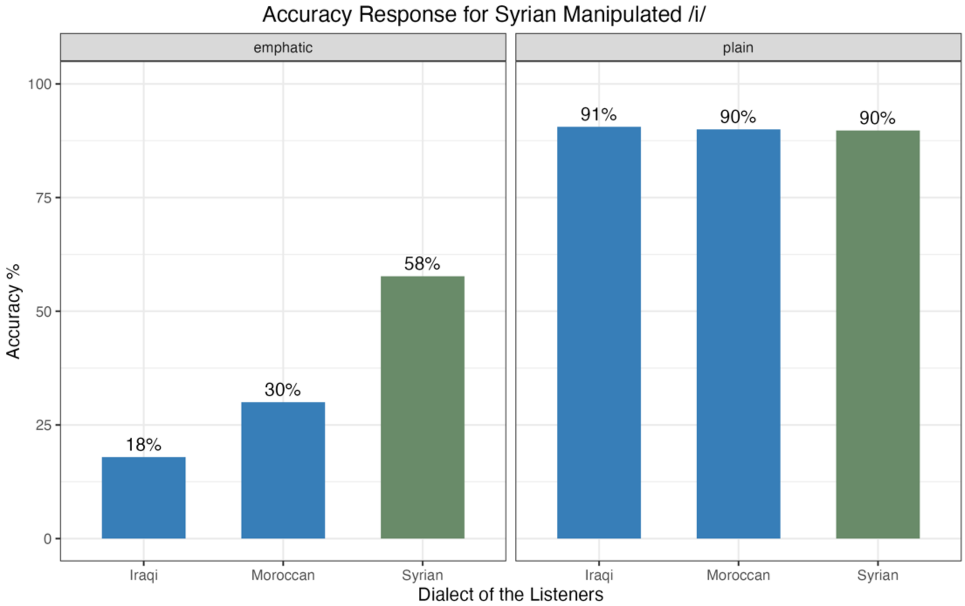


Figure 75 Percentage of the responses for each dialect listeners for the Syrian trajectories

Figure 75 above presents the percentages of correct responses for all dialects with regards to the Syrian emphatic and plain trajectories of the vowel /i/. Regarding the emphatic F2 trajectory condition, Iraqi listeners showed the lowest accuracy of 18%, indicating consistent misidentification of the cues of the Syrian emphatic F2 trajectory condition. Moroccan listeners performed better but still well below chance, with only 30% of responses correct. Syrian dialect listeners demonstrated the highest accuracy, with 58% of correct responses in response to the emphatic F2 trajectory condition. Table 37 shows that the Syrian-Iraqi contrast is not significant (β -1.9002, $SE = 0.359$, $p < 0.0001$), suggesting that Iraqi listeners find it more difficult to identify the Syrian emphatic cues. Similarly, the contrast between Syrian and Moroccan indicates that Moroccan also listeners perform worse than the Syrian listeners (β -1.4531, $SE = 0.339$, $p < 0.0003$). The overall analysis suggests that only Syrian listeners can attend to these emphatic F2 trajectory cues.

Turning to the plain condition, all dialects showed high accuracy, indicating good overall perception of the plain condition of the vowel /i/ in the manipulated stimuli. Iraqi and Moroccan listeners both exhibited 91% correct and Syrian listeners displayed 90%. These results suggest that all speakers can perceive and differentiate the presented plain F2 trajectory condition of the vowel /i/. The pairwise comparison results in Table 32 confirm that all listeners from Iraqi and Moroccan show no significant contrast with the Syrian listeners.

Overall, while the Syrian-like emphatic F2 trajectory condition poses challenges for Iraqi and Moroccan listeners, whereas Syrian listeners exhibit high accuracy in perceiving their own emphatic condition. However, all dialect listeners attend to the plain condition with high accuracy. Figures 76 and 77 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Syrian dialect and each dialect of the listener.

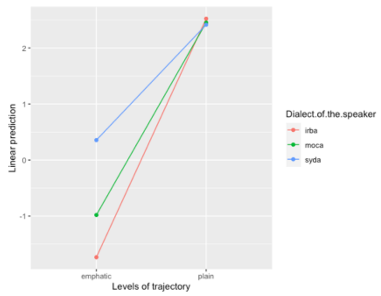


Figure 76 Linear Predicted: Syrian by each Listener Group Dialect

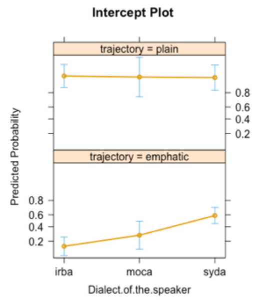


Figure 77 Predicted Probability: Syrian by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Syrian	Iraqi - Syrian	-1.6258	0.235	-5.275	<.0001
Emphatic	Syrian	Moroccan - Syrian	-1.4531	0.339	-2.484	0.0003
Emphatic	Syrian	Iraqi - Moroccan	-0.6921	0.493	-1.403	0.3394
Plain	Syrian	Iraqi - Syrian	0.0936	0.510	0.184	0.9816
Plain	Syrian	Moroccan - Syrian	0.0288	0.729	0.039	0.9991
Plain	Syrian	Iraqi - Moroccan	0.0649	0.708	0.092	0.9954

Table 37 Pairwise comparisons Results: Syrian manipulated /i:/ stimuli by each dialect Listener Group

4.5.6.3 Dialect Block: Iraqi manipulated /i:/

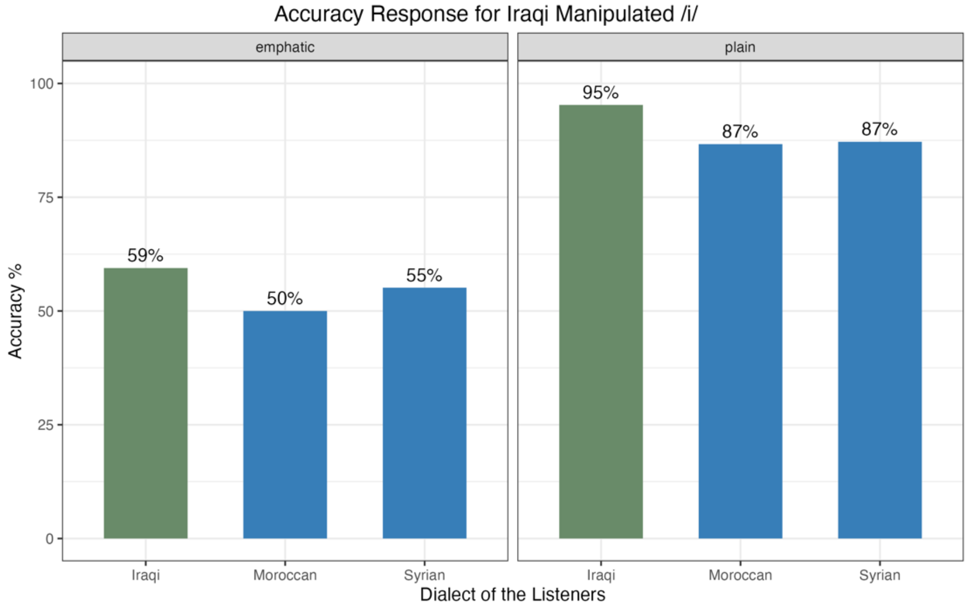


Figure 78 Percentage of the responses for each dialect listeners for the Iraqi trajectories

Figure 78 above presents the percentage of correct responses for each dialect listener group regarding the Iraqi-like F2 trajectory in emphatic and plain conditions of the vowel /i/.

For the emphatic condition, the accuracy rates were as follows: Iraqi listeners dialect achieved an accuracy rate of 59%, which is a better accuracy compared to the Moroccan and Syrian listeners who exhibited 50% and 55% accuracy, respectively, which is closer to chance. According to the pairwise analysis in Table 38, the contrast between Iraqi and Moroccan (β 0.4006, SE = 0.441, $p < 0.8007$) and between Iraqi and Syrian (β 0.1863, SE = 0.32, $p < 0.9375$) show no significant differences, and. This suggest that all listener groups stand an equally poor chance of accurately perceiving the 'Iraqi-like' emphatic F2 trajectory cues.

Turning to the Plain condition, the accuracy rates were as follows: Iraqi listeners demonstrated an accuracy rate of 95%, while Moroccan and Syrian dialect listeners exhibited 87% accuracy; although the Iraqi listeners appear to be better able to attend to their own cues the contrasts with the Moroccan and Syrian groups were not significant, as shown in Table 38 below.

In summary, the results indicate that listeners of all dialects can attend equally well to the Iraqi-like F2 trajectory cues in both conditions but with much lower accuracy in response to the emphatic F2 trajectory. Figures 79 and 80 show the Linear predicted mean and the intercept plot for the interaction between the trajectory of the Iraqi dialect and each dialect of the listener.

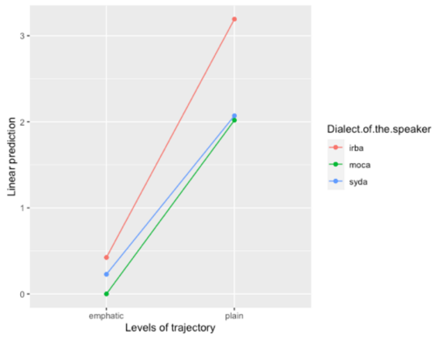


Figure 79 Linear Predicted: Iraqi by each Listener Group Dialect

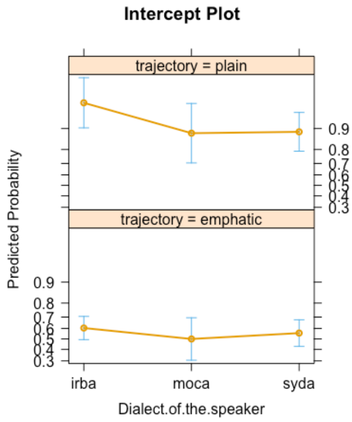


Figure 80 Predicted Probability: Iraqi by each Listener Group Dialect

Trajectory	Dialect	Contrast	Estimate	SE	z-ratio	p-value
Emphatic	Iraqi	Iraqi - Moroccan	0.3901	0.439	0.889	0.6473
Emphatic	Iraqi	Iraqi - Syrian	0.1807	0.318	0.568	0.8373
Emphatic	Iraqi	Moroccan - Syrian	-0.2093	0.455	-0.460	0.8897
Plain	Iraqi	Iraqi - Moroccan	1.1501	0.720	1.598	0.2465
Plain	Iraqi	Iraqi - Syrian	1.1051	0.579	1.909	0.1360
Plain	Iraqi	Moroccan - Syrian	-0.0449	0.651	-0.069	0.9974

Table 38 Pairwise comparisons Results: Iraqi listeners by each dialect Listener Group

4.5.6.4 Summary and Interim Discussion (Manipulated Stimuli)

The general conclusion from the above results indicates that all dialects are better at attending to their own emphatic F2 trajectory cues. This effect of listener dialect can only be observed in the emphatic condition where native listeners of a dialect can attend to emphatic F2 cues manipulated to resemble those of their own dialect significantly better than other dialect listener groups (except for Iraqi-like emphatic /i:/). For the plain stimulus, listeners are all roughly similar at perceiving the absence of emphatic cues and this is because that the plain stimuli are the same as the base one. But when an emphatic cue is present, and if it is not the one they are used to in their own dialects, they are not very good at identifying it, and they will often hear it as plain instead. This may be attributed to how much exposure listeners have to cues other than those in their own dialect.

A further GLMER model[8] shows that the Exposure factor, when interacted with dialect block and listeners' dialect, does play a significant role in predicting listeners' accuracy in identifying their own native trajectory (β 0.026723, *SE* = 0.001234, p <2e-16). It was also observed that the other factors, Attitude (β -0.001491, *SE* = 0.007996, $p < 0.852$) and Confidence (β -0.001604, *SE* = 0.008245, $p < 0.846$), did not have a significant effect when interacted with other predictors. As for the real stimuli, the GLMER model shows that all listeners find the vowel type /i/ easier to recognise in that there is a main effect of vowel type (β 0.58815, *SE* = 0.15451, p 0.0001), see Figure 81 below.

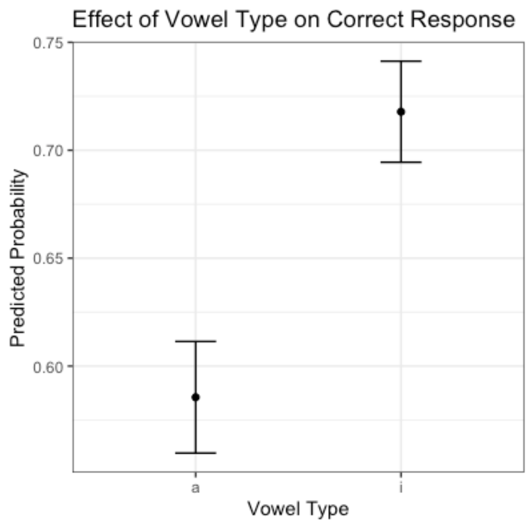


Figure 81 Predicted probability for Vowel type for manipulated stimuli

The research question presented was, 'Are listeners more 'accurate' at identifying emphatic words when given the cues to emphasis in their own dialect as opposed to other dialects, especially when presented with stimuli where only the second formant (F2) is manipulated?' Also, the hypothesis formulated posited that listeners from each dialect group would better identify the emphatic contrast when presented with F2 trajectory cues from their own dialect rather than those from another dialect. This hypothesis was accepted, as the overall results supported this conclusion, with the only exception of Iraqi listeners in the context of the vowel /i:/.

[8] Model factor <- glmer (Response ~ Dialect Block* Listeners' Dialect * Exposure +(1 | item) + (1 |subject), family = binomial, data = data)

4.5.7 **Binomial test Results for Real and Manipulated Stimuli**

This section utilized the binomial test, a statistical method used to assess whether the observed frequency of successes in a binary outcome differs significantly from a specified probability, often set at chance level (e.g., 50%). This test evaluates the likelihood that observed differences are meaningful rather than due to random variation. In this study, the binomial test was employed to determine whether listeners' response accuracy for different F2 cues of the vowels /a/ and /i/ in the two conditions—emphatic and plain stimuli—was significantly different from chance level. This approach quantified the reliability of their perceptual judgments. The analysis was conducted using the binom.test() function in R (R Core Team, 2024. By applying this test, we aimed to identify and interpret significant deviations from chance performance across the experimental conditions. Figure 82 below presents the results of the binomial test for all tested stimuli, including both real and manipulated forms; It categorizes accuracy based on p-values into "Significant ($p < 0.05$)", "Not Significant ($p \geq 0.05$)", and "Approaching Significance (p close to 0.05)".

4.5.7.1 Binomial test Results for Real Stimuli

The recognition patterns for real stimuli, as illustrated in the binomial test in Figure 82 below, reveal notable distinctions in how listeners from different dialects perceive plain and emphatic vowels. For the /i/ vowel, both plain and emphatic realizations are accurately recognized by listeners across all dialects, demonstrating the robustness of the acoustic-phonetic cues distinguishing plain and emphatic /i/. This consistent recognition underscores the clarity of these cues for /i/, irrespective of the listeners' dialectal backgrounds. In contrast, the recognition patterns for the /a/ vowel are more variable. Moroccan listeners show significant difficulty recognizing emphatic /a/ vowels across all dialects, including their own. Additionally, they struggle with plain /a/ vowels, with the exception of plain /a/ presented in Syrian stimuli. This suggests that Moroccan listeners may be less attuned to the cues marking the plain-emphatic distinction for /a/, even within their native dialect. On the other hand, Syrian and Iraqi listeners successfully recognize emphatic /a/ vowels as produced in their own dialects and each other's, indicating comparable acoustic-phonetic patterns for emphatic /a/ in these two dialects. However, both Syrian and Iraqi listeners fail to recognize Moroccan plain /a/ vowels and often misclassify them. This suggests a distinctive characteristic in Moroccan /a/ vowels that renders them difficult to interpret for non-Moroccan listeners and even for Moroccan listeners themselves.

Figure 82 Binomial test results showing response accuracy relative to chance across listeners from Moroccan, Iraqi, and Syrian dialects. The plot categorizes accuracy based on p-values into "Significant ($p < 0.05$)", "Not Significant ($p \geq 0.05$)", and "Approaching Significance (p close to 0.05)". Stimuli are divided into "real" and "manipulated" conditions: real stimuli reflect natural productions, while manipulated stimuli are classified as "Iraqi-like," "Moroccan-like," or "Syrian-like" based on their phonetic characteristics. Horizontal dashed lines at 50% represent the chance level, with listener responses above or below this threshold analysed for statistical significance.

4.5.7.2 Binomial test Results for Manipulated Stimuli

The recognition patterns for manipulated stimuli, as shown in the binomial test in Figure 82 above, differ from those for real stimuli, particularly for emphatic vowels. For manipulated plain stimuli, both /a:/ and /i:/ are recognized at above-chance levels by all listeners. However, Moroccan listeners fail to recognize manipulated Moroccan plain /a:/, further reinforcing the observation that Moroccan plain /a:/ vowels are challenging even in their real forms. In contrast, the recognition of manipulated emphatic stimuli for vowels /a:/ and /i:/ shows a general pattern of below-chance recognition or inaccurate above-chance recognition across all listeners. Syrian listeners are an exception, as they successfully classify manipulated emphatic Syrian /a:/, indicating that the emphatic cues were identifiable in this specific combination for Syrian listeners. The raw results in Section 4.5.4 for manipulated stimuli suggest that listeners are generally better at attending to their own dialect's emphatic F2 trajectory cues. This effect of listener dialect is observable only in the emphatic condition, where native listeners of a dialect can significantly better recognize manipulated emphatic F2 cues that resemble those of their own dialect compared to other dialect groups (with the exception of Iraqi-like emphatic /i:/). For plain stimuli, listeners across all groups perform similarly in perceiving the absence of emphatic cues. However, when an emphatic cue is present, and it does not match the one they are accustomed to in their own dialect, listeners struggle to identify it accurately. In such cases, they often misclassify the emphatic vowel as plain.

4.5.8 **Discussion and Conclusion**

In this section I will first revisit the aim and rationale of the study to establish a clear connection between the results and the ensuing discussion. This will involve explaining the stimulus conditions for the target dialect blocks. Then, a discussion will be presented for each dialect block regarding the vowels /a/ and /i/. Finally, I will provide the motivation for conducting a further articulatory study aimed at offering a more comprehensive explanation of the emphatic contrast phenomenon.

4.5.8.1 Revisiting the rationale and the aim of the study

The acoustic study found that aspect of the F2 trajectory are a robust acoustic correlate to distinguish the emphatic contrast. This perceptual study utilised a subset of the acoustic study data and has resynthesized some stimuli based on the acoustic results for four representative target dialect conditions which were visualised in Figure 25, reproduced in acoustic chapter 3. The resynthesized stimuli recreated different types of trajectories for the vowels /a/ and /i/. Thus, for the vowel /a/, a trajectory where the size of the F2 was medium throughout (*Medium_100*) equates to the Moroccan dialect. The trajectory where the size of F2 lowering was large throughout (*Low_100*) equates to the Syrian dialect. Finally, the trajectory with a large effect of F2 only in the first 30% of the vowel (*Low_30*) equates to Iraqi. As for the vowel /i/, a trajectory where the size effect of the F2 was medium throughout (*Meduim_100*) equates to the Moroccan dialect. The trajectory where the effect of F2 was observed to be medium only in the first 30% of the vowel (*Medium_30*) equates to the Syrian dialect. Finally, the trajectory where the effect of F2 was large only in the first 50% of the vowel (*Low_50*) equates to Iraqi.

4.5.8.2 Discussing the findings

As can be seen from the overall results of the real stimuli, all listeners from each dialect could identify their own native formant cues with higher accuracy in plain condition. Crucially however, all listeners can attend to the cues to emphasis of other dialects with considerable accuracy; this may be due to maintaining all possible phonetic cues including full formant values: F1, F2, and F3. That said, the binomial test results revealed that Moroccan plain and emphatic /a/ vowels are outliers in this acoustic-phonetic space, making them difficult to recognize for both non-Moroccan listeners and Moroccan listeners themselves. The findings also suggest that the cues distinguishing plain and emphatic realizations are clearer for /i/ than for /a/, as evidenced by the consistent recognition of /i/ vowels across all dialects. Moreover, listeners generally demonstrated the ability to interpret the plain-emphatic distinction in both their own and other dialects, with the notable exception of Moroccan /a/ vowels. This observation highlights the cross-dialectal clarity of these cues, particularly for /i/, while emphasizing the challenges posed by the distinctive acoustic-phonetic properties of Moroccan /a/. A possible explanation for the challenges associated with Moroccan /a/ vowels may lie in the findings of the acoustic study (see Figure 38 in the acoustic study chapter). The study revealed that the F2 trajectory of Moroccan /a/, for both emphatic and plain conditions, is characterized by a relatively smaller F2 effect size, with the trajectories of plain and emphatic /a/ being relatively close to each other. This reduced acoustic distinction between plain and emphatic realizations may account for the difficulty listeners—both Moroccan and non-Moroccan—face in recognizing the contrast.

For the manipulated stimuli, the F2 size and trajectory shape was observed to play a significant role for listeners to be able to identify their own cues, in emphatic condition (only). Thus, for the vowel /a/, if the size effect of the F2 trajectory was medium throughout, Moroccan listeners tended to identify it with greater accuracy. Conversely, if the effect size of F2 was large throughout, Syrian listeners were the only ones who could perceive it at above chance levels and significantly better than the other dialect listeners. Similarly, if the size effect of F2 was large only in the first 30% of the trajectory, Iraqi listeners found it easier to recognize than other listeners did. As for the vowel /i/, Moroccan listeners were seen to attend accurately to the small F2 size effect throughout. On the other hand, if the F2 effect size was small only in the first 30%, Syrian participants tended to identify it more effectively than participants from other dialects. The only exception to the above patterns was that Iraqi listeners were no better than listeners in other groups at identifying the /i/ vowel if the size effect of F2 was 'Iraqi-like' in being large only in the first 50% of the vowel. Having presented these findings, the binomial test results indicate that for the emphatic stimuli, only Syrian listeners could attend to their F2 signals with statistically significant accuracy. In contrast, listeners from other dialects, including Iraqi and Moroccan, demonstrated above-chance accuracy in attending to their F2 cues, but this was not statistically significant. Additionally, the binomial test results revealed that only Moroccan listeners were unable to reliably recognize their own plain /a/ cues, as well as the plain /a/ cues of other dialects, with significant accuracy. This reinforces the observation that Moroccan plain /a/ vowels present a considerable challenge, even in their real forms.

Overall, it is evident that the notable recognition of manipulated plain stimuli corresponds to the notable recognition accuracy observed for real plain stimuli among all listeners. This suggests that the choice of the base stimuli was appropriate, and that the splicing technique was implemented effectively without significant issues. However, the poor recognition of manipulated emphatic /a/ stimuli partially aligns with the difficulties some listeners had with certain real emphatic /a/ stimuli. Moreover, the poor recognition accuracy of manipulated emphatic /i/ stimuli represents a complete mismatch to how listeners responded to real emphatic /i/ stimuli, highlighting potential shortcomings in the manipulation process. This discrepancy may be attributed to the fact that the manipulation was conducted manually, with only the F2 trajectory being manipulated. All potential limitations will be thoroughly addressed in the limitations section, along with proposed solutions for future research.

The prediction of the current study was that listeners would attend to their own native cues with better accuracy compared to listeners who are not accustomed to hearing these cues. This has been clearly observed in the results of the manipulated stimuli, specifically in the emphatic condition, despite the lack of significance indicated by the binomial test, for all dialect groups except Syrian listeners. However, the results of the real stimuli suggested that all listeners were able to attend to the cues of both, their own and other dialects with high accuracy, despite the fact that listeners of some dialects, such as Moroccan, performed at chance level when identifying either their own or other dialect cues. This inconsistency between real and manipulated stimuli implies that F2 is not the only cue that listeners are picking up on in the real stimuli.

The main argument of prior research has been that the effect of the emphatic contrast lies primarily in the formant information of the vowel following the emphatic contrast. The acoustic study results supported this, although it was found that each dialect produces the post-emphatic vowel differently from the others. The details of the F2 trajectory have not been discussed previously in the existing literature and thus the current perceptual study was conducted to determine the extent to which listeners pay attention to these differences in the vowel. Despite this, the results did not confirm that the effect can solely be due to the F2 properties of the post-emphatic vowel. Consequently, there is a need to examine articulatory evidence, and specifically ultrasound data, to determine whether there is any articulatory rationale for the emphatic contrast. Specifically, there might be something related to the tongue position and/or something radically different in the tongue position. The potential question to address is: Is there an articulatory difference in the fricative? In the next study therefore, we move on to investigate the distinction between /s/ and /sˤ/ in terms of the tongue shapes and potential lip involvement using ultrasound imaging.

5 Chapter 5: The Articulatory Study

5.1 Introduction

The results from the acoustic study in chapter 3 indicated that F2 served as the primary and reliable cue for distinguishing between the emphatic contrast, especially in the case of s~sˤ. Additionally, the Generalised Additive Mixed Models (GAMM) results demonstrated that the effect of F2 is observable at different points in the vowel trajectory, although dialects varied in the shape and size of the trajectory slope: in other words, dialects differ in how large the F2 difference is and how much of the vowel is affected. While the results also suggested that Centre of Gravity (COG) showed differences between the emphatic contrast, these differences were not statistically significant and not consistent between dialects, suggesting that the fricative consonants do not play a substantial role in representing the distinction for the emphatic contrast. Subsequently, in the follow-up perception study in chapter four, F2 trajectory information was utilised and manipulated to (1) conform to the acoustic results, (2) determine whether the F2 information offered by the vowels (in the emphatic vicinity) is used as a perceptual cue to emphasis, and (3) to determine whether listeners across all dialects will attend differently to the manipulated and real signals. However, Inconsistencies emerged in the findings: listeners were able to attend to their manipulated signals but only in the emphatic condition, while all listers were able to attend to real-stimuli cues of their own and other dialects with high accuracy, suggesting that F2 might not be the only cue that listeners rely on. The current study will employ ultrasound imaging to explore further and substantiate these results. This approach aims to investigate whether or not there are genuine articulatory differences (in terms of both lingual and/or labial articulation) involved in the emphatic s~sˤ contrast that are not represented acoustically. i.e. investigating the covert articulation. Thus, this chapter begins by providing a general overview of studies investigating emphatic contrast from articulatory perspectives, focusing on non-ultrasound instruments employed in prior research. It then specifically reviews studies that have utilised ultrasound imaging to examine emphatic contrast. An integrated summary will follow, along with the research rationale and hypothesis. Finally, the methodology and results of this study will be presented, followed by an in-depth discussion of the findings and their implications within the broader context of the chapter on acoustic results and the existing literature.

5.2 Background

To gain a deeper understanding of articulation of emphasis, it is necessary to first revisit the foundational work of medieval Arabic grammarians such as Sibawayh and Ibn Sina (980–1037 AD), whose insights on emphasis mechanism continue to inform contemporary studies (Aldamen, 2013). They provided explanations regarding the mechanism of emphasis and showed how emphatic and non-emphatic sounds differ. "Sibawayh gives the clearest description of emphasis or /?itbag/: for 'covered' sounds, the tongue is covered (by the palate, presumably) from the place of articulation back to the place where it is raised towards the palate. The additional posterior tongue raising gives these sounds two places of articulation (Card, 1983, p. 7)." Sibawayh believed that the mechanism of the consonant itself could reveal differences in the contrast, a concept later supported by many researchers (see Card, 1983; Zawaydeh, 1999; Jongman et al., 2010). This historical perspective was a motivation to re-look at the analysis of the emphatic consonant /sˤ/ and its counterpart /s/ from an articulatory perspective.

It has been discussed in the literature that emphasis in Arabic involves two constrictions: primary and secondary. Emphatic and non-emphatic sounds share similarities in terms of their primary constriction, which typically occurs at the alveolar or dental regions (Aldamen, 2013). However, they differ significantly in terms of the secondary articulation which occur at the back of the oral cavity (Al-Solami, 2017). Despite extensive articulatory research on emphasis in Arabic, a consensus has not been reached regarding the precise nature of the secondary constriction. This lack of consensus can be attributed, in part, to cross-dialectal variation and differences in the methodologies employed in research investigations (Aldamen, 2013 & Al-Solami, 2017). Khattab et al. (2006) noted that this inconsistency may arise from speakers employing various articulatory strategies to produce emphatic sounds. These strategies are influenced by factors such as dialect, gender, phonological context, and social variables. Ghazeli (1977), Bin-Muqbil (2006), and Shar (2012) suggested that in the articulation of emphatic consonants, the tongue body is pulled backward into the upper pharyngeal region, similar to the articulation of uvulars and is depressed at the point of the palate. Additionally, Ghazeli (1977) in his study pointed out that the tongue body is pulled backwards into the upper oropharynx during the articulation of [sˤ] and the tongue body is depressed during the emphatic consonant but not during the plain coronal. To understand more on the nature of emphasis, the two sections below discuss the articulatory studies that have been conducted with different articulatory instruments.

5.2.1 Studies with Non-Ultrasound Instruments

On Iraqi Arabic, Ali and Daniloff (1972) conducted a cinefluorographic film study for the purpose of examining the position of the tongue dorsum, the movement of the pharyngeal wall and the position of the velum during the articulation of Arabic nonsense words and syllables. Their findings suggested that emphatics are articulated with simultaneous tongue depression and a rearward movement of the tongue dorsum towards the posterior wall of the pharynx. They discovered that the difference between emphatics and non-emphatics is a retraction of the tongue dorsum, which causes a narrowing of the upper pharynx. They also found that the velum and the posterior wall of the pharynx were not significantly affected in the articulation of emphatics. Ibn Sina (1037 A.D.) initially stated tongue dorsum depression in emphatics, claiming that emphatics are articulated with a depressed tongue surface behind the main articulation point (Semaan 1963). Other research (e.g. Ali & Daniloff 1972 in Iraqi Arabic; Ghazeli 1977 in Tunisian Arabic; Al-Tamimi & Heselwood 2011 in Jordanian Arabic) support this point. Conducting an MRI study on Saudi speakers, Shar (2012) discovered that emphatics are produced with dorsal retraction of the tongue, causing constant narrowing of the top region of the pharyngeal cavity; however, the tongue root is not implicated in this narrowing gesture. In their video-fluoroscopic study of emphatics in Jordanian Arabic, Al-Tamimi & Heselwood (2011) found that during the articulation of emphatics, the tongue root is seen to press against the anterior surface of the epiglottis, pushing the epiglottis towards the back of the pharynx. Additionally, they suggested that the larynx is raised in emphatics, which means that the pharyngeal volume is reduced. This finding aligns with Zawaydeh's (1998, 1999) results which demonstrated that emphatic consonants are associated with pharyngeal narrowing. However, it is difficult to judge in the already reduced pharynx whether the tongue root/epiglottis movement is independent or a result of the tongue dorsum retraction. Accordingly, tongue root retraction in emphatics appears to be a mechanical consequence of tongue dorsum retraction (Altairi, et. al. 2017; Al-Solami 2017) and thus, there is no consistency in tongue root retraction. Due to this controversial point, researchers have posited that emphatics are uvularised in Jordanian Arabic (Zawaydeh 1999), and Moroccan Arabic (Zeroual et al. 2011), velarised in Lebanese Arabic (Obrecht 1968), and pharyngealized in Iraqi Arabic (Ali & Daniloff 1972; Gianni & Pettorino 1982) and in Hijazi Arabic (Ahyad, H., & Becker, M. 2020). Focusing on examining only the emphatic contrast /s/, and /sˤ/ Hermes, et all (2015) conducted an experiment using electromagnetic articulography (EMA), investigating the primary articulation of plain-emphatic /s/-/sˤ/ in Lebanese Arabic. The study found that the front of the tongue, up to 1 cm from the tongue tip, remains low behind the front teeth in approximately the same position in /s/. However, the position of the tongue at those points during /sˤ/ is lower than during /s/, which leads to a depression in the tongue blade. The differences between the emphatic contrast, /sˤ/ and /s/ are statistically significant in all speakers. With similar instrument but different results, Zeroual et. al. (2011) examined the emphatic stop /dˤ/ and /tˤ/ and their plain counterparts in Moroccan dialect and found that plain sounds are articulated with a more laminal contact than the emphatic counterparts, while emphatic sounds are more apical. They also reported slight labialisation during the emphatic sounds. Similarly, some research reported some degree of lip protrusion associated with emphasis (see El-Halees, 1985; Hetzron, 2013; Jakobson 1957). Lehn (1963) for example, reported a slight lip rounding in Carine dialect while Bellem & Watson (2014) reported a labialisation in Ṣanʿāni Arabic and in southern (gilit) Iraqi Arabic. All these instances of lip-rounding (except for Zeroual et. al., 2011) were reported based on visual observation and the theory of emphatic spread and its effect on tongue movements but not based on instrumental evidence. Thus, in this current study an ultrasound imaging examination will be utilised alongside simultaneous lip camera recording to investigate the relationship between lingual and labial articulation in the production of the emphatic contrast.

5.2.2 Studies Using Ultrasound Imaging

A study proposed by Altairi, et. al. (2017), recruited eight speakers of different Arabic dialects, (2 Saudi, 2 Yemeni, 2 Egyptian, 1 Palestinian) and examined the tongue movements of various sound groups (/tˤ/, /sˤ/, /t/, /s/, /ħ/, /ʕ/, /χ/, /ʁ/) using the Smooth Spline ANOVA (SS-ANOVA) analysis and comparing among different plain/emphatic groups and compare them to their neutral position of the tongue. The researchers here followed Gick et al. (2004) of using the ‘inter-speech posture’ (ISP) to act as a baseline from which the postures of speech sounds were compared and measured. The findings suggested that there is a significant difference in tongue root (TR) position between the emphatic and non-emphatic consonants for all subjects. The subjects articulated the emphatic consonants with substantial TR retraction compared to the non-emphatic counterparts. Additionally, six subjects exhibited a significant difference in tongue dorsum (TD) position between the emphatic and non-emphatic consonants, where TD in the emphatics is backed and higher than that of the non-emphatic counterparts. As seen from the performances of almost all participants, the emphatics are produced with a consistently posterior tongue root relative to the ISP.

Additionally, the retraction of the tongue root for emphatics involves a consistent lowering of the tongue body position compared to the ISP. However, there is no difference between the emphatics and the ISP with regards to the tongue dorsum except for two participants who have the emphatics articulated with the tongue dorsum raised compared to the ISP. When it comes to the non-emphatics, the tongue root is advanced, and tongue dorsum lowered compared to the ISP.

A similar use of ultrasound tongue imaging was also seen in Al-Solami, (2017). He conducted a study of three participants from three Arabic dialects: Saudi Arabia, Egyptian, and Palestinian. The aim of his study was to examine differences in tongue movements for the emphatic, uvular and pharyngeal sounds. He hypothesised that these sounds involve a tongue retraction as the main articulatory component but that they differ in the degree of retraction. The findings of the emphatic contrast suggested that the tongue dorsum is more raised and retracted and the blade behind the point of main constriction is depressed during the emphatics. Alongside the EMA and endosopic approaches taken by Zeroual et al. (2011), described earlier, un ultrasound investigation was also performed in the same study. In order to answer a number of questions, Zeroual et al. (2011) collected data from Moroccan Arabic speakers. The relevant point here is the investigation of the nature of secondary articulation in Moroccan emphatics. They compared the properties of Moroccan emphatic coronals /tˤ, dˤ, sˤ/ to their plain counterparts /t, d, s/, as well as uvulars, and pharyngeals. They recruited two Moroccan speakers for the ultrasound research, eliciting from them both real words and nonsense words that include emphatic sounds with the goal of examining the positions of the tongue and the epiglottis. Their results indicate that the articulation of emphatics is more similar to uvulars than pharyngeals. Furthermore, emphatic sounds included a backward movement of the tongue towards the posterior pharyngeal wall, whereas pharyngeals included a backward movement of the tongue and the epiglottis.

Another ultrasound study was conducted by Alfaifi et al. (2020) on two Saudi Hijazi participants; their research aim was to examine the primary constriction of the voiceless stop /tˤ/ and the fricative /sˤ/ along with their corresponding plain counterparts /t/ and /s/. They also aimed to investigate the impact of emphasis on the adjacent vowel and vice-versa. Their findings indicate that the tongue root in emphatics is elevated and more retracted compared to the non-emphatics and short vowels are more susceptible to the emphasis impact than the long vowels. In other words, short vowels have their tongue root more retracted in the context of emphasis. They also found that high vowels had an impact on the tongue body shape when producing the emphatics, but this was mostly observed with one of the participants. In this current study, however, the tongue splines will be examined by considering the vowel type factor across all participants to account for the shape of the tongue.

5.2.3 Summary and Rationale/Hypothesis

In summary, most of the aforementioned studies have primarily focused on investigating the differences among guttural sounds, comparing tongue movements within and between different sound classes (emphatics, pharyngeals, uvulars, and laryngeals). Although they have provided some accounts about emphatic and non-emphatic distinctions, this was not their main focus. The relevant point here is to determine what these studies have found regarding the distinctions between (s~sˤ) in terms of tongue positions. All of the research agrees on the tongue configurations for the primary constriction, but there is slight variation in the description of tongue positions during the secondary constriction. This might lead researchers to characterize Arabic dialects as pharyngealized, uvularised, or velarised. The general observation is that in the case of emphatic sounds, the tongue dorsum is retracted and raised, and the area behind the front part of the tongue (blade) is depressed, forming a concavity compared to non-emphatic sounds. Additionally, tongue root retraction is more observable in emphatic sounds compared to non-emphatics; among guttural sounds, this is dependent on the degree of retraction of the tongue dorsum. While some accounts of lip-rounding during the articulation of emphatics have been observed in the literature, these investigations have primarily relied on acoustic or auditory inspection, and this has never been studied with direct articulatory evidence. As a result, this study aims to examine the distinction between /s/ and /sˤ/ in terms of lingual articulation as well as labial articulation. The goal is to determine whether the differences in tongue configurations between emphatic and non-emphatic sounds are consistent across the target dialects and to what extent ultrasound imaging results directly reflect the acoustic speech signal. For the purposes of this study, it is hypothesised that the emphatic /sˤ/ is articulated with the retraction of the tongue dorsum, depression of the tongue surface behind the main articulation point, and lip rounding. This mechanism is proposed to distinguish emphatic sounds from their plain counterpart, /s/.

5.2.4 Research Questions

1. Is there an articulatory difference between the fricatives /s/ and /sˤ/?
   - Is there any inter-dialectal variation in how the emphatic contrast is articulated?
   - Is there any individual-specific patterns of articulatory mechanisms?
2. Does lip rounding contribute in distinguishing between emphatic contrast among Arabic dialects?
3. Do the articulatory findings align with the acoustic conclusion that fricative consonant has no main effect in distinguishing the emphatic contrast?

5.3 Method

5.3.1 Participants

This experiment included fifteen male participants: one Moroccan, five Syrians, two Egyptians, and three Iraqis. The choice of these dialects was based on the fact that they were the same dialects examined in the perception study. The study also recruited 4 participants from Hijazi Saudi; the inclusion of Hijazi participants was motivated by previous research (e.g., Alfaifi et al., 2020), which utilised ultrasound imaging to examine Saudi participants and revealed that elevation of the tongue root was crucial in distinguishing between emphatic and non-emphatic sounds, marking the dialect as pharyngealized. Alfaifi et al. (2020) also sought to examine the relationship of the tongue root to the primary constriction in articulation. However, their analysis did not yield comprehensive results due to the exclusion of four participants due to technical challenges. In this study, the aim was to compare the results of Alfaifi et al. (2020) to the current findings and to see the extent to which this conforms to the results of other Arabic dialects. For this study, all participants used to live in the northeast of England at the time of recording. They were all aged between 18 and 45 years and spoke their respective dialect natively; it was ensured that none of the participants had a speech or hearing impairment. It is worth noting that female participants were not included in this study, primarily due to the challenges of recruitment and the difficulty of accessing participants within the limited timeframe. Future studies will consider strategies to facilitate broader participation, including that of female participants. Please see details for speakers' codes and their respective dialects in Table 39 below.

Speaker code	Age	dialect	Duration of living in L1 Country	Duration of living in the UK
hh01	36 - 44	Saudi, Hijazi	35 y	7 y
hh02	25 - 35	Saudi, Hijazi	24 y	5 y
hh03	45 and above	Saudi, Hijazi	45 y	15 y
hh04	18 - 24	Saudi, Hijazi	10 y	8 y
eg01	18 - 24	Egyptian, Alexandria	16 y	8 y
eg02	25 - 35	Egyptian, Aswan	26 y	6 m
sy01	25 - 35	Syrian, Damascus	10 y	4 y
sy02	25 - 35	Syrian, Darah	28 y	16 m
sy03	18 - 25	Syrian, Jawlan	11 y	8 y
sy04	25 - 35	Syrian, Aleppo	15 y	6 y
sy05	25 - 35	Syrian, Jawlan	20 y	5 y
ir01	45 and above	Iraqi, Basra	33 y	8 y
ir02	36 - 44	Iraqi, Baghdad	35 y	8 y
ir04	25 - 35	Iraqi, Basra	30 y	3 y
mo02	18 - 25	Moroccan, Tanja	18 y	4 y

Table 39 Information for speakers' codes and demographic information. Y = years and m = month

5.3.2 Stimuli

The stimuli for this experiment consisted of monosyllabic minimal pair words. Each word took the form C_1 V_1C_2, where C_1 represented either the emphatic /sˤ/ or the non-emphatic /s/, and V_1 represented either the low vowel /a/ or the high vowel /i/. To ensure participants articulated the target words in their native dialects, words were embedded in three carrier phrases specific to corresponding dialects. These phrases are:

- **Egyptian and Hijazi:** "howwa galli X marratiin" ("he told me X two times")
- **Syrian:** "howwa Hakaali X marratiin" ("he told me X two times")
- **Moroccan:** "howwa galli X jooj marrat" ("he told me X two times")

This resulted in a total of 14 words x 15 participants x 3 repetitions = 630 tokens[9]. Table 40 shows examples of the stimuli used in the experiment.

[9] The stimuli also included distractors, "ta:b and tˤa:b", and "ti:n and tˤi:n" as shown in Table 40.

Plain			Emphatic			
IVAr	IPA	translation	IVAr	IPA	translation	vowel
saar	sa:r	‘he walked’	Saar	sˤa:r	‘it became’	a:
saam	sa:m	*‘poisonous’*	Saam	sˤa:m	*‘he fasted’ use*	a:
saab	sa:b	‘he left it’	Saab	sˤa:b	‘he shot something’.	a:
taab	ta:b	‘he repented’	Taab	tˤa:b	‘he got better’	a:
siin	si:n	‘the letter s’	Siin	sˤi:n	‘China’	i:
siib	si:b	‘a corridor’	Siib	sˤi:b	‘hit something!’	i:
tiin	ti:n	‘fig’	Tiin	tˤi:n	‘mud’	i:

Table 40 Stimuli that are used for the UT experiment

5.3.3 Equipment and Procedure

Ultrasound Tongue Imaging (UTI) was employed to capture the midsagittal plane of the tongue during each repetition, utilising Articulate Assistant Advanced (AAA) software, version 219.08 (Articulate Instruments Ltd. 2022). The imaging process utilised a MC4-2R20S-3 microconvex-array ultrasound transducer, with probe frequencies between 2 MHz and 4 MHz, to obtain the tongue recordings. In order to hold the ultrasound probe in place an UltraFit headset (Spreafico et al. 2018) was used, while simultaneous audio was recording with a Rode smartLav+ wired lavalier microphone at a Nyquist frequency of 22050 Hz. To prevent any air from coming between the surface of the probe and the skin, a non-toxic water-based gel was applied (Stone 1997). Finally, simultaneous video footage of the lips was captured using a small camera fixed to a bracket on the head stabilising unit, offering a frontal view of the lips and recording at approximately 30 frames-per-second.

Participant recordings took place in a quiet booth in the Speech Articulation lab at the Department of Language and Linguistic Science at the University of York. Before the recording session, all participants were instructed to read phrases containing the target stimuli. This was done to ensure that participants read the stimuli correctly. All phrases were written in Arabic script and presented to participants on PowerPoint slides, as AAA does not support Arabic text. Participants were seated in a comfortable chair facing the laptop on which they could see the stimuli. With the help of the researcher, participants were asked to wear a head stabilising unit with an ultrasound transducer placed beneath their chin, above the larynx and was set to 90° to view the complete shape of the tongue. This transducer emits inaudible high-frequency soundwaves to create images of the tongue's surface. At the start of recording, participants were instructed to briefly bite down on a disposable wooden tongue depressor, which will then be discarded, in order to act as a fiducial marker for contextualising the angle of recording.

5.3.4 Data Synchronisation and Pose Estimation

The audio recordings and the ultrasonic videos of the tongue are synchronised by default in AAA, but lip footage requires manual synchronisation. Following data acquisition of ultrasound and video recordings of tongue and lip movements during each repetition of the stimuli, DeepLabCut™ [10] software was employed to automate the extraction of tongue and lip splines from each recording using a process of markerless pose estimation (see Wrench & Balch-Tomes, 2022). Figure 83 exemplifies the pose-estimated tongue contour generated by DeepLabCut™ for the first repetition of the Arabic word "Saar" spoken by a Hijazi speaker. This figure depicts eleven key anatomical landmarks along the tongue surface, namely the vallecula, two locations on the tongue root, two points on the tongue body, two points on the tongue dorsum, two points on the tongue blade, and two points on the tongue tip. These points are tracked dynamically for each video frame, resulting in a large and complex time-series dataset.

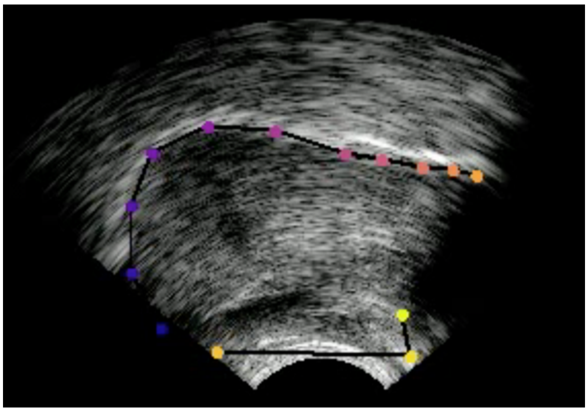


Figure 83 Pose estimated tongue contour for the word ‘Saar in Arabic (hh3-Hijazi speaker)

5.3.5 Data Analysis

After exporting the tongue and lip spline coordinates into CSV format using DeepLabCut™, the respective audio recordings were manually labelled using Praat text grids, with interval

[10] The Mathis Group, & The Mathis Lab. (2022). *DeepLabCut* (Version 2.2.2) [Computer software]. Switzerland: The Mathis Group, Swiss Federal Institute of Technology Lausanne. Retrieved December 2023 from http://www.mackenziemathislab.org/deeplabcut

tiers at the word and phoneme levels. Analysis is conducted on the video frames corresponding to the target segment midpoint; all analysis was carried out in R, including the execution of Praat scripts (Boersma and Weenink, 2010) for acoustic measurement extraction that were executed within the R software environment.

5.3.5.1 Lingual Analysis

For the lingual analysis, a Geometric Morphometric (GMM) approach was taken in order to help with the comparison of tongue shapes between individuals. Introduced by Bookstein in 1996 (as referenced in Adams et al., 2004), GMM provides a statistically solid method for quantifying and comparing shapes. GMM involves two key steps:

1. Generalised Procrustes Analysis (GPA): The initial step involves (GPA), as described by Adams et al. (2004), which is a superimposition technique that reduces shape discrepancies by removing variations in landmark arrangement that do not pertain to shape, such as scale, translation, and rotation. This technique also standardises any positional differences between recordings that might arise from movements of the ultrasound probe (Polly, 2012).

2. Principal Component Analysis (PCA): In this step, the variation in tongue shape among speakers and different fricative sounds is quantified using (PCA). For shapes normalised through GPA, PCA identifies the primary directions of shape variation within the dataset (Polly, 2012). The most significant variation is captured by the first principal component (PC1), with subsequent components (e.g., the second principal component) capturing progressively lesser variations. In this study, the possible interpretation is that if the amount of variation for one fricative type is high in PC1 while it is low for the other fricative type, this indicates a separation in tongue shapes. In other words, if the emphatic instances in this study have significant positive PC1 values and the plain instances have significant negative PC1 values, this suggests that the two fricative types exhibit distinct tongue shapes. While PCA has been employed to measure shape in the vocal tract (Gully, 2021), its use with data derived from ultrasound represents a novel approach.

Following the acquisition of PCA scores, linear models were applied to these scores for null-hypothesis significance testing. Estimated marginal means were then obtained for pairwise comparisons of environments. The fitted model (shown below) analysed the interaction between the emphatic contrast /s/ ~ /sˤ/ and the vowel types /a/ and /i/ for each subject, analysing both PC1 and PC2 as separate dependent variables.

Model formula:

- mod <- lm(Comp1 ~ Emphatic_contrast * Vowel_type, data = data)

5.3.5.2 Acoustic Analysis

Acoustic information was also extracted from the speech signal at every video frame, for both target fricatives: various spectral measures were extracted, including skew, kurtosis and Centre of Gravity (COG), with the analysis here conducted solely on the latter of these measures. It is worth noting that these measures are the same as those examined in the acoustic study chapter. The obtained COG values for the emphatic contrast were modelled and analysed using the following linear regression model:

Model formula:

- mod <- lm(COG ~ Emphatic contrast * Vowel type, data = data)

5.4 Results

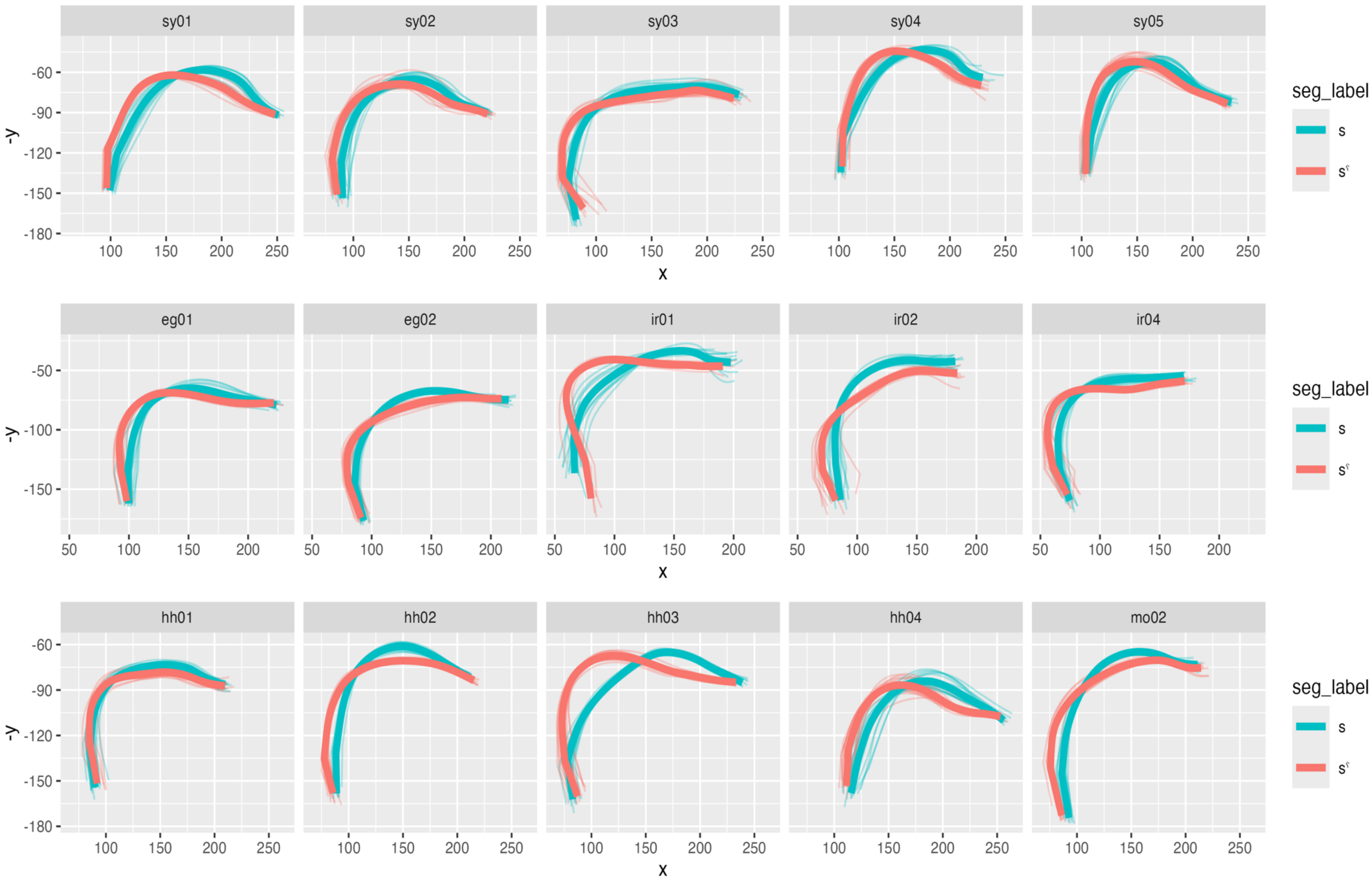
This section presents the results of the ultrasound experiment covering both articulatory and acoustic analysis, specifically comparing the COG and mid-sagittal tongue posture across both conditions: emphatic /sˤ/ and the non-emphatic /s/. To illustrate the differences in lingual articulation for the emphatic contrast /s/ versus /sˤ/, a visual analysis will be conducted, emphasising inter-speaker variability in three key tongue mechanisms: tongue dorsum retraction, tongue root retraction, and tongue blade depression (the surface point posterior to the anterior tongue region). It is important to note that all figures will depict the tongue tip positioned to the right. Subsequently, the Principal Component Analysis (PCA) along with the statistical testing of these values will be discussed to identify significant differences in the emphatic contrast. Following this, a visual analysis of the acoustic results will be presented, followed by a statistical analysis of the COG values for the emphatic contrast across all speakers. This will be accompanied by a description of the acoustic-to-articulatory mapping findings: in other words, to what extent do the acoustic results reflect any differences in tongue posture from the ultrasound data. A description of the lip camera images will then investigate any relevant labial gestures (rounding or spreading) that may co-occur with the mechanisms of the tongue, and how this correlates with the acoustics. For clarity, each subsection will discuss selected figures from each speaker, with all relevant figures provided in the Appendix 11.4.

5.4.1 **Visual Inspection of Lingual Articulation**

Figure 84 below displays the average tongue tracings for the emphatic /sˤ/ and plain coronal /s/. The differences between emphatic /sˤ/ and plain /s/ are noticeable in the dorsum and blade of the tongue. Specifically, for the emphatic sound, the dorsum of the tongue is more retracted compared to the non-emphatic /s/ across all subjects. However, this retraction varies among subjects. For instance, subjects like 'hh03', 'ir01', and 'ir02' exhibit a significant difference in tongue shape between emphatic and non-emphatic sounds, with greater retraction for the emphatic sounds. Other subjects such as 'hh02', 'sy01', and 'eg02' also demonstrate a notable difference in the shape of the emphatic tongue, while the remainder show a moderate to minor difference in tongue shape for the emphatic contrast. Regarding the tongue blade, despite some articulatory variability, only the emphatic /sˤ/ is associated with a depression at the point behind the main constriction, a mechanism not observed with the non-emphatic /s/.

As previously described, the primary articulatory component of emphatic sounds is the retraction of the tongue dorsum. However, the degree and direction of this retraction vary across dialects. Accordingly, Figure 84 illustrates that for all speakers, the back of the tongue is retracted towards the posterior wall of the pharynx. Subjects hh03, ir01, and sy04 exhibit more tongue dorsum retraction, while subjects sy02, sy03, hh02, mo02, ir02, and hh04 show more retraction of the tongue root. Although there appears to be inter-speaker (or inter-dialectal) variation in the degree and manner of the articulatory contrast between /s/ and /sˤ/, it is clear to see from Figure 84 that *all* speakers show at least some articulatory difference in the fricative.

Figure 84 Smoothed average tracings of tongue shapes for emphatic [sˤ] and plain coronal [s] consonants, as produced by speakers from various dialect groups. The figure illustrates the mean tongue shape for each speaker, averaged across repeated tokens, with the colour blue representing plain [s] and the colour red representing emphatic [sˤ]



5.4.1.1 Principal Component Analysis (PCA)

PCA is a technique used for dimensionality reduction; in other words, it transforms a complex tongue shape into a set of 'principal component' scores. These scores represent the most relevant axes or dimensions of variation in tongue shape (Jolliffe & Cadima, 2016). It is worth noting that the observation in Figure 84 provides a qualitative picture, while the PCA data offers a quantitative analysis to support and strengthen the qualitative findings. Appendices 11.1-11.3 present the principal component scores for all speakers. As previously mentioned, R software was utilised to extract these PCAs. Each point represents a tongue shape for a single prompt, either a fricative /s/ or /sˤ/ before the vowel /a/ or /i/. Appendix 11.4 illustrates the tongue shapes corresponding to the extremes of the PC axes. For simplicity, the PCA results for each dialect are presented sequentially.

5.4.1.2 Hijazi Subjects

Figures 85 and 86 below present PCA results along with their respective traced PCA scores for subjects hh01, hh02, hh03, and hh04. Specifically, in Figure 85 the plots on the left show the values of the first two principal components (PC1 and PC2) for each prompt. If the points cluster together by fricative type, particularly with non-overlapping distributions, then this suggests a greater separation in terms of tongue shapes. The plots on the right help in interpreting what those PCs actually represent in terms of the tongue shape differences. So, we can see what a higher PC1 means in articulatory terms for a particular speaker.

In Figure 85, we observe that PC2 accounts for only 4-20% of the variation across the four speakers, compared to PC1, which accounts for 48-86%. it is noticeable that PC2 does not significantly impact the variations in tongue shape between the emphatic contrast; however, most emphatic instances are associated with high PC1 scores, whereas most instances of the plain /s/ are linked to low PC1 scores. This indicates that the retraction of the tongue body is prominently observed within emphatic instances. This observation is corroborated by Figure 86, where the green points signify the retraction shape of the tongue, corresponding in this instance to the emphatic sounds. Considering variability across subjects, the contribution of PC1 to the tongue shape variation is notable. For subject hh03, approximately 86% of the variation is attributed to PC1, showing a clear separation between the tongue shapes of emphatic and non-emphatic. This is followed by subject hh02, which accounts for about 70% of the variation due to PC1. For subjects hh04 and hh01, PC1 accounts for 64% and 48% of the variation, respectively. The former indicates that some non-emphatic instances are marked with positive PC1 scores, suggesting that the PCA identifies these instances as emphatic due to slight tongue retraction. Conversely, the latter demonstrates that plain instances are closely clustered with consistent PC1 scores, but a significant number of emphatic instances overlap with non-emphatic ones. Nonetheless, PC1 still captures their tongue body retraction, as illustrated in Figure 85. It is worth mentioning that PC2 captures something intriguing for speaker hh04: within the emphatic /sˤ/ category—and to a lesser extent, the plain /s/ category—low PC2 scores are associated with fricatives preceding /i/, rather than /a/. This suggests that for this speaker, PC2 is capturing differences in tongue dorsum elevation related to coarticulation with the following high /i/ vowel.

Although PC1 accurately captures the tongue shape for the emphatic /sˤ/ across these four Hijazi speakers, there is noticeable variation among them regarding the specific appearance of that tongue shape. For instance, it appears that there is greater retraction at the root for speaker hh02 compared to, for example, hh03, where the emphasis is more on the tongue dorsum.

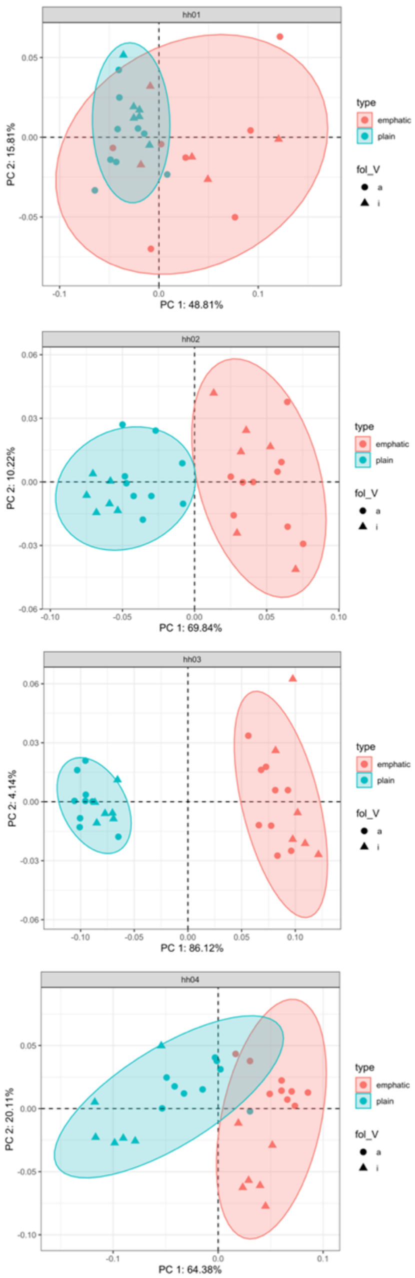


Figure 85 PC1 and PC2 scores of the xy coordinates during /sˤ/ and /s/ for Hijazi subjects

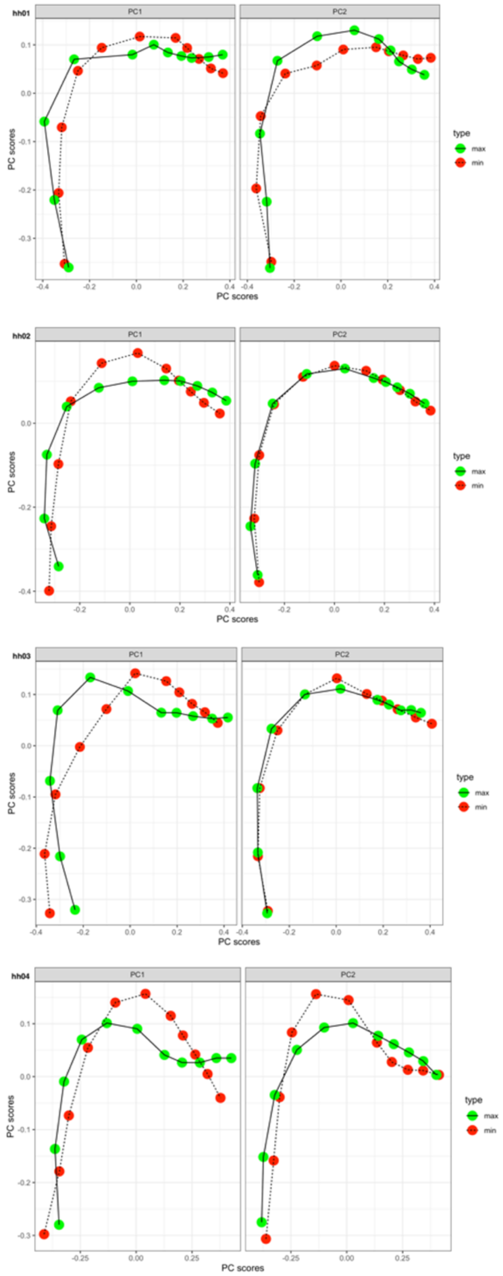


Figure 86 PC1 and PC2 tongue tracing for Hijazi subjects

The linear regression model output suggests that for subject hh03, the PC1 score is significantly lower for non-emphatic sounds compared to emphatic ones ($\beta = -0.075$, $SE = 0.011$, $p < 0.001$). This indicates greater tongue shape variation associated with the emphatic contrast. Similar significant differences in tongue variation were observed for subjects hh02 and others, as shown in Table 41.

subject	Estimate	StdError	t value	P Value
hh01	**-0.07184**	**0.020592**	**2.452**	**0.002**
hh02	**-0.08279**	**0.007801**	**-10.614**	**0.0001**
hh03	**-0.17165**	**0.00569**	**-30.168**	**0.0001**
hh04	**-0.07493**	**0.010822**	**-6.924**	**0.0079**

Table 41 Linear regression results for PC1 scores for Hijazi subjects

5.4.1.3 Syrian Subjects

Figure 87 shows that for all subjects, the emphatic instances have a mostly positive PC1. This indicates that the tongue shape of the emphatic is retracted. However, subjects again differ in terms of the extent to which PC1 correlates with the plain~emphatic contrast. For example, the plain and emphatic fricatives of sy01 and sy04 show a clear or nearly clear separation along the PC1 dimension which accounts for a large proportion of variation in tongue shape (77% and 61%, respectively). PC1 accounts for less variation among subjects sy02, sy03, and sy05 (45%, 47%, and 48%, respectively) though the PCA traces in Figure 88 indicate that the retraction of the tongue shapes is still clearly captured but it differs across subjects. For example, subjects sy02 and sy03 have the tongue root retracted, while subjects sy01, sy04, and sy05 have tongue dorsum retraction. Notably, PC2 accounts for only 5-27% of the variation across the five speakers. However, low PC2 scores correspond to the pre-/i/ tokens for subject sy05. This is clearly evident from Figure 88, where the tongue shape appears higher during the articulation of the emphatic that followed by vowel /i/. The same subject also exhibits some overlap between the plain and emphatic fricatives, specifically before the vowel /a/ (represented by circle shapes). However, before the vowel /i/ (represented by triangle shapes), the contrast is observed in both PC1 and PC2. Thus, there is much less separation before /a/, but clear separation before /i/. Refer to Appendix C.

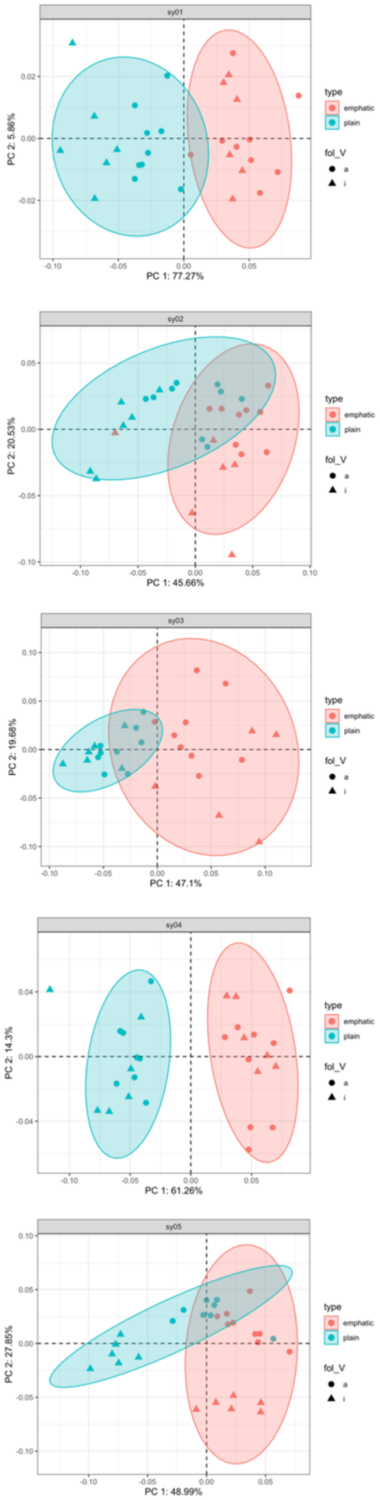


Figure 87 PC1 and PC2 scores of the xy coordinates during /sˤ/ and /s/ for Syrian subjects

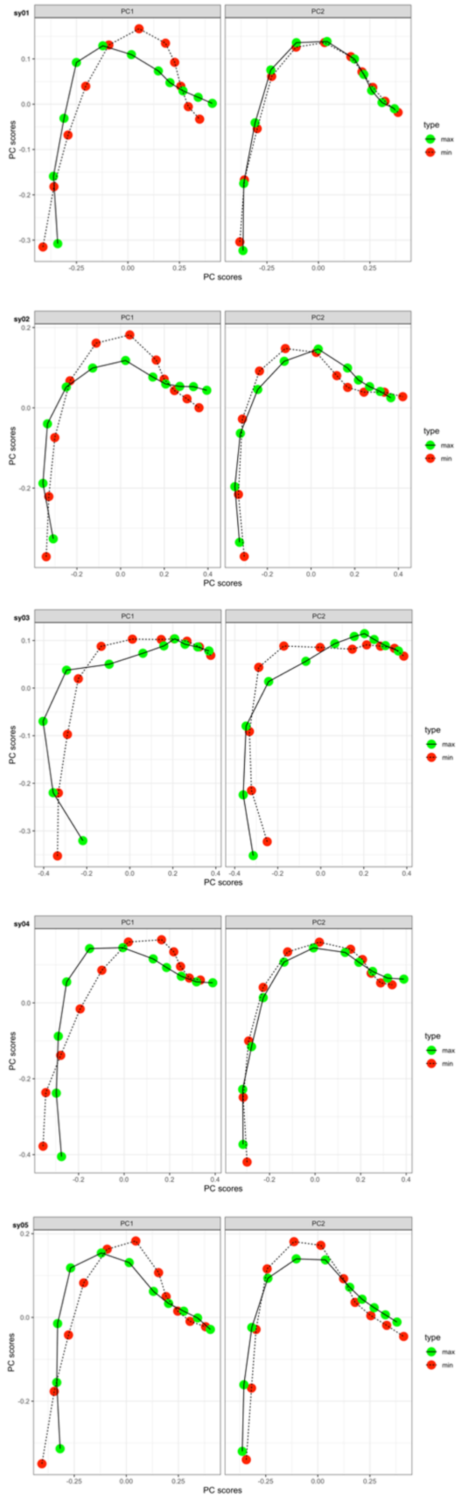


Figure 88 PC1 and PC2 tongue tracing for Syrian subjects

The linear regression model output reveals significant differences in PC1 scores for the emphatic contrast across all subjects, with plain instances exhibiting significantly lower PC1 values. Notably, subject sy04 exhibit an even more pronounced decrease in PC1 for non-emphatic sounds compared to emphatic ones, indicating that this speaker had the greatest degree of separation. Refer to Table 42 for details.

subject	Estimate	StdError	t value	PValue
sy01	**-0.07344**	**0.007864**	**8.639**	**0.001**
sy02	**-0.04367**	**0.012752**	**-3.425**	**0.002**
sy03	**-0.07015**	**0.01243**	**-5.644**	**0.001**
sy04	**-0.10051**	**0.008606**	**-11.678**	**0.001**
sy05	**-0.03135**	**0.0095**	**-3.300**	**0.003**

Table 42 Linear regression results for PC1 scores for Syrian subjects

5.4.1.4 Iraqi Subjects

Figures 89 and 90 present the PC scores for Iraqi subjects. In Figure 89, we can see that PC2 accounts for only 12-19% of the variation across all Iraqi speakers. In contrast, PC1 does show clearly that there is a differentiation between the two tongue shapes, accounting for 41-65% variation across all speakers: one corresponding to the emphatic sounds with higher PC1 values, and the other corresponding to plain /s/ sounds with lower PC1 values. This indicates that PC1 captures tongue shape variations associated with the emphatic contrast. To illustrate, Figure 89 shows that subject ir01 exhibits the highest variation in PC1 scores (around 65%) compared to the other subjects, indicating a clear separation between the two fricative types. Subjects ir02 and ir04 show less variation in PC1 scores than subject ir01, at 41% and 47%, respectively. Ir02 is characterized by a negative PC1 for non-emphatics and a positive PC1 for emphatics, indicating a retracted tongue body for emphatics and an advanced position for non-emphatics. In contrast, ir04 displays an even more retracted tongue shape for emphatics, with a positive PC1 for emphatics and a negative PC1 for non-emphatics. The degree of tongue retraction is observed to vary across Iraqi subjects. For example, PC1 tracing scores in Figure 90 show that subjects ir02 and ir04 exhibit tongue root retraction, while subject ir01 shows tongue dorsum retraction. These variations in tongue shape are further supported by the linear regression results in Table 43. The table shows that subject ir01 has the highest significant PC1 score, followed by ir02 and then ir04.

subject	Estimate	StdError	T Value	P Value
ir01	0.227433	0.034121	6.666	0.001
ir02	0.090115	0.013632	6.610	0.001
ir04	-0.07164	0.013271	-5.399	0.003

Table 43 Linear regression results for PC1 scores for Iraqi subjects

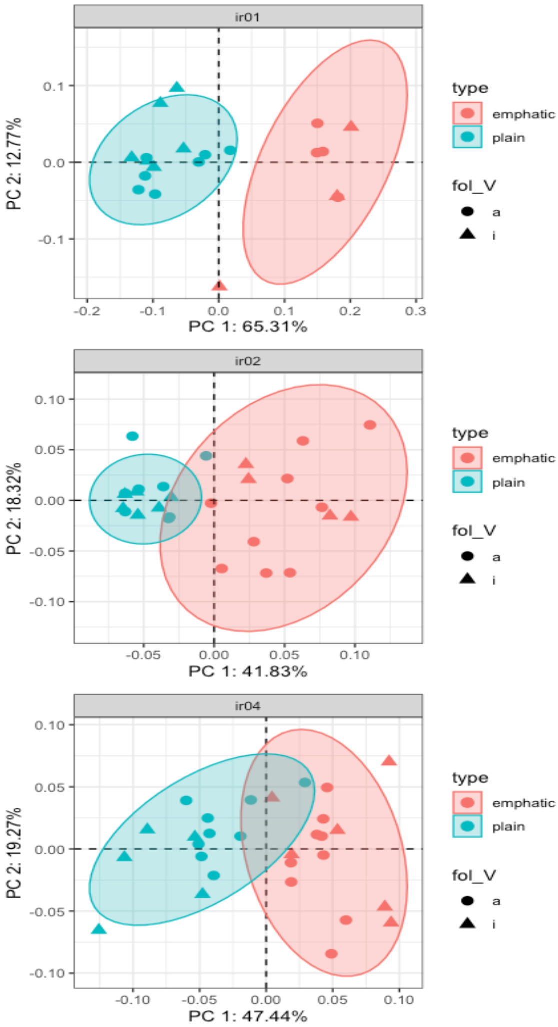


Figure 90 PC1 and PC2 scores of the xy coordinates during /sˤ/ and /s/ for Iraqi subjects

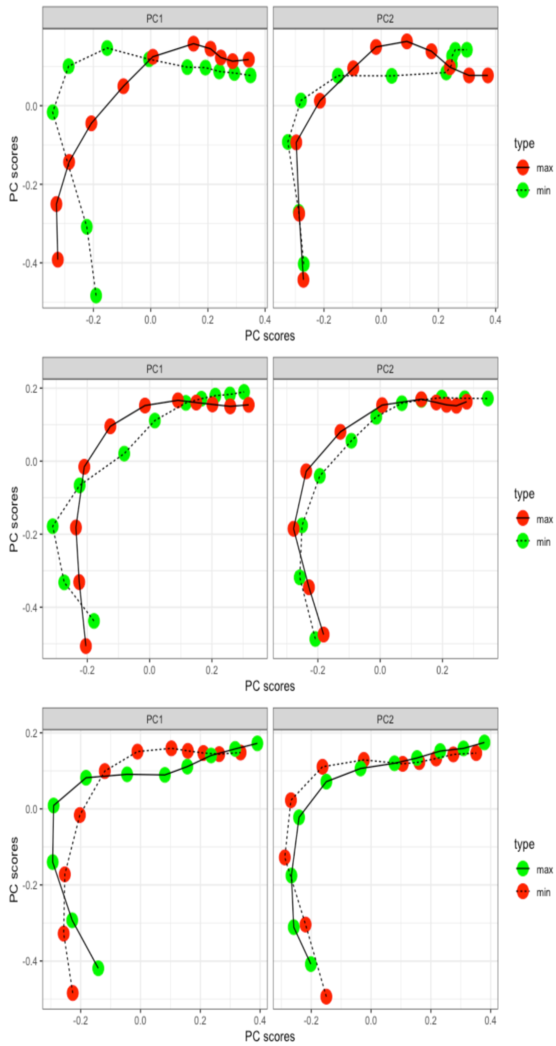


Figure 89 PC1 and PC2 tongue tracing for Iraqi subjects

5.4.1.5 Egyptian Subjects

In Figures 91 and 92, it is evident that subjects vary in terms of their PC1 scores. Subject eg01 demonstrates a partial separation between the tongue shapes for emphatic and non-emphatic sounds, with a slight overlap. The PC1 scores for emphatic instances are positive compared to those for plain instances, capturing about 70% of the variation, which is significantly higher than that for PC2 (12%). This is clearly illustrated in Figure 92, where a higher PC1 clearly corresponds to a retraction of the tongue dorsum. However, subject eg02 exhibits a distinct separation between the two tongue shapes, as shown in Figure 91. This suggests that the emphatic sounds have higher PC1 scores than the non-emphatic sounds, indicating that the tongue dorsum is retracted toward the root position for emphatics. In contrast, PC2 accounts for only 16% of tongue shape variation, as illustrated in Figure 91, where no observable tongue shape differences are evident.

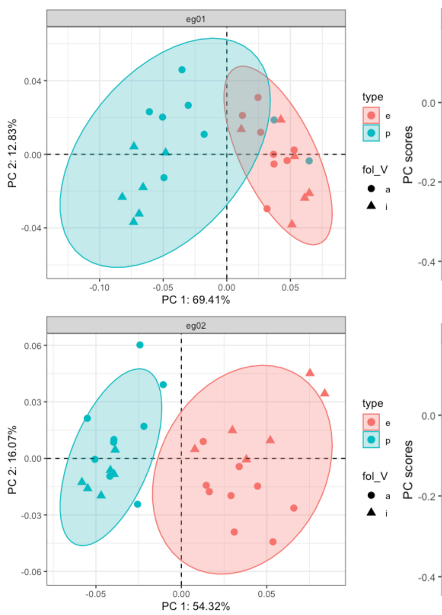


Figure 91 PC1 and PC2 scores of the xy coordinates during /sˤ/ and /s/ for Egyptian subjects

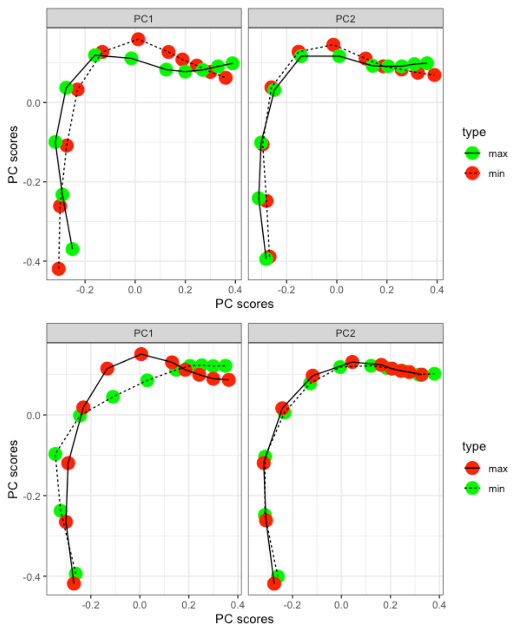


Figure 92 PC1 and PC2 tongue tracing for Egyptian subjects

The linear regression model (lm) was used to assess the relationship between fricative type (plain vs. emphatic) and PC1, which captures tongue shape variation. Specifically, the model compares PC1 scores for plain and emphatic fricatives to determine if there are significant differences in tongue shape. The Estimates reflect the magnitude and direction of these differences, with negative estimates indicating that plain fricatives have lower PC1 scores compared to emphatic fricatives. As shown in Table 44, plain fricatives generally have significantly lower PC1 scores than emphatic fricatives, indicating distinct tongue shapes. Notably, the difference in PC1 scores is more pronounced for subject eg02 compared to eg01, suggesting that tongue shape variations between emphatic and non-emphatic sounds are more strongly differentiated for eg02.

subject	Estimate	StdError	T Value	PValue
eg01	**-0.05142**	**0.013657**	**-3.765**	**0.001**
eg02	**0.067859**	**0.008676**	**7.821**	**0.0001**

Table 44 Linear regression results for PC1 scores for Egyptian subjects

5.4.1.6 Moroccan Subject

Figure 93 illustrates that PC1 captures 61% of the tongue shape variation for this speaker's articulation. There is clear separation between the two categories, with emphatic sounds exhibiting higher PC1 values and plain sounds exhibiting lower PC1 values. Notably, PC2 has minimal influence, with capturing only 13% of the tongue shape variation. Figure 94 further supports the observation of tongue retraction in emphatic sounds, suggesting that the tongue is retracted to the root position. Additionally, the linear regression results indicate a significantly lower PC1 scores for plain sounds compared to emphatic ones ($\beta = -0.081$, SE $= 0.008$, $p < 0.001$).

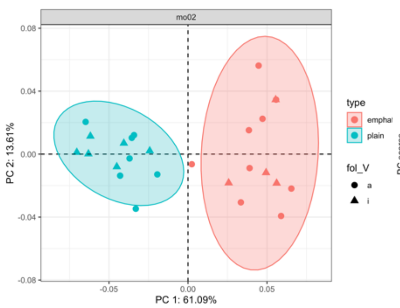


Figure 93 PC1 and PC2 scores of the xy coordinates during /sˤ/ and /sˤ/ for Moroccan subjects

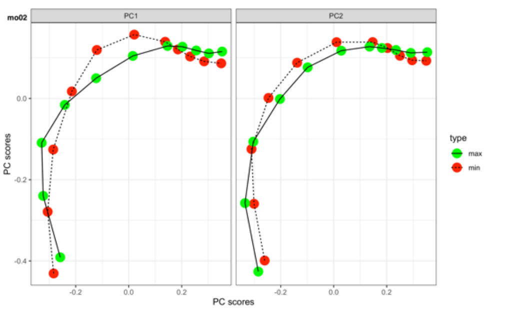


Figure 94 PC1 and PC2 tongue tracing for Moroccan subjects

5.4.2 Inspection of Acoustic Results

Having stated that all speakers show some articulatory differentiation between the tongue shapes for plain /s/ and emphatic /sˤ/, the question remains as to whether this is represented in acoustic differentiation in the speech signal. It is observed from the Figure 95 below that the COG values of the non-emphatic /s/ are generally higher than for emphatic /sˤ/, in almost all the speakers. Although the differences between the emphatic contrast are not consistent, subjects 'eg02' and 'hh02' show relatively bigger differences in the value of COG, where non-emphatic /s/ was considerably higher than it is with the emphatic /sˤ/. Also, subjects 'ir02', 'sy05', 'mo02', and 'hh01' show a smaller, though still noticeable, difference between /s/ and /sˤ/, while the rest of the subjects show only minimal differences between the two categories. Also, subject 'ir01' reveals that the emphatic /sˤ/ is produced with a higher COG than the non-emphatic /s/. Table 45 below presents the absolute difference in COG values from highest to lowest, indicating the differences between the /sˤ/ and /s/ for each speaker, categorised into High, Medium, and Low to illustrate the range of differences observed; the categorisation of the differences into "High," "Medium," and "Low" was based on the distribution of the absolute differences in mean COG values between the /sˤ/ and /s/ for each speaker. Specifically, a quantiles method was used to divide the data into three equal parts:

- High: Represents the top one-third of differences. These are the largest differences in COG values between /sˤ/ and /s/ among the speakers, indicating the most significant variability.
- Medium: Represents the middle one-third of differences. These differences are significant but not as pronounced as those in the "High" category.
- Low: Represents the bottom one-third of differences. These are the smallest differences in COG values, indicating the least variability between /sˤ/ and /s/ segments.

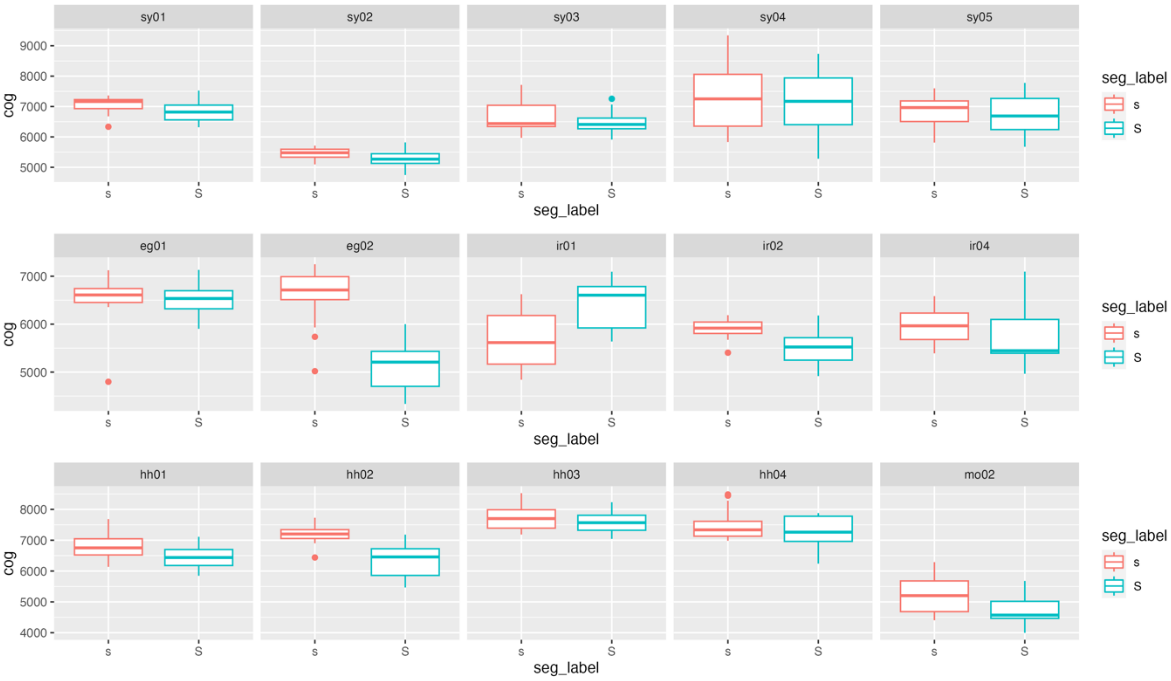


Figure 95 The distribution of COG value for fricatives /sˤ/ and /s/ for all subjects

Speaker	COG for 'S'	COG for 's'	Absolute Difference	Category
eg02	4835.007	5842.057	1007.050	High
hh02	5610.745	6255.620	644.875	High
ir02	4839.011	5254.758	415.747	High
sy05	5388.104	5780.000	391.896	High
hh01	5931.864	6267.375	335.511	High
hh03	6663.493	6998.701	335.208	Medium
sy01	6237.065	6541.061	303.996	Medium
ir04	5138.676	5360.883	222.207	Medium
eg01	5903.859	6099.394	195.535	Medium
ir01	5288.742	5102.990	185.752	Medium
sy02	4411.926	4588.537	176.611	Low
sy04	6331.644	6488.422	156.778	Low
mo02	4587.476	4725.480	138.004	Low
sy03	5757.263	5835.571	78.308	Low
hh04	6391.341	6350.658	40.683	Low

Table 45 Absolute difference in COG values from highest to lowest across subjects

While the table above suggests larger COG value differences for the first five subjects, the linear regression model reveals statistically significant differences in COG values between the emphatic contrast only for subjects eg02, hh02, sy02, and ir02. Another thing to point out is that subject sy02 was categorised as s ‘low’ in the earlier table and the fact it still comes out as significant suggests that, while the magnitude of the difference is small, the speaker is very consistent within each of the plain/emphatic categories, i.e. low standard deviation. Complete results for all subjects are presented in Table 46, which has been reorganised with the highest p-value listed first.

subject	Estimate	StdError	PValue
eg02	1260.858	253.94	0.001
hh02	737.777	208.854	0.002
sy02	266.193	105.067	0.018
ir02	395.504	160.927	0.022
mo02	547.305	281.91	0.065
hh04	360.317	252.703	0.166
sy01	208.128	150.189	0.178
ir01	-480.636	398.219	0.245
sy03	255.058	226.4	0.271
ir04	247.193	238.651	0.31
eg01	-164.619	206.925	0.433
hh03	-45.83	150.404	0.763
sy05	88.744	299.707	0.77
sy04	92.373	523.475	0.861
hh01	5.817	207.819	0.978

Table 46 Linear regression results for the COG values for all subjects

5.4.3 Acoustic -to-Articulatory Mapping Results

As has been observed the analysis of tongue splines revealed significant differences between emphatic and non-emphatic consonants across all subjects. This finding suggests a consistent articulatory distinction for these two fricative types. However, the acoustic analysis using Centre of Gravity (COG) values only identified statistically significant differences for four subjects (eg02, hh02, sy02, and ir02. Interestingly, these four subjects exhibit significant differences in both, their articulation and acoustics results between emphatic and non-emphatic sounds (refer to Table 47 for details). The remaining subjects displayed significant articulatory differences for the emphatic contrast, but these differences were not reflected in their acoustic COG values.

speaker	Acoustics	Articulatory	Potential mapping
eg01	0.433	0.001	NO
eg02	0.0001	0.001	YES
hh01	0.978	0.002	NO
hh02	0.002	0.001	YES
hh03	0.763	0.001	NO
hh04	0.166	0.0079	NO
ir01	0.245	0.001	NO
ir02	0.022	0.001	YES
ir04	0.31	0.003	NO
mo02	0.065	0.001	NO
sy01	0.178	0.001	NO
sy02	0.018	0.002	YES
sy03	0.271	0.001	NO
sy04	0.861	0.001	NO
sy05	0.77	0.003	NO

Table 47 Linear regression results for COG and PC1 scores for all subjects

This absence of a direct one-to-one acoustic-articulatory mapping might be attributable to the role of lip movements, potentially interfering with what could have been a more straightforward mapping between lingual articulation and acoustic output. The following section explores the lip states to examine if they could explain the observed disconnect between acoustics and articulation.

5.4.4 **Lip Video Results**

It is worth noting that an R script was created and utilised to extract frames corresponding to the articulation of target words. Specifically, the frame that includes the sounds /s/ and /sˤ/ was captured at the midpoint of the fricative. As shown in Figure 96, only observable differences between emphatic articulation and lip movements were observed in subjects eg02, hh01, and hh02. These subjects displayed slight lip rounding during the articulation of the emphatic /sˤ/, while their lips were unrounded for the plain /s/. This was particularly evident in eg02 and hh01. For subject hh02, the lip was spread during the articulation of the plain /s/ and slightly rounded during the emphatic /sˤ/. For the remaining subjects, no observable associations between lip and tongue shapes were found.

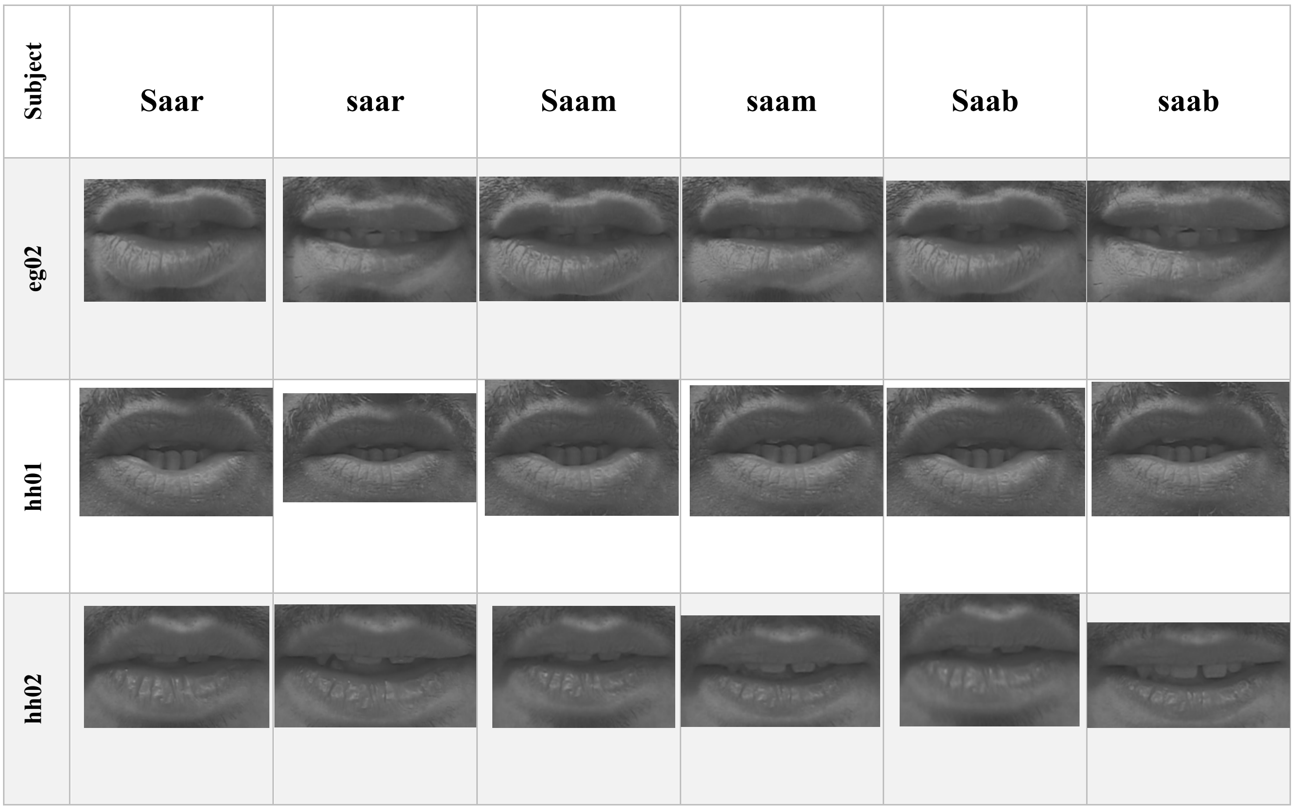


Figure 96 Pictures of lip associations to the articulation of the emphatic and non-emphatic fricatives

To draw some comparison, refer to Figure 97, which demonstrates that some subjects, such as ir01, sy02, sy03, and sy05, do not demonstrate any lip association with the articulation of emphatic contrasts; the lip state remains nearly identical for both emphatic and non-emphatic sounds. All the target lip picture frames for all subjects are presented in Appendix 11.5.

Overall, the lip analysis did not fully explain the disconnect between acoustic and articulatory mapping, nor did it explain the association of lip rounding with the articulation of emphatic sounds for some subjects. However, it did: 1) reflect on the results for the subjects hh02, and eg02, who showed significant differences in both acoustics and articulation, while also demonstrating lip rounding for the emphatic sounds, and 2) provide an explanation for the lack of mapping, for example, in subject hh01, who exhibited significant articulatory differences but not a significant acoustic differences, yet showed lip association with the articulation of emphatic sounds.

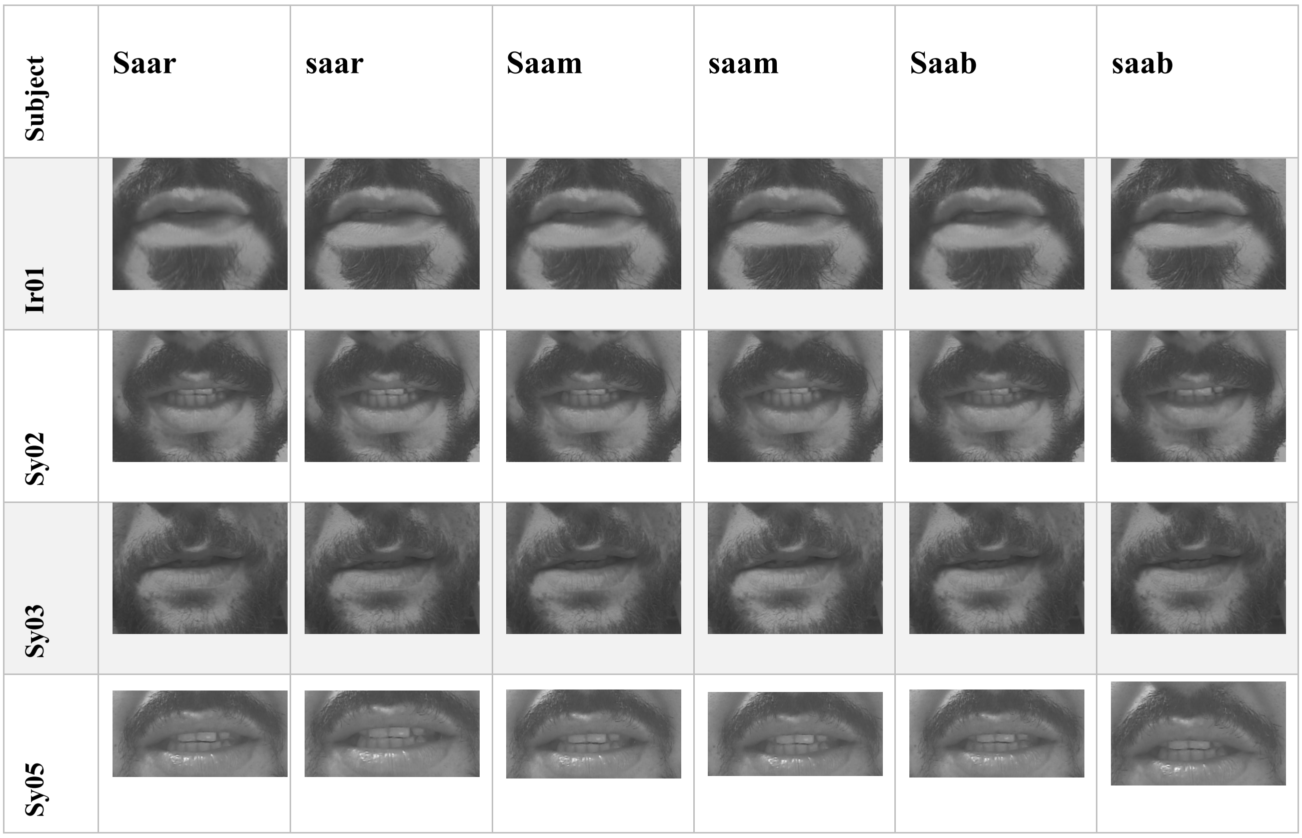


Figure 97 Pictures of lip lack of associations to the articulation of the emphatic and non-emphatic fricatives

5.4.5 **Summary and Interim Discussion**

5.4.5.1 Lingual Results

Taken collectively, we have seen that Principal Component Analysis has revealed significant differences between the articulation of emphatic /sˤ/ and its plain counterpart /s/. In other words, tongue shape variations were the primary key in distinguishing the emphatic contrast; as previously hypothesised, the tongue dorsum is retracted during the articulation of the emphatics and the surface of the tongue blade was seen depressed across most of the subjects.

This was expected since the secondary articulation in emphasis includes pulling the tongue into the upper oropharynx area (Ali & Daniloff 1972; Ghazeli 1977); this aligns with previous research, including the interpretations by Ibn Sina (Avicenna), who died in 1037 A.D. (as cited by Semaan in 1963), as well as studies by Ali & Daniloff (1972) on Iraqi Arabic, Ghazeli (1977) on Tunisian Arabic, and Al-Tamimi & Heselwood (2011) on Jordanian Arabic. Figure 98 below, adapted from Ghazeli (1977), distinctly illustrates that during the articulation of the emphatic /sˤ/, there is noticeable tongue depression, a feature that is absent in the articulation of the non-emphatic /s/.

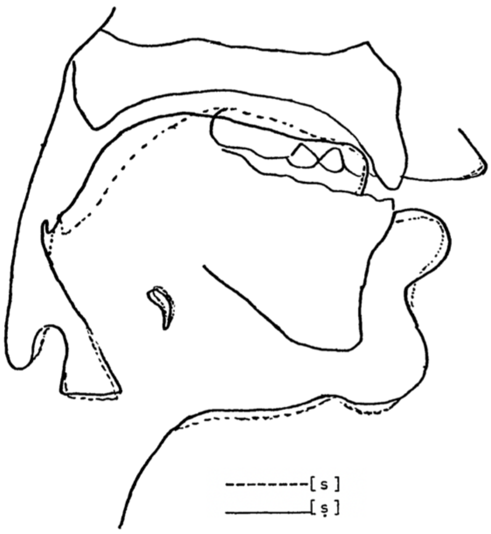


Figure 98 Tongue positioning for the emphatic /sˤ/ versus the non-emphatic /sˤ/ in the terms "Saamla" and "saamla," based on Ghazeli (1977, p. 70).

Despite variability among subjects in the extent of tongue retraction, the observed differences in tongue shape proved to be a consistent indicator of the emphatic contrast across all speakers. This variability in retraction and directionality of the back of the tongue aligns with previous research acknowledging the lack of consensus regarding the precise nature of the secondary constriction (Al-Solami, 2017). Such inconsistency can be attributed to various factors, including cross-dialectal variations and individual differences in articulatory strategies employed by speakers when producing emphatic sounds (Aldamen, 2013; Al-Solami, 2017; Khattab et al., 2006). The findings suggest that during emphasis, the movement of the tongue towards the oropharynx region is primarily horizontal. PCA tracing scores reveal that in the Hijazi dialect, two subjects (hh02 and hh03) demonstrate effects on the tongue root for the former and tongue dorsum for the latter, while the remaining subjects exhibit only tongue body effects, with no active retraction of the tongue root observed. This inconsistency is in line with expectations, as tongue root retraction typically follows the retraction of the dorsum (Al-Solamy, 2017). This suggests that if the tongue root is actively retracted, it is likely due to a more pronounced retraction of the tongue dorsum, which influences the root to retract as well. Additionally, Shar (2012) noted in her study on Saudi speakers that the tongue root does not play an active role in the articulation of emphatic sounds. Based on the findings, the suggestion that the Hijazi dialect should be classified as pharyngealized, as proposed by Alfaifi et al. (2020), is subject to debate. Regarding the roles of the pharynx and tongue root in producing emphatic sounds, evidence from some subjects suggests they are not actively involved (Norlin, 1987). Instead, it is the retraction or backing movement of the tongue dorsum that generates the necessary constriction for emphatic articulation. Thus, for some subjects, the retraction of the tongue root in emphatic sounds appears to be a secondary effect of the overall retraction of the tongue dorsum, rather than an independent articulatory gesture. However, evidence from other subjects indicates that the pharynx and tongue root do play a role in producing emphatic sounds, as noted by Al-Tamimi and Heselwood (2011, p. 174). The varying results across studies align with the current finding of speaker variation in tongue positions.

The obtained results suggest that dialects exhibit variations in both the degree and direction of tongue retraction. Examination of tracing scores in the Iraqi dialect reveals that the tongue is retracted toward the root position, with the exception of ir01, which exhibits tongue dorsum retraction, consistent with findings from Ali and Daniloff (1972). Similarly, in the Syrian dialect, all subjects showed tongue root retraction, except for sy01, sy05, and sy04, who exhibited dorsum retraction. Among all Moroccan, Egyptian, and Hijazi subjects, tongue root retraction was observed, except for subject hh03, who exhibited tongue dorsum retraction. Consistent with prior observations (Zeroual et al., 2011), in the Egyptian and Moroccan dialects, the back of the tongue was observed retracting towards the posterior wall of the pharynx without significant elevation.

Thus, the answer to the research question, 'Is there any inter-dialectal variation in how emphatic contrast is articulated?' is affirmative, based on the overall results. This conclusion is premised on the observation that within each dialect, some subjects exhibit tongue root retraction while others show tongue dorsum retraction. This suggests that the variation in articulating emphatic contrast is rather characterised by individual-specific patterns. Therefore, this finding addresses the other research question: 'Are there individual-specific patterns in articulatory mechanisms?

5.4.5.2 Acoustic-to-Articulatory Results (COG Value for Fricative)

When utilising the acoustic correlate COG to examine the differences in the emphatic contrast, it was observed that consistency was lacking across dialects and subjects; only four subjects demonstrated a significant acoustic distinction between the emphatic contrast /s/ versus /sˤ/. Specifically, subjects eg02, hh02, sy02, and ir02 exhibited significant differences in both their acoustic and articulatory results for the /s/ versus /sˤ/ contrast. Despite the articulatory analysis proving valuable for distinguishing between emphatic and non-emphatic sounds, the acoustic results did not consistently reflect these distinctions. Additionally, it is noteworthy that the analysis of the relationship between the regression model estimates for acoustics and the articulatory results yielded a crucial finding. Contrary to the initial assumption that stronger acoustic estimates would correspond with stronger articulatory estimates, indicative of a one-to-one relationship, no such correlation was observed, as illustrated in Figure 99 below. For instance, subject eg02 displayed the largest acoustic difference among all subjects in terms of the estimate of the regression model. However, in terms of their articulation estimate plotted on the Y-axis, eg02 was positioned in the middle relative to other subjects. Conversely, for the Hijazi subject hh03, the acoustic estimate was nearly zero, yet hh03 demonstrated one of the most pronounced articulatory differences among the subjects. Referring to PCA Figure 85, it was evident that the tongue splines for hh03 were distinctly different, albeit with minimal acoustic difference. This acoustic difference was significantly lower than that observed in other subjects, such as eg02. Similarly, subject ir01 exhibited strong estimates for lingual articulation but lower acoustic estimates. This disparity indicated a lack of correlation between the two sets of estimates for these subjects. This observation also extended to the remaining subjects, whose estimates for both lingual articulation and acoustics did not show any correlation with one another.

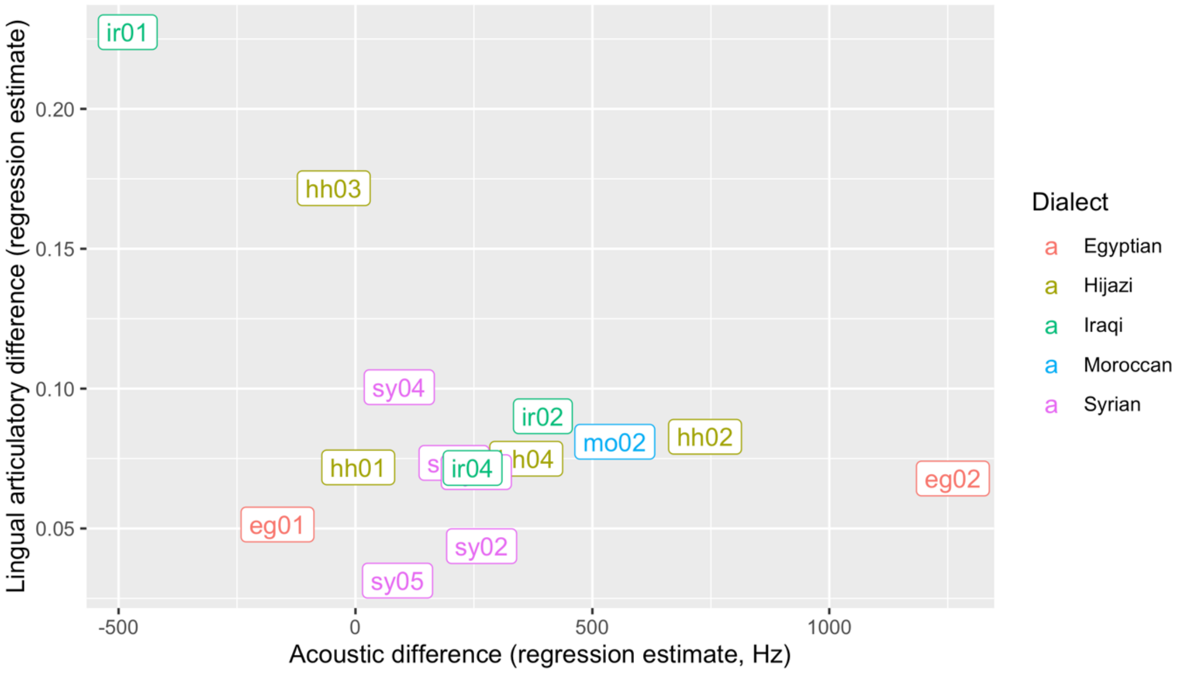


Figure 99 The regression estimates for acoustics and articulatory results across all subjects

Thus, the answer for the research question: ‘Do the results of the articulatory study map onto the acoustic study?’ is that the articulatory results do not completely map onto the acoustic findings. This discrepancy can be attributed firstly to the challenge of recovering certain aspects of speech production solely through acoustic analysis, and secondly, to the advantages offered by articulatory instruments like Ultrasound Imaging, which can uncover covert features of speech articulation (Lawson et al., 2014). Covert articulatory variation, as described by Mielke et al. (2017), occurs when speakers employ different articulatory strategies to produce sounds that are acoustically or auditorily indistinguishable. An illustrative example of this discrepancy is provided by Strycharczuk and Scobbie (2017) in their study on the fronting of the high/back vowels /u/ and /ʊ/ in Southern British English dialect. Despite clear differences in tongue position for /u/ and /ʊ/, they found no significant corresponding difference in F2, indicating that the lack of mapping between acoustic and articulatory data may be reasonable.

5.4.5.3 Labial Results

Regarding the relationship between lip movement and tongue shape during the articulation of emphatic and non-emphatic sounds, a hypothesis proposed that lips would adopt a rounded configuration for emphatic sounds and remain unrounded or spread for non-emphatic sounds. This hypothesis was based on previous research findings, which observed some degree of lip protrusion associated with emphatic sounds (El-Halees, 1985; Hetzron, 2013; Jakobson, 1957). Additionally, Lehn (1963) noted slight lip rounding in the Carine dialect, while Bellem & Watson (2014) identified labialisation in Ṣanʿāni Arabic and Southern (gilit) Iraqi Arabic.

However, only a subset of subjects, such as eg02, hh01, and hh02 (see Figure 96), exhibited minimal lip rounding for the emphatic /sˤ/, while their lips remained unrounded for the plain /s/. This pattern was particularly pronounced in eg02 and hh01. Notably, in the case of hh02, lips were spread for the plain /s/ and displayed slight rounding for the emphatic /sˤ/. This observation raises important questions: Does the occurrence of lip rounding during the emphatic contrast, especially among these three subjects, imply that they manifest more significant lingual differences compared to other subjects? Alternatively, could it indicate that a subject with notably higher lingual PC1 compared to another exhibits more pronounced lip rounding? Exploration of these questions reveals insights: Among Hijazi speakers, subject hh02 exhibited slight lip rounding, despite displaying significantly high PC1 results, indicating a distinct separation between emphatic and non-emphatic sounds. In contrast, subject hh01, who showed evident lip rounding, did not exhibit a greater PC1 difference compared to hh02 (refer to Table 41), and lingual results showed some overlap between emphatic and non-emphatic sounds (see Figure 85). Consequently, according to this hypothesis, one might expect that Hijazi subject hh03, registering the highest PC1 scores among all Hijazi subjects, would exhibit clear lip rounding. However, subject lip state appeared unrounded for both types of fricatives, emphatic and non-emphatic. Similarly, Syrian subject sy04 and Moroccan subject mo02, recording the highest PC1 among all dialects ($\beta = -0.10051$, $SE = 0.008606$, $p < 0.001$, and $\beta = -0.081$, $SE = 0.008$, $p < 0.001$, respectively), also showed a lack of lip rounding during the articulation of emphatic fricative sounds. Therefore, the hypothesis suggesting that the articulation of emphatic fricatives necessarily entails lip rounding does not hold as a valid prediction. This overarching finding addresses the research question, "Does lip rounding contribute to distinguishing between emphatic contrasts among Arabic dialects?" It appears that lip rounding does not significantly contribute to distinguishing emphatic contrasts, which may account for the individual variability observed among subjects.

5.5 Conclusion

In summary, this study utilised ultrasound imaging to analyse the articulatory patterns of the emphatic /sˤ/ in comparison to its plain counterpart /s/ across speakers representing five distinct Arabic dialects. The goal was to uncover potential differences in fricative articulation among these dialectal variations. Through an acoustic analysis focusing on the Centre of Gravity (COG) for different fricative types, the study explored the relationship between acoustic properties and articulatory patterns. The findings revealed that articulatory examination effectively distinguished differences in the emphatic contrast, showcasing variability in tongue shapes. Specifically, emphatic sounds demonstrated tongue body retraction, albeit with variations in how this retraction was manifested; some subjects retracted towards the root positions while others towards the tongue dorsum. However, acoustic analysis showed inconsistency in COG results among subjects, with only a few displaying significant differences in the emphatic contrast. As a result, no clear alignment was observed between acoustic and articulatory findings, as indicated by the regression model estimates for both results. Furthermore, the investigation into the association of lip rounding with emphatic fricative articulation found no significant correlations, contradicting the hypothesis suggesting a link between emphatic articulation and lip rounding. These findings collectively indicate that examining lingual articulation provides deeper insights than acoustic analysis alone, and the observed variation is more attributed to individual-specific patterns rather than inter-dialectal differences.

6 Chapter 6: General Discussion

6.1 Chapter Outline

This chapter will begin with an overall summary of the results obtained from each conducted study: production, perception, and articulatory. This will be followed by a discussion of the results, hypotheses and research questions for each chapter. A reflection on the existing literature will also be included, linking it with the findings obtained, to highlight the contributions of the current thesis. Finally, a discussion of the mapping between articulatory and acoustics, will be presented.

6.2 Results Summary

6.2.1 Acoustic Study Results

This study investigated the realisation of the emphatic /sˤ/ and its plain counterpart /s/ across eight Arabic dialects. The aim was to investigate whether the realisation of this emphatic contrast the same across these Arabic dialects and to see whether we can find a similar typological divide as was found with the plosive consonants; it was found that there is a typological divide which classified Arabic dialects into two groups: some dialects show an acoustic difference in the plosive (VOT) as well as in the following vowel, while other dialects only show a difference in the following vowel (see Bellem, 2014). To achieve this aim, multiple acoustic correlates were utilised. For fricative, the duration of the frication, the intensity, the spectral properties which includes Centre of Gravity COG, and spectral peak location were analysed. Additionally, for vowels, the first three formants (F1, F2, and F3) were examined. The findings suggested that most of the variation in this dataset is due to main effects between dialects, genders, and vowel type. The acoustic correlates that accounted for fricative contrast s~sˤ were all came out not significant, except for the second and third formants measures (F2 and F3) for the following vowel, which were consistent acoustic correlates found in this cross-dialectal dataset for the emphatic contrast. Given that the acoustic correlates of fricative consonants do not significantly contribute to identifying the distinction for the emphatic contrast, the examination of formant information for the vowels following the emphatic contrast was conducted. The results from Generalized Additive Mixed Models (GAMMs) revealed that the majority of the emphatic contrast was found in the entire trajectory of the vowel following the s~sˤ contrast, with dialects varying in the magnitude of the F2 difference and the extent to which the vowel is affected. while dialects exhibit variations in the emphatic contrast, these differences are specific to dialect and gender. Moreover, the findings did not reveal a similar typology as observed in emphatic plosives (Bellem, 2014).

6.2.2 **Perception Study Results**

Since the findings of the acoustic study revealed that the second formant (F2) of the vowels following the fricative consonants was one of the only acoustic correlates that can play a significant role in identifying the differences between the emphatic /sˤ/ and its plain counterpart /s/ across all the eight Arabic dialects the perception study was designed and structured to know to what extend these Arabic participants can recognise the F2 signals of their own and other dialects. Also, to see if there is any mapping between the results of the production and the perceptions. Thus, the F2 signal was resynthesised and manipulated, and the key findings suggested that, in real stimuli, all listeners from each dialect could identify their own native formant cues with higher accuracy in plain condition. Additionally, all listeners can attend to the cues to emphasis of other dialects with considerable accuracy. However, for manipulated stimuli, the F2 size and trajectory shape was observed to play a significant role for listeners to be able to identify their own cues, in emphatic condition (only). Having mentioned that, the binomial test results revealed that for the real stimuli, the cues distinguishing plain and emphatic realizations are more distinct for /i/ than for /a/, as evidenced by the consistent recognition of /i/ vowels across all dialects. Moreover, listeners generally demonstrated the ability to perceive the plain-emphatic distinction in both their own and other dialects, with the notable exception of Moroccan /a/ vowels. Specifically, Moroccan plain and emphatic /a/ vowels were found to be challenging to recognize for both non-Moroccan listeners and Moroccan listeners themselves. The results of the manipulated stimuli indicated that only Syrian listeners could attend to their F2 signals with statistically significant accuracy. In contrast, listeners from other dialects, including Iraqi and Moroccan, demonstrated above-chance accuracy in attending to their F2 cues, but this was not statistically significant. Thus, this inconsistency between real and manipulated stimuli suggests that F2 may not be the sole cue relied upon by listeners

6.2.3 **Articulatory Study Results**

This study was motivated by the findings of the perception study and was designed to explore whether articulatory differences in tongue shape, as well as lip associations, can better describe the distinction between emphatic and non-emphatic fricative consonants.

Ultrasound Imaging analysis was conducted, involving fifteen speakers from five different Arabic dialects. Additionally, acoustic analysis was carried out to examine the Centre of Gravity (COG) values for the speakers during the articulation of the emphatic /sˤ/ and the non-emphatic /s/.

The findings suggested several points:

1. Tongue shape variations were key in distinguishing the emphatic contrast. Specifically, the tongue dorsum was retracted during the articulation of the emphatic consonants, and the surface of the tongue blade was observed to be depressed across most subjects, unlike the non-emphatic consonants where the tongue was advanced and raised.

2. Acoustic analysis revealed inconsistency in the COG results among subjects, with very few showing a significant difference in the emphatic contrast. Consequently, no clear mapping was found between the acoustic and articulatory results, as indicated by the regression model estimates for acoustics and articulation.

3. The labial results did not play a significant role in connecting the associations between the articulation of the emphatic consonants and lip rounding.

4. Subjects differed in the degree of tongue retraction and the directionality of movement. These variations were attributed to individual behaviour, with no observation of inter-dialectal variations.

6.3 Discussion and interpretation of the results

6.3.1 Acoustic Study

As observed, the findings from the acoustic chapter clearly demonstrated that the formant information of the vowels following the emphatic contrast could significantly distinguish between the emphatic and the plain fricatives[11]. This aligns with previous literature suggesting that the F2 frequency of the vowel following the emphatic consonants tends to decrease, while it increases in the non-emphatic consonants. Thus, the effect of F2 has become a very salient feature that marks emphasis. It was predicted from the acoustic chapter that F2 would behave in this way.

[11] Huge body of previous research have reported this effect of F2 of the following vowels in emphatic contexts like: (Zawaydeh, 1999;Card, 1983; Al-Tamimi and Heselwood, 2011; Alioua, 1995; Yeou, 1997)

However, an interesting point to note is that previous studies have typically utilized only one or two Arabic dialects to examine these phenomena, resulting in varied findings, particularly regarding other formants, such as F3. Some studies have reported high frequency of F3 (e.g., Kriba, 2010; Yeou, 2001; Jongman et al., 2007), while others have found lower F3 frequencies (e.g., Card, 1983; Norlin, 1987). For a detailed discussion, refer to McCarthy (1994) and Watson (2007).

The current study provides insightful results as it incorporates eight Arabic dialects to investigate whether the emphatic contrast realizations are consistent across these dialects. Interestingly, one of the findings is that F3 was observed to significantly differentiate between the emphatic and non-emphatic consonants. In other words, these eight Arabic dialects have shown a consistent F3 effect, indicating that they share a commonality beyond F2 lowering. This suggests that the location of the pharyngeal constriction for the emphatic consonants likely occurs in the lower pharynx and is relatively consistent across all these dialects.

As Kent and Read (1992) claimed, there is a correlation between the location of the pharyngeal constriction and the increase in F3. This might lead one to argue that the inconsistencies in previous research regarding F3 results imply variations in the location of the pharyngeal constriction, ranging from mid (minimal effect of F3) to high (no effect of F3) (Kent and Read, 1992). However, the current study shows consistent values of F3, even for the Egyptian speakers, contrary to Norlin (1997), who reported that F3 was low and had no effect in distinguishing between the emphatic contrast.

An important point to highlight is that while the majority of previous research on sound variation within Arabic dialects has claimed that the effect of emphasis is related to the vowels following the consonants, and particularly that F2 was observed to distinguish between the emphatic contrast, only very few studies have utilized a dynamic approach to study the vowel trajectory. Studies such as those by Al-Tamimi (2007), Almurashi et al. (2024), and Sakr (2024) have limited themselves to examining one or two Arabic dialects. Additionally, the number of trajectory points explored in these studies varied between two to seven, with Sakr (2024) aiming for 11 points. Al-Tamimi (2007) and Almurashi et al. (2024) focused on confirming the effectiveness of using the dynamic feature to distinguish and classify vowels in their examined dialects, rather than examining the context in which fricative emphasis occurs and relating it to the vowel's behaviour.

Thus, it can be argued that only one study has used the dynamic approach of vowels to account for the emphatic contrast, which is the study by Sakr (2024) that examined the Lebanese dialect. Sakr showed that the whole F2 trajectory of the vowel preceded by a plain consonant and followed by an emphatic one is consistently lower. However, a limitation of Sakr's findings is that he analyzed the emphatic fricative without attention to the contexts in which the target sounds occurred (e.g., monosyllabic versus disyllabic words, stressed versus unstressed syllables). In contrast, in this thesis, the stimuli were controlled and balanced, and the overall results revealed interesting patterns, especially with the Generalized Additive Mixed Model (GAMM) findings.

While the findings conformed with those of Sakr (2024), revealing that the majority of the emphatic contrast was found in the whole trajectory of the vowel following the emphatic fricatives, an interesting pattern was also shown among the eight Arabic dialects examined. Dialects differed in the magnitude of the F2 difference and how much of the vowel trajectory was affected. Figure 37 from the acoustic chapter (section 3.4.7) showed that, for example, Egyptian, Jordanian, and Moroccan dialects exhibited significant time-varying emphasis starting from the onset to the offset of the vowel /a:/, whereas the significance of height was observed only in the initial parts of the trajectories of dialects such as Iraqi, Kuwaiti, Omani, and Syrian. Similar observations were made for the vowel /i:/, with some dialects showing significant differences in height at different intervals.

By considering the fact that emphatic coronals have an impact on adjacent segments and that the impact would affect anything from a single syllable to the entirety of a word within Arabic dialects (Almuhaimeed, 2021), one could argue that, in light of the current findings, the emphasis may spread rightward across the vowel to the following consonants, particularly, with the vowel quality /a/ since emphatic~plain trajectories were observed to be apart across all dialects, regardless of the height. However, the same is not predicted when it comes to the vowel /i/ as the emphatic~plain trajectories tended to overlap at diffrent points throughout the trajectories. This observation aligns with the results reported by Al-Masri and Jongman (2004), who found that the vowels /i/ and /u/ block the spread of emphasis in the rightward direction.

When discussing the acoustic correlates of fricative consonants (emphatic and plain), Chapter 2 reported that very few studies have found that Peak location can distinguish between the emphatic contrast (e.g., Norlin, 1983). Conversely, locus equations were the only acoustic correlates that differentiated between the emphatic and plain consonants (e.g., Yeou, 1997). However, a vast body of other studies found that there are no effects of acoustic measures (e.g., friction duration, COG, intensity, peak location) on the emphatic contrast (see Khan, 1975; Norlin, 1987; Al-Khairy, 2005; Kriba, 2010; Sakr, 2024). Despite these varying findings, the prediction offered in the acoustic chapter was that the duration of the frication noise would be longer for the emphatic and shorter for the non-emphatics. This prediction was motivated by the principle of the source-filter theory by Fant (1960), suggesting that emphatic fricatives exhibit co-articulation, with two sources of energy: primary (turbulence occurring in the vocal tract) and secondary (extra energy in the pharyngeal area), leading to slowed airflow in the back area of the oral cavity compared to plain fricative sounds.

Although the overall findings suggested that there is no significant effect of duration across the board, evidence of dialect-specific behaviour was found. In Chapter 3, section 3.4.1.1, it was discovered that among Egyptian speakers, the duration of the emphatic fricative was significantly longer than that of the non-emphatics [ß = 7.269e-03, SE = 1.634e-03, P > 46e-06] [12]. It is interesting to note that no previous study concerning the Egyptian dialect has reached this finding.

To provide a comprehensive view of the other acoustic correlates and their effects, Table 48 summarizes the predictions for each measure and its effect across all dialects. It is worth noting that only dialects with significant effects of the target measure are included. The up-arrow symbol (↑) represents high frequency/longer duration, the down arrow symbol (↓) represents lower frequency/shorter duration, and the check mark symbol (√) indicates that the dialect effect agrees with the prediction.

[12] For example, Nolin (1987) examined Egyptian speakers and found no effects of acoustic correlates of spectral moments on the emphatic contrast distinction.

Acoustic measurement	Prediction	Dialect effect	Agree to the prediction
Duration	/sˤ/ ↑ /s/ ↓	**Egyptian** /sˤ/ ↑ /s/ ↓	√
		Iraqi /sˤ/ ↓ /s/ ↑	
COG	/sˤ/ ↓ /s/ ↑	**Egyptian** /sˤ/ ↓ /s/ ↑	√
		Moroccan /sˤ/ ↑ /s/ ↓	
Intensity	/sˤ/ ↑ /s/ ↓	**Iraqi** /sˤ/ ↓ /s/ ↑	
		Moroccan /sˤ/ ↑ /s/ ↓	√
Peak Location	/sˤ/ ↓ /s/ ↑	**Egyptian** /sˤ/ ↓ /s/ ↑	√

Table 48 A summary of the dialect effects for each of the acoustic measurements. Only dialects with significant effects are included. Symbols (↑ =high/longer), (↓ =low/shorter) and (√ = agree to the prediction)

Table 48 indicated that three out of eight dialects showed a significant use of acoustic measures, with Egyptian speakers standing out from other dialects followed by Moroccan and then Iraqi speakers, all of whom have used several correlates showing the different realisations of the fricative s~sˤ.

The key point from this section is that no acoustic correlates were consistently observed within the fricatives themselves to mark the emphatic contrast across all dialects. Although there is some evidence of dialectal variation, this variation remains limited. Therefore, it is difficult to argue that there is a typology within this variation, and this differs significantly from the typology observed in the realization of plain~emphatic plosives (see Bellem, 2014) [13]. Overall, as stated earlier, the only consistent effect observed that could mark the emphatic contrast was in the following vowels. Therefore, one could argue that emphasis in this acoustic study is not a consonantal phenomenon but rather a vocalic one.

[13] Bellem (2014) found that there is a typological divide which classified Arabic dialects into two groups: some dialects show an acoustic difference in the plosive (VOT) as well as in the following vowel, but other dialects only show a difference in the following vowel.

6.3.2 Perception Study

In an attempt to investigate the argument that emphasis is a vocalic phenomenon, the perception study in Chapter 4 was designed with the aim to ascertain whether the results of perception confirm what was found in acoustic. However, before interpreting the current findings of the perception study, it is worth discussing what previous research has presented. As stated in the background chapter (Section 2.6), Jongman et al. (2011) used a technique involving cross-spliced CVC syllables and found that syllables with emphatic CV or VC portions received significantly more emphatic responses than words with plain CV or VC portions. They assumed that perception study results somehow mapped onto acoustic's findings, which emphasized the role of emphasis effect on F2. However, this assumption may not be entirely valid since the researchers did not specifically explore the perceptual effect of the formant information of the vowel itself that occurs in the vicinity of the emphasis. Instead, they relied on entire words or syllables containing both consonants and vowels to assess listeners' judgments of emphasis identification. Therefore, claiming that the vowel plays a role in identifying the emphatic contrast in their study, without providing conclusive justification, is not entirely accurate, even though the argument itself is valid as tested in other studies, such as Obrecht (1968). Obrecht examined the effect of the second formant on the perception of emphasis and found that listeners can identify emphasis if the F2 value was below 1560Hz, indicating a categorical perception of the emphatic contrast. However, there remains the question of whether it is solely about the F2 lowering or if there are other characteristics of the F2 information that listeners may pay attention to in making their judgments. Similarly, the same questions should also be addressed in the study of Ali and Daniloff (1974); they compared minimal pair emphatic/non-emphatic words and found that listeners correctly identified the missing consonant from truncated stimuli with about 68% accuracy (73% for emphatic stems, 62% for plain stems). Their findings indicate that the distinction between the emphatic contrast can be accurately perceived without the emphatic consonant, suggesting that the vocalic portion of the plain/emphatic words comprises sufficient cues for listeners to detect the distinction. I would argue that these studies have not provided insightful information about the characteristics of the vowel formant information and what types of information, particularly in F2, listeners rely on to identify the distinction between the emphatic contrast. Furthermore, it is noteworthy that no single study has ever examined more than two dialects and utilized the F2 trajectory characteristics to investigate listeners' perception of the emphatic contrast. Intriguingly, none of the previous studies have also explored the perception of listeners when presented with unfamiliar cues from other dialects.

To address this gap, the current study has employed a novel technique that has not been utilized in any previous research. Specifically, it investigated the entire trajectory of the F2 for the vowels and resynthesized the F2 signals with different heights and slopes based on the findings from the acoustic results in Chapter 3. Consequently, four manipulated stimuli were created for the vowel /a:/, and another four were created for the vowel /i:/. A summary of all resynthesized conditions for both vowels /a:/ and /i:/ can be found in the Table 49 below (For visual inspections, refer to Appendices 10.4 to 10.7).

Vowel /a:/		
Condition	**The pattern**	**F2 manipulation**
Iraqi-like stimuli	**Low_30**	F2 trajectory is lowered by around 500 Hz from the onset and starts to rise after 30% of the vowel duration, merging to the base stimuli.
Syrian-like stimuli	**Low_100**	F2 trajectory is manipulated to start low with a 500 Hz difference (from the plain trajectory) and never becomes high; the large F2 difference is seen throughout the whole trajectory.
Moroccan-like stimuli	**Medium_100**	F2 trajectory is lowered by approximately 200 Hz compared to the plain base stimuli and maintains a medium height difference throughout the entire vowel.
Plain = Base stimuli	**---**	This condition is the manipulated plain version with values set to resemble the plain base stimulus trajectory. (Refer to Figure 29)
Vowel /i:/		
Iraqi-like stimuli	**Low_50**	F2 trajectory is manipulated to have a 500 Hz difference compared to the plain trajectory and begins to rise at around 50% through the vowel.
Syrian-like stimuli	**Medium_30**	F2 trajectory is lowered by 200 Hz compared to the plain base stimulus and merges with the base stimulus trajectory after 30% of the vowel duration.
Moroccan-like stimuli	**Medium_100**	F2 trajectory is 200 Hz lower than the base stimuli and remains at this medium height difference from the plain trajectory for the entire vowel.
Plain = Base stimuli	**---**	This condition is the manipulated plain version with values set to resemble the plain base stimulus trajectory. (Refer to Figure 29)

Table 49 The created conditions for the vowel F2 trajectory for the vowels /a:/ and /i:/

As for the results, Table 50 below presents the overall summary for each condition. Thus, in the real stimuli results, we can see that all listeners from each dialect could identify their own native formant cues with higher accuracy in both plain and emphatic conditions. Crucially, however, all listeners demonstrate a notable ability to attend to the emphasis cues of other dialects with considerable accuracy. This may be attributed to the retention of all potential phonetic cues, including full formant values: F1, F2, and F3. Also, it is important to point out that listeners often misidentify a plain /a:/ as emphatic across dialects, which may be due to variations across dialects in the size and shape of the overall vowel triangle.

			emphatic		
Condition	**Vowel**	**stimuli:**	Moroccan listeners	Syrian listeners	Iraqi listeners
real	a:	Moroccan	67%	72%	79%
		Syrian	73%	90%	100%
		Iraqi	73%	92%	98%
	i:	Moroccan	87%	90%	98%
		Syrian	93%	92%	98%
		Iraqi	73%	87%	91%
manipulated F2	a:	Moroccan-like	67%	17%	20%
		Syrian-like	27%	67%	32%
		Iraqi-like	33%	19%	56%
	i:	Moroccan-like	63%	42%	15%
		Syrian-like	30%	58%	18%
		Iraqi-like	50%	55%	59%
			plain		
Condition	**Vowel**	**stimuli:**	Moroccan listeners	Syrian listeners	Iraqi listeners
real	a:	Moroccan	67%	23%	21%
		Syrian	87%	82%	60%
		Iraqi	40%	79%	77%
	i:	Moroccan	93%	92%	96%
		Syrian	100%	96%	96%
		Iraqi	93%	82%	94%
manipulated F2	a:	Moroccan-like	77%	72%	83%
		Syrian-like	77%	88%	81%
		Iraqi-like	67%	72%	77%
	i:	Moroccan-like	97%	88%	93%
		Syrian-like	90%	90%	91%
		Iraqi-like	87%	87%	95%

Table 50 The overall summary of the perception findings; the correct responses for listeners to the cues presented for real and manipulated stimuli

For the manipulated stimuli, the findings indicated that all dialects were better at attending to their own emphatic F2 trajectory cues (only). In other words, the size and shape of the F2 trajectory played a significant role for listeners to identify their own cues in the emphatic condition. For the vowel /a:/, if the size effect of the F2 trajectory was medium throughout, Moroccan listeners tended to identify it with greater accuracy. However, if the effect size of F2 was large throughout, Syrian listeners were the only ones who could perceive it at above chance levels and significantly better than listeners from other dialects. Similarly, if the size effect of F2 was large only in the first 30% of the trajectory, Iraqi listeners found it easier to recognize than other listeners did. As for the vowel /i/, Moroccan listeners were seen to attend accurately to the small F2 size effect throughout. Conversely, if the F2 effect size was small only in the first 30%, Syrian participants tended to identify it more effectively than participants from other dialects. The only exception to the above patterns was that Iraqi listeners were no better than listeners in other groups at identifying the /i/ vowel if the size effect of F2 was 'Iraqi-like', being large only in the first 50% of the vowel.

For the plain stimulus, listeners were all roughly similar at perceiving the absence of emphatic cues, likely because the manipulated plain stimulus was set to resemble the plain base stimulus trajectory. However, when an emphatic cue is present and not that the one they are used to in their own dialects, they do not perform well at identifying it, often perceiving it as plain instead. This may be attributed to the extent of exposure listeners have, to cues other than those in their own dialects. As reported in section 4.5.6.4, the exposure factor does play a significant role in predicting listeners' accuracy in identifying their own native cues (ß 0.026723, SE = 0.001234, $p < 2e-16$).

When comparing these findings with the binomial test results, it becomes evident that for the real stimuli, Moroccan listeners were the only group who experienced difficulties identifying their own F2 cues for the vowels /i/ and /a/ in both conditions. In contrast, for the manipulated stimuli, Syrian listeners were the only group who could significantly attend to their emphatic cues for both vowels, while the remaining listeners performed at chance level when identifying their own cues. Finally, Moroccan listeners consistently performed at chance level in identifying their own cues, whereas listeners from other groups performed significantly above chance level.

Thus, the answer to the research question, " Are listeners more 'accurate' at identifying emphatic stimuli when given the cues to emphasis in their own dialect versus other dialects, when presented with stimuli in which only F2 is manipulated?" is that listeners are indeed more 'accurate' at identifying emphatic stimuli when given the cues to emphasis in their own dialect as opposed to other dialects. Additionally, the hypothesis that listeners from each dialect group would better identify the emphatic contrast when presented with F2 trajectory cues from their own dialect rather than those from another dialect was accepted, as the overall results supported this conclusion.

However, the answer to the other research question, "Are listeners more 'accurate' at identifying emphatic words when given the cues to emphasis in their own dialect versus other dialects, when presented with real stimuli?" is that listeners were all proficient with identifying emphatics; the differences lie only in the relative accuracy of identifying the plain /a:/. This suggests that there is something related to the overall shape and size of the vowel trajectory across dialects, rather than cues to emphasis. Additionally, the prediction was that listeners of such a dialect are better in attending to their own vowels and can easily identify the cues of their own trajectory types. However, because there was no consistency in the results, this prediction was rejected. Thus, this inconsistency between real and manipulated stimuli implies that F2 is not the only cue that listeners are picking up on in the real stimuli.

6.3.3 Articulatory Study

As discussed previously, the acoustic findings suggested that vowel formant information, specifically F2 trajectory, plays a major role in differentiating the emphatic contrast. The perception study has partially confirmed that (since the effect of F2 was only observed within the manipulated stimuli of emphatics). So, both studies have emphasized the vocalic impact and found no effect for the consonant. Therefore, the articulatory study was conducted to investigate if there is any other articulatory justification for that; it has then utilized Ultrasound imaging analysis to explore the variations of the tongue body as well as the lip associations during the articulation of the emphatic contrast.

It is crucial to note that none of the previous research that concerned with examining the Arabic emphatic fricatives has ever examined the lip association during the articulation of the emphatic and non-emphatic consonants with direct articulatory evidence. The study proposed some research questions (see section 5.2.4) and the answer was that all speakers from all tested dialects have shown significant differences between the emphatic contrast; the nature of the differences lay in the tongue shape variation and overall, the tongue dorsum was observed to be retracted during the articulation of the emphatic consonant while it was seen advanced when speakers articulating plain consonants (as presented in the results of PCA scores in section 5.4.1.1).

Although all speakers showed this tongue variation, they varied in terms of the degree and direction of the tongue backing. For example, among Hijazi speakers, subject hh02 showed more tongue root retraction while hh03 showed tongue dorsum retractions. The same was also observed within other dialects. See the Table 51 below that shows the summary of the speakers with tongue root retraction and others with tongue dorsum retraction within each dialect.

Dialect	Subjects with Tongue root retraction	Subjects with Tongue dorsum retraction
Hijazi	hh02	hh03
Syrian	sy02, sy03	sy01, sy04, sy05
Iraqi	ir02, ir04	ir01
Egyptian	eg02	eg01
Moroccan	mo02	

Table 51 Observed subjects with different direction of tongue retraction

The point I want to emphasize here is that the variations occur among speakers within a dialect, suggesting that the observed variation in tongue shape is due to individual behaviours rather than inter-dialectal variation. This can be attributed to several factors, including cross-dialectal variations and individual differences in articulatory strategies employed by speakers when producing emphatic sounds (Aldamen, 2013; Al-Solami, 2017; Khattab et al., 2006). This suggests that the variation in articulating the emphatic contrast is characterized by individual-specific patterns. Therefore, this finding addresses the research question: 'Are there individual-specific patterns in articulatory mechanisms?' Another important point to highlight is that this inconsistency in tongue root retraction is expected since previous research has suggested that tongue root retraction often occurs due to the retraction of the tongue dorsum (Al-Solamy, 2017).

As for the lip association, it was hypothesized that lips would be rounded during emphatic sounds and unrounded or spread for non-emphatic sounds. Although this hypothesis was motivated by the results from previous research[14], only a few subjects, such as eg02, hh01, and hh02 (refer to Figure 96), exhibited lip rounding for the emphatic /sˤ/, whereas their lips remained unrounded for the plain /s/. Interestingly, some subjects showed significant lingual differences for emphasis, yet they did not exhibit clear lip rounding. In contrast, others showed clear lip rounding during emphasis but did not show greater lingual differences. As shown in the summary of Table 52 below, there is no clear association between the

[14] El-Halees (1985), Hetzron (2013), and Jakobson (1957) observed some degree of lip protrusion associated with emphatic sounds. Additionally, Lehn (1963) noted slight lip rounding in the Carine dialect, while Bellem and Watson (2014) identified labialization in Ṣanʿāni Arabic and Southern (Gilit) Iraqi Arabic.

articulation of emphatic contrast and lip rounding. Therefore, the hypothesis that the articulation of emphatic fricatives necessarily includes the additional feature of lip rounding does not hold as a valid prediction.

Subjects	Lingual difference	Lip Movement During Emphasis
hh02	Significant difference, greater than hh01	Slightly rounded
hh01	Significant difference, but less than hh02	Clearly rounded
hh03	The most significant difference among Hijazi subjects	Unrounded
sy04	The most significant difference across all dialects	Unrounded
mo02	The most significant difference across all dialects	Unrounded

Table 52 The association between lip rounding and the relative significant difference in tongue shape for the emphatic contrast

This overall finding addresses the research question, "Does lip rounding contribute to distinguishing between emphatic contrasts among Arabic dialects?" It appears there is no significant contribution of lip rounding to distinguishing emphatic contrasts, which may explain the individual variability among subjects. This mismatch in associations could be attributed to the methodological limitations of the study; specifically, it did not utilize actual spline analyses of lip movements, relying instead, on static lip imagery. Additionally, the camera used to record lip movements was positioned to capture only the frontal aspect of the lips. These limitations will be discussed in more detail in the limitations section.

6.4 Articulatory -to-Acoustic

6.4.1 Acoustics Analysis in the Articulatory Study

As we have discussed in 'the Articulatory study' in chapter 5, along with collecting the articulatory data the acoustic data for COG correlates was also collected simultaneously. Interestingly, even though only four subjects—namely, eg02, hh02, sy02, and ir02—exhibited significant differences both acoustically and articulatorily for the /s/ versus /sˤ/ contrast (refer to Table 47, section 5.4.3), our analysis of the relationship between the regression model estimates for acoustic and articulatory results revealed a crucial insight. Specifically, no correlation was observed, as demonstrated in Figure 100 (reproduced from Articulatory Study, Section 5.4.5.2).

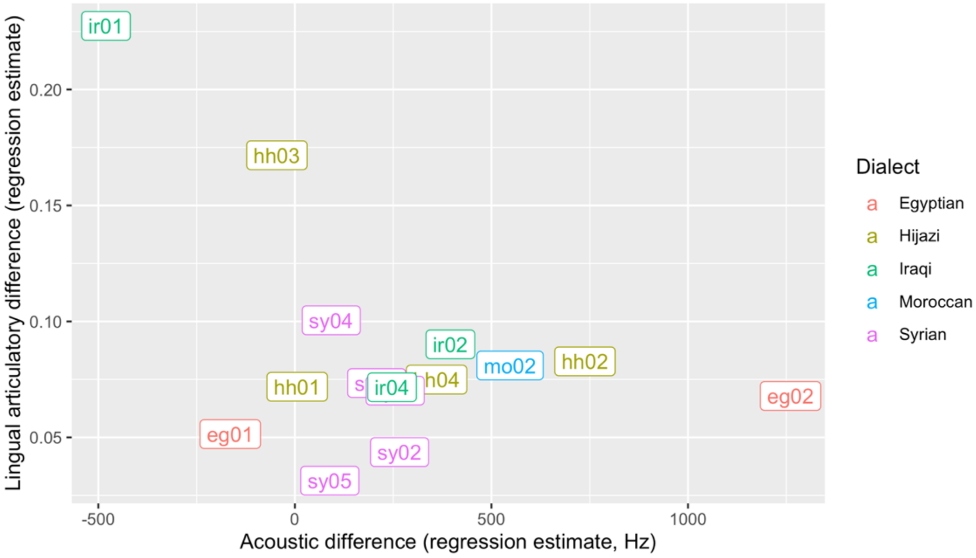


Figure 100 The regression estimates for production and articulatory results across all subjects

To illustrate, subject eg02 shows the largest acoustic difference in the regression model but is average in articulation. Meanwhile, subject hh03 has an almost zero acoustic estimate but exhibits significant articulatory differences [15]. Similarly, subject ir01 has strong lingual articulation but lower acoustic estimates, indicating a disconnect between lingual articulation and acoustic results for these subjects. This lack of correlation is also seen in other subjects, leading to the conclusion that articulatory outcomes do not align closely with acoustic findings.

This mismatch may be attributed firstly to the fact that some aspects of speech production cannot be easily recovered from acoustic analysis, and secondly, to the implications of using articulatory instruments like Ultrasound Imaging which is capable of identifying the covert features of speech articulation (Lawson et al., 2014). Covert articulatory variation, as described by Mielke et al. (2017), occurs when speakers employ different articulatory strategies to produce sounds that are acoustically or auditorily indistinguishable. One illustrative example of this mismatch is offered by Strycharczuk and Scobbie (2017) in their study. They investigated the fronting of the high/back vowels /u/ and /ʊ/ in Southern British English dialect. Their goal was to compare articulatory findings with the corresponding acoustic measurements of the second formant (F2) for both vowels. Despite the clear

[15] In Chapter 5, as shown in PCA Figure 85 and Table 46, it is evident that the tongue splines for subject hh03 are distinctly different, albeit with a minimal acoustic difference.

distinction in tongue position for /u/ and /ʊ/, they found no significant corresponding difference in F2. This indicates that the lack of mapping between acoustic and articulatory data may be justifiable.

6.4.2 Acoustics Analysis in the Acoustic Study

The observed inconsistency in acoustic results in the Articulatory study differs from the findings offered in the acoustic chapter 3 (section 2) of this thesis. There, the analysis of Centre of Gravity (COG) values did not reveal any significant influence of fricative type (emphatic vs. non-emphatic) on COG, despite dialectal variability [β= -90.369, SE= 55.542, $p = 0.1218$]. As shown in Figure 101 below (previously presented in Chapter 3), COG values exhibited an occurrence of overlap between plain and emphatic fricatives, with only the effects of gender [β= -599.852, SE= 78.678, $p < 2.82e-11$], dialect [β= 906.534, SE= 197.141, $p < 1.39e-05$], and vowel type [β= 425.416, SE= 74.795, $p < 2.75e-05$] observed on COG values.

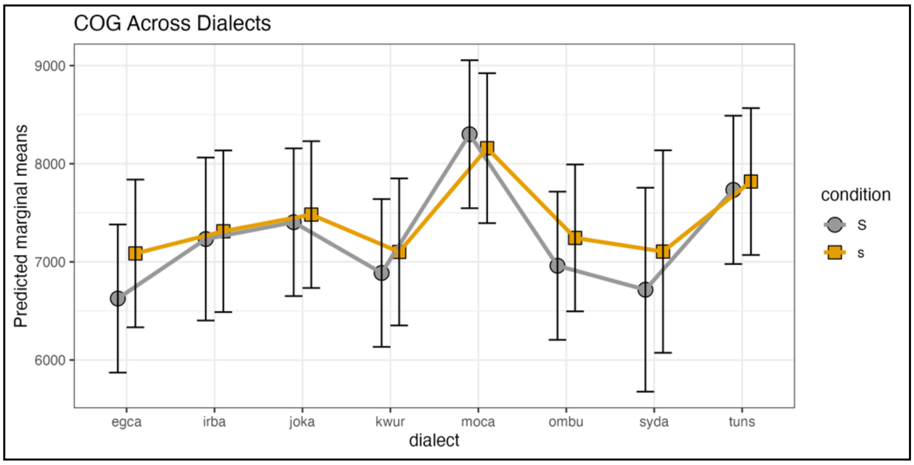


Figure 101 Results from LMER for fricative COG frequencies of /sˤ/ and /sˤ/ across dialects

These findings, coupled with previous research by Norlin (1983) and Sakr (2024), suggest that CoG may not be a reliable acoustic correlate for distinguishing emphatic and non-emphatic consonants due to the observed overlap in values. One potential explanation for the discrepancies between the two acoustic outcomes may lie in the differences between the recording conditions. In the articulation study, participants were simultaneously recorded acoustically while undergoing ultrasound imaging. This setup required participants to sit in a fixed position, use a head-stabilizing unit, and employ an ultrasound transducer, potentially resulting in less natural articulation. In contrast, the acoustic study allowed for more typical and unconstrained speech, which could explain the differentiation in CoG values. Moreover, the artificial constraints of the ultrasound study may have amplified individual variability in articulatory strategies, contributing to the overlap in CoG values. The influence of these controlled conditions on natural speech production highlights the potential impact of task effects on acoustic outcomes. Thus, the differences in CoG results between the two studies underscore the importance of considering recording environments when interpreting acoustic data.

6.4.3 **Acoustic-to-Articulatory Results (Vowel Formant Information)**

As previously discussed, the articulatory analysis played a crucial role in highlighting the differences between the articulation of emphatic /sˤ/ and the non-emphatic /s/. Despite the consistency with which an articulatory distinction in the fricatives is present, the extent to which this is represented by significant acoustic differences is highly variable. In Chapter 3, it has been shown that the reliable acoustic correlates can be found in formant information of the vowels following the emphatic contrast. According to resonance models (Kent & Read, 1992; Pickett, 1999), F1 frequency is inversely correlated with the degree of constriction in the oropharyngeal cavity. This suggests that a greater constriction in the front of the vocal tract (achieved by raising the tongue body), which reduces the oral space and expands the pharyngeal space, will lower F1. However, lowering the tongue body (lessening the pharyngeal cavity) will increase F1. In simpler terms, increased pharyngeal space corresponds to a lower F1. In a similar term, the perturbation theory that formulated by Chiba and Kajiyama in (1958) stated that narrowing the space at or close to a formant's antinode reduces the formant frequency, whereas a constriction close to the node of a formant increases its frequency. Conversely, expanding the space at these points has the opposite impact. Thus, expanding the area near a formant's antinode will raise its frequency, and expanding near a formant's node will decrease its frequency. The locations of nodes and antinodes for the first two formants, F1 and F2, are depicted in Figure 102 below. This was clearly shown from the PCA results, particularly when examining PC1 figures such as Figures 103 and 104 (taken from Figures 85, 86 for subject hh02), demonstrate that the tongue body is lowered during the articulation of emphatic fricatives. The height of the tongue body significantly influences the pharyngeal cavity of the vocal tract. As Pickett (1999) notes, lowering the tongue body forces the tongue volume towards the pharyngeal wall, leading to tongue root retraction. Both these actions contribute to an increase in F1. Consequently, F1 is influenced by both the height of the tongue body and the retraction of the tongue root.

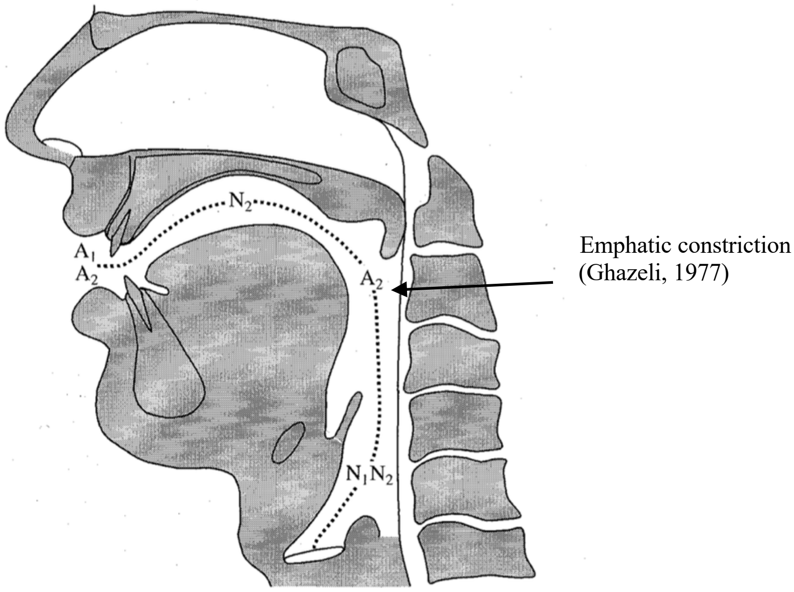


Figure 102 Positions of nodes and antinodes for the F1 and F2 formants, based on Bin-Muqbil (2006, p. 11)

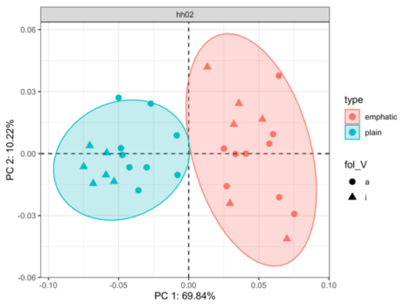


Figure 103 PC1 and PC2 scores of the xy coordinates during /sˤ/ and /sˤ/ for Hijazi subject hh02

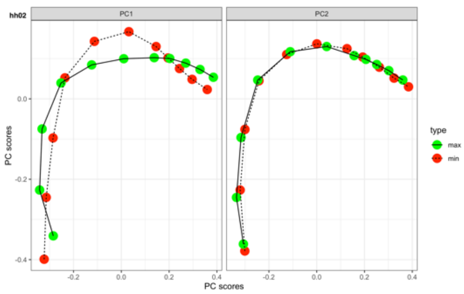


Figure 104 PC1 and PC2 tongue tracing for Hijazi subject hh02

This observation aligns with the acoustic results discussed in Chapter 3; although F1 was not identified as a primary factor in differentiating the emphatic contrast [β= 17.19249, SE= 10.10031, p = 0.10376], the overall F1 mean value for the vowel following the emphatic /sˤ/ was found to be comparatively higher than that for the non-emphatic /s/. Refer to Figure 105 below, which is adopted from the acoustic study in Chapter 3.

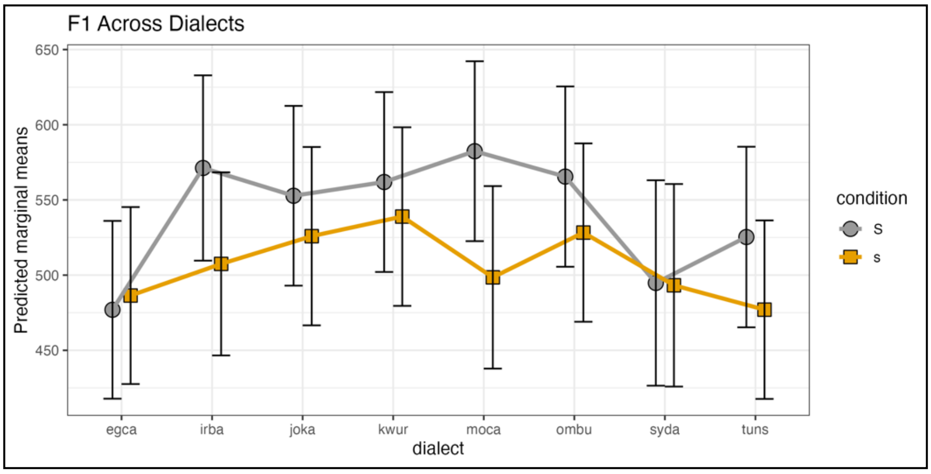


Figure 105 Results from LMER for fricative F1 frequencies of /sˁ/ and /sˁ/ across dialects

F2 is also associated with the degree of constriction within the oropharyngeal area of the vocal tract. According to Pickett (1999), F2 frequency is influenced by the location of the constriction and the resulting length of the oral cavity in front of the constriction. A constriction in the front of the vocal tract shortens the oral cavity and lengthens the pharyngeal cavity, leading to a higher F2. Conversely, a constriction further back in the oral cavity has the opposite effect, lengthening the front cavity and shortening the pharyngeal cavity, resulting in a lower F2. To simplify, a more forward tongue position (shorter front cavity) corresponds to a higher F2, while a retracted tongue position (longer front cavity) leads to a lower F2. Therefore, F2 reflects the front-back movement of the tongue body. Thus, the lingual findings of this study demonstrated clear tongue retraction with emphatic sounds, as indicated by the PC1 scores, where emphatics consistently show positive PC1 (indicating to the tongue body retraction) in comparison to non-emphatics. Referring to the acoustic study, chapter 3, the linear regression model suggested that F2 formant information in the following vowel showed a significant difference between the plain and emphatic fricative contrast [β= -126.908, SE= 47.350, p = 0.01424], including factors such as dialect, gender, and vowel type. Observing Figure 106 below, it is apparent that all dialects utilize this measure to distinguish the emphatic contrast.

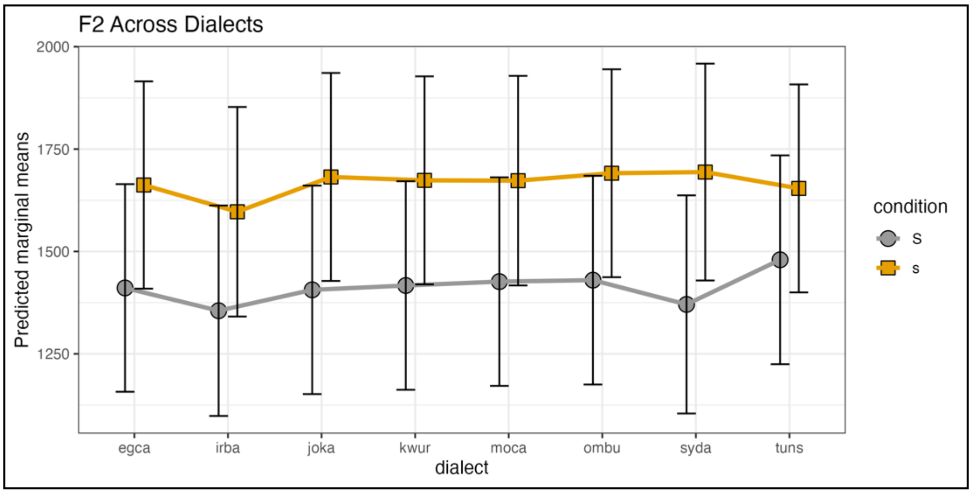


Figure 106 Results from LMER for fricative F2 frequencies of /sˁ/ and /sˁ/, across dialects

Also, the Generalized Additive Mixed Models (GAMM) analysis, as detailed in the acoustic study in Chapter 3 of this thesis, provided insightful findings concerning the articulation patterns of different vowels, specifically /a:/ and /i:/, following emphatic versus non-emphatic fricatives. The analysis presented a notable distinction in the F2 trajectory patterns when comparing emphatic to non-emphatic fricatives, highlighting a systematic approach employed by speakers of various dialects to acoustically mark the emphatic contrast. This significant variance underscores the importance of F2 trajectory as a reliable acoustic marker for identifying emphatic sounds within speech. Such findings not only enrich our understanding of the phonetic nuances inherent in different dialects but also validate the use of F2 trajectory as a critical parameter for distinguishing between these fricative types. The implications of these results are visually represented in Figure 37 in Chapter 3, which depicts the F2 trajectory effect of the vowels following the fricatives, offering a clear visual affirmation of the significant differences observed between emphatic and non-emphatic sounds across the examined dialects.

Thus, spotting the distinction between the emphatic contrast is all about formant information of the following vowels as stated by huge body of study (see Chapter 2 section 4). This may be because of the implication of the covert articulatory variation (discussed earlier) by which the sound variation can easily be discovered articulatory but not easily distinguished acoustically (for more details see (Mielke et al., 2017).

As for the third formant information (F3), several studies have explored the connection between tongue placement and F3. Kent and Read (1992) also discuss how the F3 pattern correlates with the location of pharyngeal constriction, noting that a lower constriction in the pharynx is linked to higher F3 frequency. Therefore, a constriction in the lower pharynx raises F3, whereas it is reduced with a constriction located in the middle of the pharynx. However, A constriction in the upper pharynx either has no effect on F3 or may cause a slight increase. Similarly, Lindblom and Sundberg (1971) highlight that F3 rises when the tongue moves from the velar to the pharyngeal area while Stevens (2000) observed an increase in all formant frequencies associated with lower pharyngeal constriction. This potentially explains the observed variation in tongue shapes across subjects when producing emphatic sounds; some subjects demonstrated that the tongue dorsum is retracted and raised toward the velum during the articulation of emphasis, indicating that the constriction occurs in the upper pharynx. Meanwhile, others retracted their tongue toward the root position, indicating that the location of the constriction occurs in the lower pharynx. Furthermore, the acoustic findings from the acoustic chapter indicated that the rise in F3 for emphatic sounds was not consistently observed across dialects. Refer to Figure 107 below for illustration.

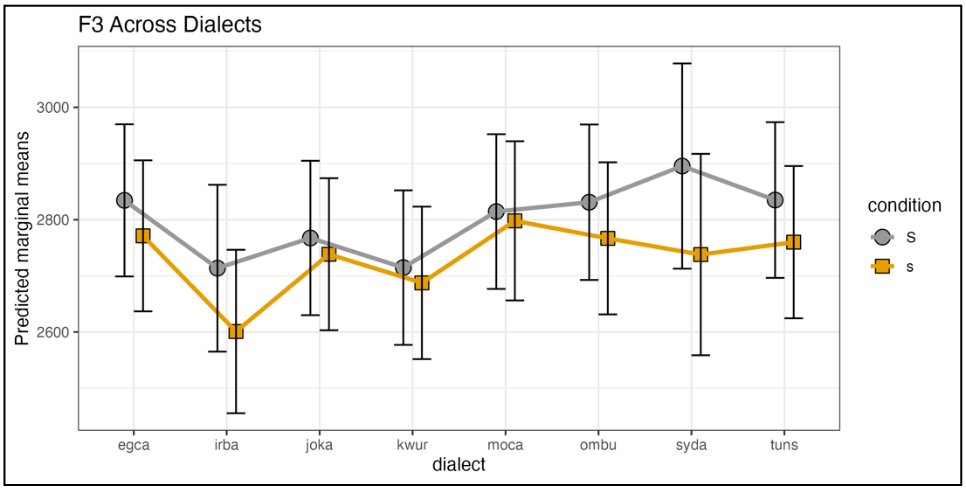


Figure 107 Results from LMER for fricative F3 frequencies of /$s^ʕ$/ and /$s^ʕ$/ across dialects

7 Chapter 7: Conclusion

7.1 What is it that varies?

A recurring question that arises during this thesis is, "How do the emphatic contrasts s~sˤ vary?" The answer to this question varies, depending on the methodology used in the study. From an acoustic standpoint, the realization of the emphatic contrast can be described by the second formant information (F2) of the vowel following the contrast; F2 is lower with emphatics compared to non-emphatics. Perceptually, the differentiation in the realization of the emphatic contrast among Arabic listeners is based on the F2 trajectory (differing shapes and sizes); dialect listeners pay attention to the various shapes and sizes of the F2 trajectory. Articulatorily, the retraction of the tongue dorsum is key in differentiating the emphatic contrast; the tongue is retracted during the articulation of emphatic consonants and advanced for plain consonants. Despite these findings, it remains complex to specify exactly what accounts for these emphatic differences. The acoustic results suggest it is a vocalic phenomenon, while articulatory analysis indicates a consonantal aspect, suggesting that there may be other measures or methodologies yet to be explored. Moreover, the complexity of emphatic contrast phenomena is reflected in the varied findings of researchers, attributed to the multifaceted nature of the phenomenon. It is crucial to note that in articulatory analysis, observed variations in emphatic contrast are due to individual behavioural differences, potentially influenced by the nature of the methodology. I would also posit that this individual specificity is probably present in other studies but is obscured by measures of central tendency, that researchers may not pay sufficient attention to report the degree of variation in their data.

7.2 Contributions

This thesis makes several unique contributions to the understanding of the emphatic contrast across Arabic dialects.

7.2.1 Multimodal Approach

One of the key distinguishing features of this thesis is its adoption of a multimodal approach to investigate the emphatic contrast. While prior studies often relied on a single methodology, usually acoustic analysis, this research integrates three distinct methodologies: acoustics, perception, and articulation. By utilising multiple methodologies, we gain a more comprehensive understanding of the intricacies surrounding the emphatic contrast phenomenon. Acoustic analysis provides insights into the acoustic properties of emphatic sounds, perception studies shed light on how listeners perceive these contrasts, and articulatory analysis unveils the underlying mechanisms involved in their production. This multimodal approach allows for a holistic examination of the emphatic contrast from various perspectives, enhancing the depth and breadth of our insights into this complex linguistic phenomenon. Overall, the multimodal approach adopted in this thesis represents a significant methodological advancement, paving the way for a more comprehensive understanding of the emphatic contrast across Arabic dialects.

7.2.2 **Dialectal Diversity:**

In contrast to previous studies that often focused on a limited number of dialects (one or two), with some, such as Embarki et al. (2011), examining up to four dialects, this thesis adopts a more inclusive approach by examining eight distinct Arabic dialects. This broad coverage of dialectal diversity provides a richer and more nuanced understanding of the emphatic contrast phenomenon across the Arabic language spectrum. By encompassing a diverse range of dialects, including Syrian, Iraqi, Egyptian, Moroccan, Jordanian, Tunisian, Hijazi, Omani, and Kuwaiti, this research captures the diverse linguistic landscapes within the Arabic-speaking world. Each dialect brings its unique phonetic characteristics and variations to the table, enriching our appreciation of the multifaceted nature of the emphatic contrast. Furthermore, by comparing and contrasting the emphatic contrasts across different dialects, we can identify commonalities and differences, leading to insights into the underlying linguistic and cultural factors shaping these variations.

Moreover, the examination of dialectal diversity allows for a more comprehensive analysis of the robustness and universality of the emphatic contrast phenomenon. By testing the consistency of these contrasts across various dialects, we can assess the extent to which they are invariant or subject to dialectal variation. This exploration not only contributes to our understanding of phonetic variation within Arabic but also has broader implications for theories of language variation and change. By highlighting the richness and diversity of Arabic dialects, this thesis underscores the importance of considering dialectal factors in linguistic research and emphasizes the need for a more inclusive approach to studying language phenomena.

7.2.3 F2 Trajectory Analysis:

This thesis breaks new ground by delving deeper into the analysis of F2 trajectories, surpassing the conventional focus solely on F2 lowering in the vowel following fricative consonants. Rather than examining two points—onset and midpoint, or midpoint and offset—in the F2 trajectory (e.g., Al-Tamimi, 2017; Embarki et al., 2011; Zeroual et al., 2011), this research explores the entire trajectory, including its size effect and overall shape. By scrutinising the full extent of F2 movement, we gain a more nuanced understanding of how emphatic contrasts manifest across dialects. This comprehensive approach allows for a detailed characterization of the emphatic contrast phenomenon, shedding light on subtle variations in F2 trajectories that may have previously gone unnoticed. Furthermore, by analysing the size effect and shape of the F2 trajectory, we can uncover additional layers of complexity in the realization of emphatic contrasts, providing valuable insights into the acoustic correlates used by speakers to articulate emphasis in speech. Overall, the F2 trajectory analysis undertaken in this thesis represents a significant methodological advancement, offering a more detailed and comprehensive examination of the emphatic contrast phenomenon.

7.2.4 Unfamiliar Cues in Perception:

An interesting aspect of this thesis is its investigation into listeners' perception of emphatic contrast when presented with unfamiliar cues from different dialects. Unlike previous studies that focused on presenting familiar cues to the listeners (often from their own native dialects), this research employs an innovative approach by exploring how listeners perceive cues that they are not accustomed to hearing. The findings not only enhance our understanding of cross-dialectal perception but also shed light on the robustness of perceptual processes in accommodating variation in linguistic input. Overall, the exploration of unfamiliar cues in perception represents a novel and impactful contribution to the study of emphatic contrast in Arabic dialects.

7.2.5 Articulation and Lip Movement:

While previous research has explored the articulatory aspects of the emphatic contrast, this thesis investigated the association between lip movement and emphatic contrast articulation using direct articulatory evidence. By employing advanced articulatory techniques such as Ultrasound Imaging, this study offers interesting insights into the articulatory mechanisms involved in producing emphatic contrasts. The findings not only enrich our understanding of the physiological processes underlying speech production but also provide valuable clues about the connect between different articulatory organs during the articulation of emphatic sounds. Overall, the exploration of articulation and lip movement represents an innovative contribution that enhances our understanding of the complex interplay between articulatory gestures and linguistic contrasts in speech.

7.3 Limitations

This thesis, like any scholarly work, is not impervious to limitations, which are acknowledged here. These limitations not only highlight areas for improvement but also serve as a foundation for future research.

7.3.1 Acoustic Study Limitations

The acoustic study (Section 3.3.1) encountered several limitations, notably regarding sample size differences among dialect groups. While some dialects had full representation with twelve participants each, others, such as the Syrian and Iraqi groups, had smaller sample sizes, including only six and ten participants, respectively. This imbalance in sample sizes could potentially impact the generalisability of the findings across different dialects. Additionally, acoustic measurements for fricatives were only taken at the midpoint of the fricative consonants. However, to fully capture the nuances of emphasis spread and distinguish emphatic contrasts accurately, measurements at various points along the fricatives, including onset and offset, are crucial to account for co-articulatory effects and temporal dynamics.

7.3.2 Perception Study Limitations

In the perception study, several limitations were encountered. Firstly, the manipulation focused exclusively on the second formant (F2) information, overlooking the potential contributions of manipulated F1 and F3 to the perception of emphasis (Section 2.4.1). Additionally, the stimuli manipulation process was conducted manually, introducing the possibility of inconsistency and human error. Developing a Praat script for automatic manipulation would enhance the consistency and reliability of the manipulation process. Moreover, all manipulated stimuli utilized the initial target sound /s/ from an Omani dialect recording, chosen for its proximity to the Centre of Gravity (COG) across all dialects (Section 4.4.2). However, this approach may limit the generalisability of the findings to other dialects. Finally, the perception study lacked repetition of real stimuli or inclusion of fillers, potentially confounding the results. The absence of fillers may make the /s~sˤ/ contrast overly salient to listeners, affecting their ability to accurately perceive emphasis, as discussed in Section 4.4.5.

7.3.3 **Articulatory Study Limitations**

The articulatory study utilized a frontal face lip camera, which may have limitations in capturing the full range of articulatory movements. To address this limitation, future research could consider incorporating side-profile lip videos in addition to frontal views. This would provide a more comprehensive perspective on articulatory gestures and enhance the understanding of the mechanisms involved in producing emphatic contrasts.

Furthermore, the exclusion of some speaker recordings due to issues with lip video and audio synchronization may have introduced biases or gaps in the data. Efforts to improve synchronization techniques or alternative methods for data collection could mitigate this limitation and ensure a more comprehensive analysis of articulatory patterns across speakers

7.4 Suggested Future Work

7.4.1 **Acoustic Study:**

Future investigations should delve deeper into the emphatic contrast present across diverse Arabic dialects, employing a comprehensive range of acoustic correlates. Specifically, researchers could consider incorporating measures such as COG, peak location, and intensity, measured at multiple points along the fricative consonants' duration: onset, midpoint, and offset. By examining these acoustic properties at different points of articulation, researchers can more effectively capture the co-articulatory effects inherent in speech production, providing a more nuanced understanding of the emphatic contrast phenomenon.

7.4.2 **Perception Study:**

Future research would design an experiment where the stimuli incorporate conflicting information across context, vowels, and consonants. The objective would be to elucidate the relative relationship between context and the vowel. Specifically, the experiments would investigate whether listeners are able to accurately identify the target words when presented with wrong cues for vowels and consonants. For example, listeners would hear the word /Siin/ "China, the country" with manipulated F2 for the vowel (giving a plain-like cue), but presented with a different context (such as showing a map of China).

7.4.3 **Articulatory Study:**

Future research endeavours should adopt dynamic analysis through Ultrasound imaging, to provide a comprehensive examination of the articulatory gestures involved in producing emphatic contrasts. Rather than focusing solely on the midpoint of fricative consonants, researchers could explore the entire duration of fricatives and the subsequent vowels. By capturing the dynamic articulatory movements throughout these segments, researchers can gain deeper insights into the intricate articulatory mechanisms underlying emphatic sounds. Additionally, incorporating side-profile videos of lip movements would offer valuable insights into the association between lip articulation and the production of target sounds, further enriching our understanding of the articulatory correlates of emphatic contrasts.

In summary, future research endeavours should adopt a multifaceted approach encompassing acoustic, perceptual, and articulatory analyses to advance our understanding of the emphatic contrast phenomenon in Arabic dialects. By addressing the suggested areas for future research outlined above, researchers can uncover new insights into the phonetic, perceptual, and physiological dimensions of emphatic sounds, contributing to a more comprehensive understanding of this linguistically significant phenomenon.

8 References

- Abu-Al-Makarem & Cooper (2006). The Acoustic Characteristics of Voiceless Fricatives of Gulf Spoken Arabic. Bowling Green State University, Bowling Green, OH Poster Presented at the 2006 ASHA Convention November 16-18. Miami, Florida.
- Adams, D., Rohlf, F., Slice, D. (2004). Geometric morphometrics: Ten years of progress following the 'revolution'. Italian Journal of Zoology, 71(1), 5-16.
- Ahyad, H., & Becker, M. (2020). Vowel unpredictability in Hijazi Arabic monosyllabic verbs. Glossa: a journal of general linguistics, 5(1).
- AI-Nassir, A. (1993). Sibawayh the phonologist. London: Kegan Paul International Ltd.
- Al-Ani, S., & el-Dalee, M. S. (1984). Tafkhim in Arabic: The acoustic and psychological parameters. In *Proceedings of the 10th International Congress of Phonetic Sciences* (pp. 385-389). Dordrecht: Foris.
- Al-Khairy, M. A. (2005). Acoustic characteristics of Arabic fricatives. University of Florida.
- Al-Masri, M., & Jongman, A. (2004). Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results. Proceedings of the 2003 Texas Linguistics Society Conference. Somerville, MA: Cascadilla Proceedings Project, 96–106.
- Al-Solami, M. A. (2017). Ultrasound study of emphatics, uvulars, pharyngeals and laryngeals in three Arabic dialects. Canadian Acoustics, 45(1), 25-35.
- Al-Tamimi, F., & Heselwood, B. (2011). Nasoendoscopic, videofluoroscopic and acoustic study of plain and emphatic coronals in Jordanian Arabic. *Instrumental studies in Arabic phonetics*, *319*, 166-191
- Al-Tamimi, J. (2007). Indices dynamiques et perception des voyelles: étude translinguistique en arabe dialectal et en français. *Unpublished PhD Dissertation, University Lyon*, *2*.Al
- Al-Tamimi, J. (2007)b. Static and dynamic cues in vowel production: A cross dialectal study in Jordanian and Moroccan Arabic. In *16th International Congress of Phonetic Sciences* (pp. pp-541).
- Al-Tamimi, J. (2017). Revisiting acoustic correlates of pharyngealisation in Jordanian and Moroccan Arabic: Implications for formal representations. Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8(1).

- Al-Tamimi, J. E., & Ferragne, E. (2005, September). Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. *Proceedings of Interspeech 2005*, 2465–2468.
- Al-Wer, E., (2000). Education as a speaker variable in Arabic variationist studies.
- AlDahri, S. S. (2013). A study for the effect of the emphaticness and language and dialect for voice onset time (VOT) in Modern Standard Arabic (MSA). *arXiv preprint arXiv:1305.2680*.
- Aldamen, H. A. K. (2013). The production of emphasis by second language learners of Arabic (Doctoral dissertation, University of Kansas).
- Alfaifi, A. H., Cavar, M. E., & Lulich, S. M. (2020, December). Tongue root position in Hijazi Arabic voiceless emphatic and non-emphatic coronal consonants. In Proceedings of Meetings on Acoustics (Vol. 42, No. 1). AIP Publishing.
- Alhammad, R. (2014). Emphasis Spread in Najdi Arabic.
- Ali, L. H., & Daniloff, R. G. (1972). A contrastive cinefluorographic investigation of the articulation of emphatic-non emphatic cognate consonants. *Studia Linguistica*, *26*(2), 81-105.
- Alioua, A. (1995). L'effet des consonnes d'arriere et des emphatiques sur la nature acoustique des voyelles longues de l'Arabe litteral moderne. Unpublished doctoral dissertation, Universite Laval, Montreal, Canada.
- Alkuwaiz, L. (2022). Laryngeal contrast and phonetic voicing: a cross-dialectal laboratory phonology approach for Arabic (Doctoral dissertation, University of York).
- Almashaqba, B. M. (2015). The phonology and morphology of Wadi Ramm Arabic (Phd, University of Salford).
- Almuhaimeed, A. (2022). *Emphasis Spread in the Najdi Arabic Dialect* (Doctoral dissertation, University of Manchester
- Almurashi, W., Al-Tamimi, J., & Khattab, G. (2020). Static and dynamic cues in vowel production in Hijazi Arabic. *The Journal of the Acoustical Society of America*, *147*(4), 2917-2927.
- Almurashi, W., Al-Tamimi, J., & Khattab, G. (2024). Dynamic specification of vowels in Hijazi Arabic. *Phonetica*, (0).
- Alosh, M. M. (1987). The perception and acquisition of pharyngealized fricatives by American learners of Arabic and implications for teaching Arabic phonology (PhD Thesis). The Ohio State University.

- Altairi, H., Brown, J., Watson, C., & Gick, B. (2017, May). Tongue retraction in Arabic: An ultrasound study. In Proceedings of the Annual Meetings on Phonology (Vol. 4).
- Alwan, A. (1986). Acoustic and perceptual correlates of uvular and pharyngeal consonants. Unpublished S.M. thesis, MIT.
- Anwyl-Irvine, A.L., Massonié, J., Flitton, A., Kirkham, N.Z., Evershed, J.K. (2019). Gorilla in our midst: an online behavioural experiment builder Behavior Research Methods. Doi: https://doi.org/10.3758/s13428-019-01237-x
- Articulate Instruments Ltd. (2022). Articulate Assistant Advanced (Version 219.08) [Computer software]. Edingurgh: Articulate Instruments Ltd. Retrieved April 01, 2022, from http://www.articulateinstruments.com/downloads/
- Behrens, S. J., & Blumstein, S. E. (1988). Acoustic characteristics of English voiceless fricatives: A descriptive analysis. Journal of Phonetics, 16(3), 295–298.
- Bellem, A. (2014). Triads, emphatics and interdentals in Arabic sound system typology. In Journal of Semitic studies. (Vol. 34, pp. 9–41). Oxford University Press.
- Bellem, A., & Watson, J. C. (2014). Backing and glottalisation in three SWAP Language varieties. In Journal of Semitic studies. (Vol. 34, pp. 169-207). Oxford University Press.
- Bellem, A. (2008). Towards a Comparative Typology of Emphatics: Across Semitic and Into Arabic Dialect Phonology. University of London.
- Bin-Muqbil, M. S. (2006). *Phonetic and phonological aspects of Arabic emphatics and gutturals*. The University of Wisconsin-Madison.
- Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer. (version 5.1.3.1). Available online at: www.fon.hum.uva.nl/praat/. Computer Program Available At< Http://Www. Praat. Org. www.fon.hum.uva.nl/praat/
- Bookstein F. L., (1996)a - Biometrics, biomathematics and the mor-phometric synthesis. Bull. Math. Bid., 58: 313-365.
- Bukshaisha, F. A. M. (1985). An experimental phonetic study of some aspects of Qatari Arabic. *Annexe Thesis Digitisation Project 2017 Block 15*.
- Card, E. (1983). A Phonetic and Phonological Study of Arabic Emphasis. PhD thesis, Cornell University.
- Carignan, C. (2018). Using ultrasound and nasalance to separate oral and nasal contributions to formant frequencies of nasalized vowels. *The Journal of the Acoustical Society of America*, *143*(5), 2588-2601.

• Carignan, C., & Zellou, G. (2023). Sociophonetics and vowel nasality. In *The Routledge Handbook of Sociophonetics* (pp. 237-259). Routledge

• Carré, R., & Mrayati, M. (1992). Distinctive regions in acoustic tubes. Speech production modelling. *Journal d'acoustique (Les Ulis)*, *5*(2), 141-159.

• Chejne, A. G. (1969). *The Arabic language, its role in history / Anwar G. Chejne.* Minneapolis: University of Minnesota Press.

• Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*(4), 2360-2370.

• Chiba, T., & Kajiyama, M. (1958). *The vowel, its nature and structure*. Phonetic society of Japan.

• Crumbie, L., Salvador, F., & Rad, A. (2019). Tongue. In. Germany: Kenhub GmbH.

• Davis, S. (1995). Emphasis Spread in Arabic and Grounded Phonology. Linguistic Inquiry, 26(3), 465–498.

• Delattre, P. (1971). Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English. Phonetica, 23, 129-155.

• Delattre, P. (1951). The physiological interpretation of sound spectrograms. *Pmla*, *66*(5), 864-875.

• Delvaux, V. (2009). "Perception du contraste de nasalite vocalique en franc ¸ais" ("Perception of the vowel nasality contrast in French"), French Lang. Studies 19, 25–29.

• El-Halees, Y. (1985). The role of F1 in the place-of-articulation distinction in Arabic. Journal of Phonetics, 13(3), 287-298.

• Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America*, *126*(3), 1379-1393.

• Fant, G. (1960/1971). Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations, second edition. Walter de Gruyter.

• Feng, G., & Castelli, E. (1996). Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *The Journal of the Acoustical Society of America*, *99*(6), 3694-3706.

• Ferguson, C. (1956). The emphatic f J.f in Arabic. Language, 32 (2), 446-452.

• Fujimura, O., & Lindqvist, J. (1971). Sweep-tone measurements of vocal-tract characteristics. *The Journal of the Acoustical Society of America*, *49*(2B), 541-558.

• Ghazeli, S. (1977). Back consonants and backing coarticulation in Arabic. (PhD). The University of Texas at Austin, United States.
• Giannini, A., Pettorino, M. (1982). "The emphatic consonants in Arabic". Speech Laboratory Report IV, Istituto Universitario Orientale di Napoli.
• Gick, B, I. Wilson, K. Koch, and C. Cook. (2004). "Language-specific articulatory settings: Evidence from inter- utterance rest position." Phonetica, 61 (4): 220-233.
• Gick, B., Wilson, I., & Derrick, D. (2013). Articulatory Phonetics. Malden, MA, USA: Wiley- Blackwell.
• Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. Journal of the International Phonetic Association, 141–174.
• Gorman, K., Howell, J., & Wagner, M. (2011). Prosodylab-aligner: A tool for forced alignment of laboratory speech. Canadian Acoustics, 39(3), 192–193.
• Gully, A. (2021). Quantifying vocal tract shape variation and its acoustic impact: A geometric morphometric approach. Proceedings of INTERSPEECH 2021, 3999-4003.
• Hardcastle, W. J., & Hewlett, N. (2006). *Coarticulation: Theory, data and techniques*. Cambridge University Press.
• Harrell, R. S. (1957). The phonology of colloquial Egyptian Arabic.
• Hassan, Z. (1981). *An experimental study of vowel duration in Iraqi spoken Arabic*. PhD. Leeds University.
• Hassan, Z. M., & Esling, J. H. (2011). Investigating the emphatic feature in Iraqi Arabic: Acoustic and articulatory evidence of coarticulation.
• Hellmuth, S., & Almbark, R. (2017, November 24). Intonational variation in Arabic Corpus 2011-2017 [Data Collection]. http://reshare.ukdataservice.ac.uk/852878/10/text_checklist.xlsx
• Hermes, Z., Wong, N., Loucks, T. M., & Shosted, R. (2015). The primary articulation of plain-emphatic/s/-/sˤ/in Lebanese Arabic: An EMA study. In ICPhS.
• Herzallah, R. S. (1990). *Aspects of Palestinian Arabic phonology: A nonlinear approach*. Cornell University.
• Hetzron, R. (2013). The Semitic Languages: Routledge.
• Hixon, T. J., Weismer, G., & Hoit, J. D. (2014). Preclinical Speech Science: Anatomy, Physiology, Acoustics, Perception (Second Edition). San Diego, CA: Plural Publication.

- House, S. A., & Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders, 21*, 218–232.
- Howard, D., & Angus, J. (2009). *Acoustics and psychoacoustics*. Routledge.
- Ingham, B. (1982). Northeast Arabian dialects. London: Kegan Paul International.
- Jakobson, R. (1957). Mufaxxama: The 'emphatic' phonemes in Arabic. Volume I Phonological Studies, 510-522.
- Jastrow, O. (2002). 'Neo-Aramaic dialectology. The state of the art'. Israel Oriental Studies, 20, 347-363.
- Johnson, K. (2012). *Acoustic and auditory phonetics* (3rd ed.). Wiley-Blackwell, A John Wiley & Sons, Ltd., Publication.
- Johnson, K., & Johnson, K. (2004). Acoustic and auditory phonetics. *Phonetica*, *61*(1), 56-58.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202.
- Jongman, A., Herd, W., & Al-Masri, M. (2007). Acoustic correlates of emphasis in Arabic.
- Jongman, A., Herd, W., Al-Masri, M., Sereno, J., & Combest, S. (2011). Acoustics and perception of emphasis in Urban Jordanian Arabic. Journal of Phonetics, 39(1), 85–95.
- Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. The Journal of the Acoustical Society of America, 108(3), 1252–1263.
- Kahn, M. (1975). Arabic emphatics: The evidence for cultural determinants of phonetic sex - typing. Phonetica, 31, 38-50. doi:10.1159/000259648
- Kavanagh, C. (2012). *New consonantal acoustic parameters for forensic speaker comparison* (Doctoral dissertation, University of York).
- Kent, R. D. (1998). The Speech Sciences. San Diego, California: Singular Publishing Group.
- Kent, R. D., Read, (1992). "The acoustic analysis of speech" (Vol. 58). San Diego: Singular Publishing Group.
- Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006, February). Acoustic and auditory differences in the/t/-/tˤ/opposition in male and female speakers of Jordanian Arabic. In Perspectives on Arabic Linguistics XVI: Papers from the sixteenth annual symposium on Arabic linguistics (pp. 131-160). Amsterdam: John Benjamins.

- Kier, W. M., & Smith, K. K. (1985). Tongues, tentacles and trunks: The biomechanics of movement in muscular-hydrostats. Zoological Journal of the Linnean Society, 83, 307-324. Retrieved from: https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1096-3642.1985.tb01178.x
- Klatt, D., & Stevens, K. N. (1969). Pharyngeal consonants. MIT Research Laboratory of Electronics Quarterly Progress Report, 93, 207-216.
- Kriba, H. A. (2010). *Acoustic parameters of emphasis in Libyan Arabic* (Doctoral dissertation, Newcastle University).
- Kulikov, V., Mohsenzadeh, F. M., & Syam, R. M. (2021). 'Effect of Emphasis Spread on VOT in Coronal Stops in Qatari Arabic', Journal of the International Phonetic Association, 1-14
- ‘Effect of Emphasis Spread on VOT in Coronal Stops in Qatari Arabic’, Journal of the International Phonetic Association, 1–14
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). *lmerTest Package: Tests in linear mixed effects models*. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13. Retrieved November 2020.
- Laradi, W. (1983). Pharyngealisation in Libyan (Tripoli) Arabic. *University of Edinburgh. PhD Thesis*.
- Laufer, A., & Baer, T. (1988). The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and speech*, *31*(2), 181-205.
- Lawson, E., Stuart-Smith, J., & Scobbie, J. M. (2014). A mimicry study of adaptation towards socially-salient tongue shape variants. *University of Pennsylvania Working Papers in Linguistics*, *20*(2).
- Lehn, W. (1963). Emphasis in Cairo Arabic. Language: Journal of the Linguistic Society of America, 34, 29-39.
- Lieberman, P. & Blumstein, S. (1988). Speech physiology, speech perception, and acoustic phonetics. Cambridge: Cambridge University Press.
- Lindblom, B., & Sundberg, J. (1971). Acoustical consequences of lip, tongue, and larynx movement. Journal of the Acoustical Society of America, 50(4, Part 2), 1166-1179.
- LmerTest Package: Tests in Linear Mixed Effects Models—Google Search. (n.d.). Retrieved 4 November 2020, from: https://www.jstatsoft.org/article/view/v082i13
- Löfqvist, A. (1990). Speech as audible gestures. In *Speech production and speech modelling* (pp. 289-322). Dordrecht: Springer Netherlands.

• Lukose, A. M., Gully, A., & Bailey, G. (2023). Markerless tongue pose estimation in cross-linguistic spoken and mimed speech using ultrasound.

• Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In *Nasals, nasalization, and the velum* (pp. 147-167). Academic Press.

• Malmberg, B. (1963). Structural linguistics and human communication. Springer.

• Maryais, P. (1948). L'articulation de l'emphase dans un parler arabe maghrébin. Annales de l'Institut d'Études Orientales (Alger), 7, 5-28.

• McCarthy, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. A. Keating (Ed.), Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III. Cambridge: Cambridge University Press.

• McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III, 191–233.

• McCasland. (1979). Noise intensity and spectrum cues of spoken fricatives: The Journal of the Acoustical Society of America: Vol 65, No S1.

• Mielke, J., Smith, B., & Fox, M. J. (2017). Implications of covert articulatory variation for several phonetic variables in Raleigh, North Carolina English. The Journal of the Acoustical Society of America, 141(5_Supplement), 3981-3981.

• Moisik, S. (2013). The epilarynx in speech (Unpublished doctoral dissertation). Victoria: University of Victoria.

• Moulines, E., & Charpentier, F. (1990). *Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones*. Speech Communication, 9(5-6), 453-467. https://doi.org/10.1016/0167-6393(90)90021-Z

• Mrayati, M., Carré, R., & Guérin, B. (1988). Distinctive regions and modes: A new theory of speech production. *Speech Communication*, *7*(3), 257-286.

• Nearey, T. M., & Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. *The Journal of the Acoustical Society of America, 80*(5), 1297–1308. https://doi.org/10.1121/1.394433

• Norlin, K. (1983). ACOUSTIC ANALYSIS OF FRICATIVES TN CAIRO ARABIC. Working papers/Lund University, Department of Linguistics and Phonetics, 25.

• Norlin, K. (1987). A phonetic study of emphasis and vowels in Egyptian Arabic. Working papers/Lund University, Department of Linguistics and Phonetics, 30.

• Obrecht, D. (1968). Effects of the second formant on the perception of velarisation consonants in Arabic. Paris: Mouton.

• Palmer, J. (1993). Anatomy for speech and hearing. Baltimore: Williams & Wilkins.

• Palva, H. (2006). 'Dialects: Classification'. In: K. Versteegh et al (eds.). Encyclopedia of Arabic Language and Linguistics, 1. Leiden: Brill, 604- 613.

• Palva, H. (1991). 'Is there a Northwest Arabian dialect group?' In M. Forstner (ed.) Festgabe für Hans-Rudolf Singer, zum 65. Geburtstagam 6. April 1990 überreicht von seinen Freunden und Kollegen, Part 1. Frankfurt: Peter Lang: 151-166.

• Perkins, W. & Kent, R. (1986). Functional anatomy of speech, language and hearing: A primer. San Diego: College-Hill Press.

• Pickett, J. M. (1999). "The acoustics of speech communication: Fundamentals, speech perception theory, and technology". Allyn & Bacon.

• Polly, P. (2012). Procrustes, PCA, and 3D coordinates. Geometric morphometrics module lecture notes. Indiana University: Department of Earth & Atmospheric Sciences.

• R Core Team (2020). —European Environment Agency. (n.d.). [Methodology Reference]. Retrieved 4 November 2020, from https://www.eea.europa.eu/data-and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-2006

• R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing,

• Rosenhouse, J. (2006). 'Bedouin Arabic'. In: K. Versteegh et al (eds.). Encyclopedia of Arabic Language and Linguistics, 1: Leiden: Brill, 259-269.

• Sakr, G. (2024). On the Acoustics of Emphasis in Central Mount Lebanon Lebanese. *Journal of Semitic Studies*, 69(1), 601-642.

• Sakr, G. (2023). "On the Acoustics of Emphasis in Central Mount Lebanon Lebanese." Journal of Semitic Studies, September, fgad024.

• Sanders, I., & Mu, L. (2013). A three-dimensional atlas of human tongue muscles. Anat Rec (Hoboken), 296(7), 1102-1114. doi:10.1002/ar.22711

• Seikel, J. A., King, D. W., & Drumright, D. G. (2000). Anatomy and Physiology for Speech Language and Hearing (Second edition). San Diego, California: Singular Publishing Group.

• Seikel, J., King, D., & Drumright, D. (1997). Anatomy and physiology for speech and language. San Diego: Singular Publishing Group.

• Serrurier, A., & Badin, P. (2008). A three-dimensional articulatory model of the velum and nasopharyngeal wall based on MRI and CT data. *The Journal of the Acoustical Society of America*, *123*(4), 2335-2355.

- Shadle, C. H. (1985). The acoustics of fricative consonants.
- Shosted, R. K., Fu, M., Hermes, Z., Fu, M., & Hermes, Z. (2017, December 22). Arabic pharyngeal and emphatic consonants.
- Siemund, R., Heuft, B., Choukri, K., Emam, O., Maragoudakis, E., Tropf, H., Gedge, O., Shammass, S., Moreno, A., & Rodriguez, A. N. (2002). OrienTel—Arabic speech resources for the IT market. LREC 2002 (Arabic Workshop).
- Smorenburg, L., & Heeren, W. (2020). The distribution of speaker information in Dutch fricatives/s/and/x/from telephone dialogues. *The Journal of the Acoustical Society of America*, *147*(2), 949-960.
- Spreafico, L., Pucher, M., Matosova, A. (2018). UltraFit: A speaker-friendly headset for ultrasound recordings in speech science. INTERSPEECH 2018, 1517-1520.
- Stevens, K. N. (1968). The Quantal Nature of Speech: Evidence from Articulatory-acoustic Data.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of phonetics*, *17*(1), 3-45.
- Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). MIT press.
- Stone, M. (1997). "Laboratory techniques for investigating speech articulation". In Handbook of Phonetic Sciences. eds. William Hardcastle and John Laver, 11-32. Oxford: Blackwell Publishers.
- Stone, M., Woo, J., Lee, J., Poole, T., Seagraves, A., Chung, M., ... & Blemker, S. S. (2018). Structure and variability in human tongue muscle anatomy. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualisation, 6(5), 499-507.
- Strycharczuk, P., & Scobbie, J. M. (2017). Fronting of Southern British English high-back vowels in articulation and acoustics. *The Journal of the Acoustical Society of America*, *142*(1), 322-331.
- Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America*, *142*(4), 2469-2482.
- The Mathis Group, & The Mathis Lab. (2022). *DeepLabCut* (Version 2.2.2) [Computer software]. Switzerland: The Mathis Group, Swiss Federal Institute of Technology Lausanne. Retrieved December 2023 from http://www.mackenziemathislab.org/deeplabcut
- Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide.
- Van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, D. (2015). itsadug: Interpreting time series and autocorrelated data using GAMMs.

• Vienna, Austria. <https://www.R-project.org/>.

• Wahba, K. (1993). A sociolinguistic treatment of the feature of emphasis in Egypt. *Unpublished doctoral dissertation, University of Texas at Austin*.

• Watson, J. C. (1999). The directionality of emphasis spread in Arabic. Linguistic Inquiry, 30(2), 289–300.

• Watson, J. C. (2007). The Phonology and Morphology of Arabic.

• Wood, S. N. (2017). Generalised Additive Models: An Introduction with R, Second Edition. CRC Press.

• Wrench, A., & Balch-Tomes, J. (2022). Beyond the edge: markerless pose estimation of speech articulators from ultrasound and camera images using DeepLabCut. *Sensors*, *22*(3), 1133.

• Yeou, M. (1995). Trading relations between cues for the pharyngealized/non pharyngealized contrast. In *Proceedings of the 13th International Congres of Phonetic Sciences* (pp. 464-467).

• Yeou, M. (1997). Locus Equations and the Degree of Coarticulation of Arabic Consonants. Phonetica, 54(3–4), 187–202. https://doi.org/10.1159/000262221

• Yeou, M. (2001). Pharyngealisation in Arabic: Modelling, acoustic analysis, airflow, and perception. *Revue de la Faculté des Lettres Chouaib Doukkali, 6*, 51-70.

• Younes, M. (1993). Emphasis spread in two Arabic dialects. AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4, 119-119.

• Zawaydeh, B. (1998). Gradient uvularisation spread in Ammani-Jordanian Arabic. In: A. Benmamoun, M. Eid and N. Haeri, eds. Perspectives on Arabic Linguistics XI. AmsterdamlPhiladelphia: John Benjamins, 117-141.

• Zawaydeh, B. (1999). The phonetics and phonology of gutturals in Arabic. Thesis (PhD). Indiana University, Bloomington.

• Zemlin, W. R. (1998). Spech and Hearing Science: Anatomy and Physiology (Fourth edition). Boston, MA: Allyn & Bacon.

• Zeroual, C., Esling, J. H., & Hoole, P. (2011). EMA, endoscopic, ultrasound and acoustic study of two secondary articulations in Moroccan Arabic. Instrumental studies in Arabic phonetics, 319, 277.

9 Appendix A (For Acoustic Study Chapter 3)

9.1 Appendix: Models' comparison for including the random factors

It shows the models comparison for including the random factors, where M1 refers to the addition of the random effect "speaker" only, whereas M2 refers to the addition to the other random effect "item".

DVs	Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
Fricative duration	M1	-5281.7	-5266.3	2643.8	-5287.7	-	-	-
	M2	-5704.9	-5684.3	2856.4	-5712.9	425.19	1	<2.2e-16 ***
intensity	M1	6806	6822.2	-3400.4	6800.9	-	-	-
	M2	-5704.9	-5684.3	2856.4	-5712.9	425.19	1	3.371e-16 ***
Peak location	M1	22520	22535	-11257	22514	-	-	-
	M2	22436	22456	-11214	22428	86.312	1	< 2.2e-16 ***
Vowel duration	M1	-3778.1	-3762.7	1892.0	-3784.1	-	-	-
	M2	-5388.3	-5367.8	2698.2	-5396.3	1612.3	1	< 2.2e-16 ***
F1	M1	16070	16085	-8031.8	16064	-	-	-
	M2	14667	14687	-7329.4	14659	1404.8	1	< 2.2e-16 ***
F2	M1	19068	19084	-9531.1	19062	-	-	-
	M2	17116	17136	-8553.9	17108	1954.5	1	< 2.2e-16 ***
F3	M1	17524	17540	-8759.0	17518	-	-	-
	M2	17346	17367	-8669.2	17338	179.67	1	< 2.2e-16 ***

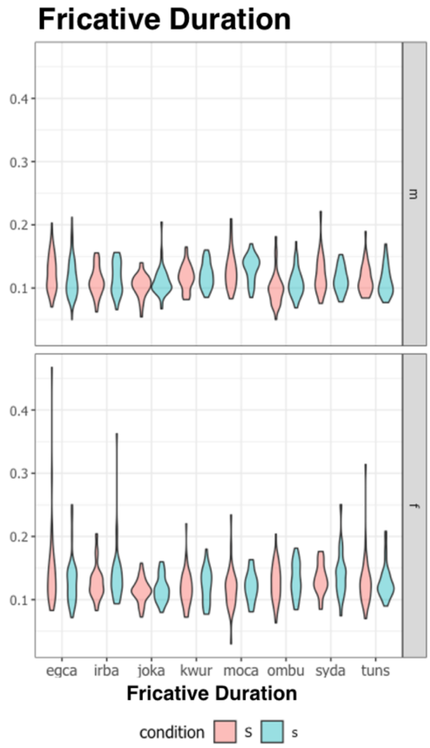
9.2 Appendix: Models comparison for including the fixed effects

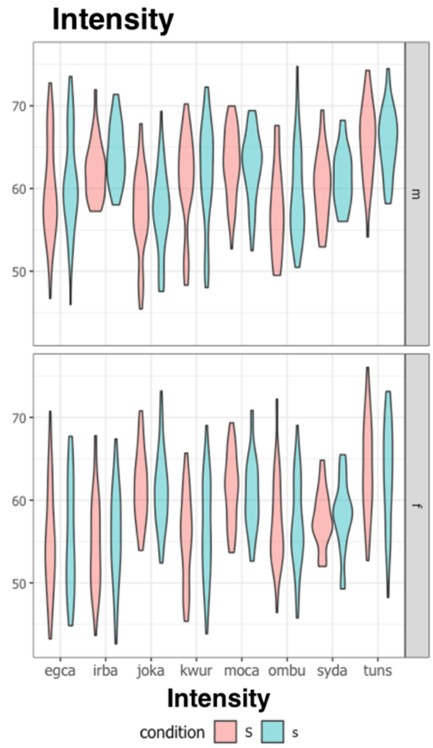
It shows the models comparison for including the fixed effects, where M1 refers to the addition of "dialect" and "condition" only, whereas M2 refers to the addition to the other fixed effects "gender" and "vowel type".

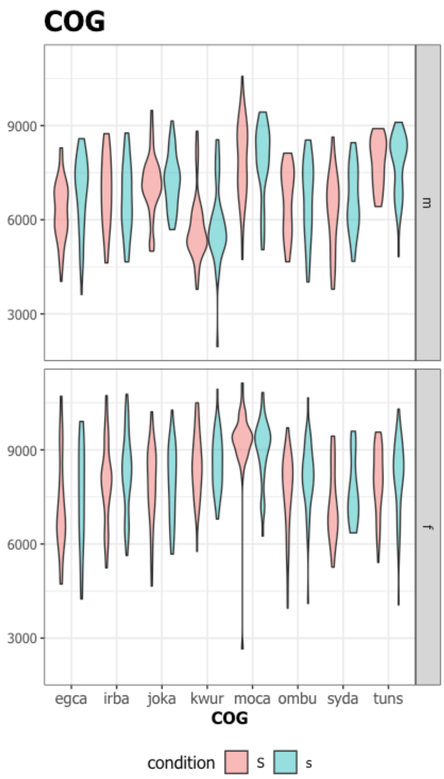
Model comparison results for the analyzed dependent variables (DVs), including Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Log-Likelihood (logLik), Deviance, Chi-squared (Chisq) statistics, Degrees of Freedom (Df), and P-values (Pr(>Chisq)).

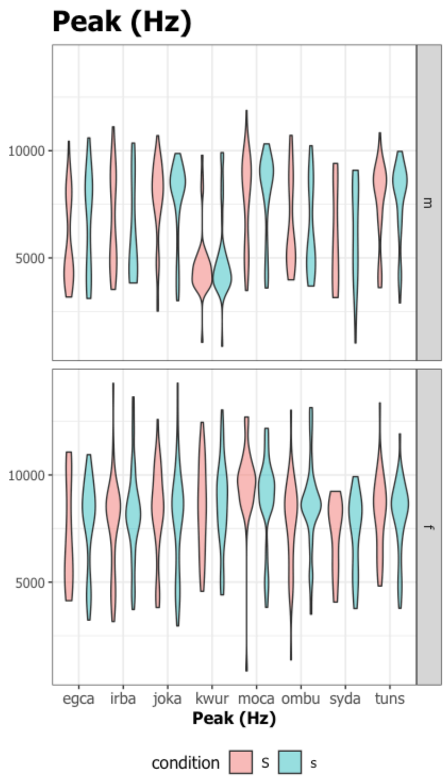
DVs	Model	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
Fricative duration	M1	-5713.6	-5616.1	2875.8	-5751.6	-	-	-
	M2	-5721.9	-5532.2	2897.9	-5795.9	44.34	18	0.0005157 ***
intensity	M1	6741.9	6839.4	-3352.0	6703.9	-	-	-
	M2	6725.6	6915.4	-3325.8	6651.6	52.335	18	3.323e-05***
Peak location	M1	22437	22534	-11199	22399	-	-	-
	M2	22380	22570	-11153	22306	92.578	18	4.959e-12***
Vowel duration	M1	-5395.9	-5298.5	2717.0	5433.9	-	-	-
	M2	-5386.6	-5196.9	2730.3	-5460.6	26.688 18	18	0.08505
F1	M1	14641	14738	-7301.3	14603	-	-	-
	M2	14507	14697	-7216.4	14433	169.69	18	< 2.2e-16 ***
F2	M1	17134	17231	-8547.8	17096	-	-	-
	M2	16999	17189	-8462.6	16925	170.34	18	< 2.2e-16 ***
F3	M1	17360	17458	-8661.0	17322	-	-	-
	M2	17277	17467	-8601.5	17203	119.09	18	< 2.2e-16 ***

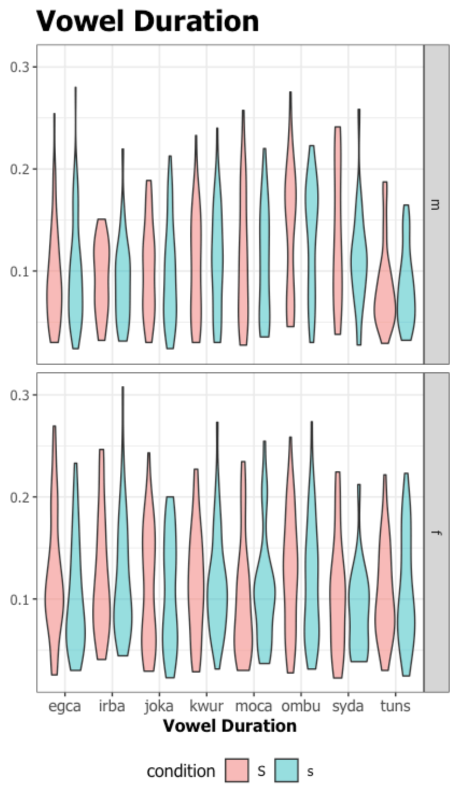
9.3 Appendix shows the Raw data for the DVs

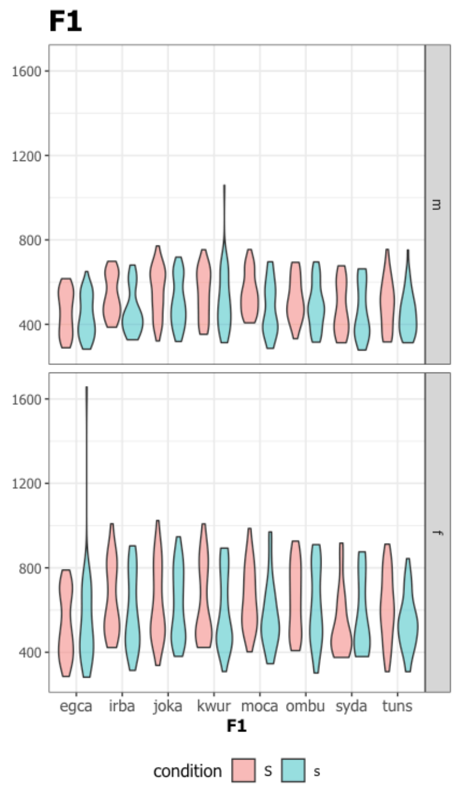


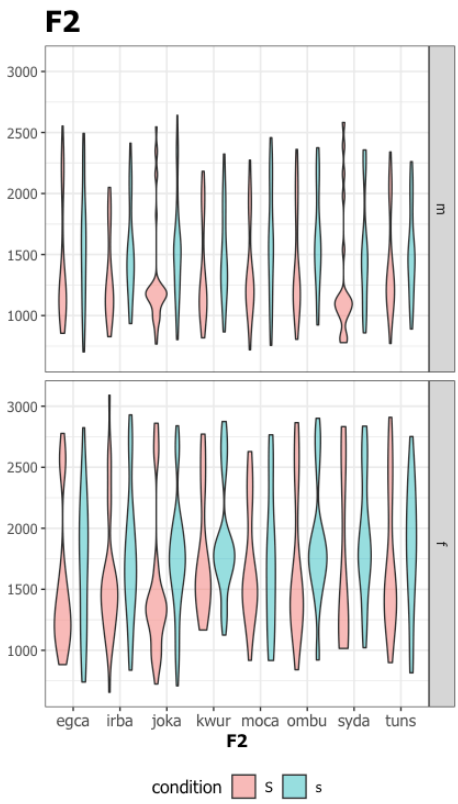


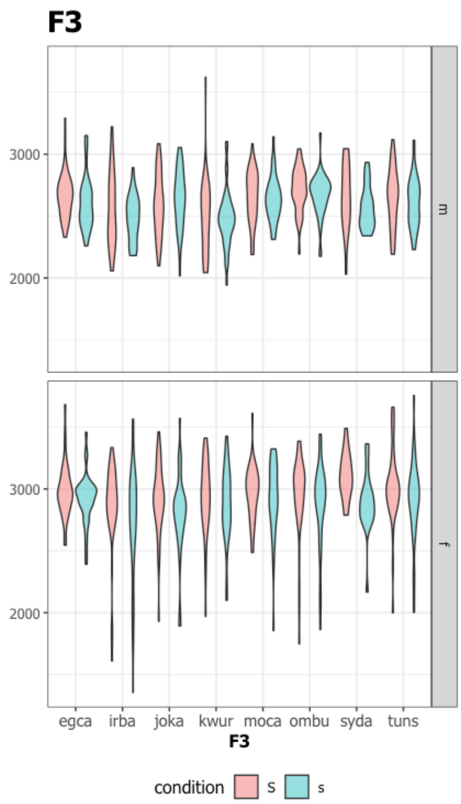





Vowel Duration
0.3
0.2
0.1
m
f
egca irba joka kwur moca ombu syda tuns
Vowel Duration
condition S s


F1
1600
1200
800
400
m
f
egca irba joka kwur moca ombu syda tuns
F1
condition S s


F2
3000
2500
2000
1500
1000
m
f
egca irba joka kwur moca ombu syda tuns
F2
condition S s


F3
3000
2000
m
f
egca irba joka kwur moca ombu syda tuns
F3
condition S s

9.4 Mean, standard deviation (SD), and range of Fricative Duration, grouped by dialect and condition.

dialect	condition	mean_fricative_duration	sd_fricative_duration	range_fricative_duration
Egyptian	$s^ʕ$	0.13346718	0.04967292	0.069994158 - 0.467452043
Egyptian	s	0.12067026	0.03351209	0.050046108 - 0.250129035
Iraqi	$s^ʕ$	0.12141273	0.02679608	0.062161063 - 0.204270343
Iraqi	s	0.12894533	0.03822538	0.065877186 - 0.362566241
Jordanian	$s^ʕ$	0.10899139	0.01895098	0.054256767 - 0.157737559
Jordanian	s	0.11341934	0.02134397	0.067109617 - 0.204238189
Kuwaiti	$s^ʕ$	0.11744559	0.02435219	0.072442628 - 0.220176194
Kuwaiti	s	0.12107395	0.02339654	0.077133826 - 0.18
Moroccan	$s^ʕ$	0.12362815	0.03129846	0.03 - 0.234051194
Moroccan	s	0.12662256	0.0218961	0.080891056 - 0.17
Omani	$s^ʕ$	0.11475014	0.03105413	0.050200446 - 0.203780094
Omani	s	0.12026358	0.02712151	0.068634835 - 0.181453346
Syrian	$s^ʕ$	0.12723106	0.02917209	0.075729825 - 0.22115728
Syrian	s	0.12674407	0.03296706	0.074517821 - 0.250309348
Tunisian	$s^ʕ$	0.12090211	0.03315958	0.07 - 0.313909045
Tunisian	s	0.11865639	0.02559905	0.077045015 - 0.208280851

9.5 Mean, standard deviation (SD), and range of Intensity values, grouped by dialect and condition

dialect	condition	mean_intensity	sd_intensity	range_intensity
Egyptian	$s^ʕ$	58.0072499	6.50147432	43.23000932 - 72.77906688
Egyptian	s	58.9308081	6.7126657	44.83340563 - 73.54789509
Iraqi	$s^ʕ$	57.6259143	6.11023431	43.64879394 - 71.93602115
Iraqi	s	59.1542718	6.47472475	42.62822994 - 71.37561446
Jordanian	$s^ʕ$	59.3997396	5.21696673	45.43687319 - 70.79315204
Jordanian	s	59.2233481	5.21283337	47.56086216 - 73.18042262
Kuwaiti	$s^ʕ$	58.2528498	6.17920943	45.36168451 - 70.23128722
Kuwaiti	s	59.2442709	6.60350881	43.85446461 - 72.2631733
Moroccan	$s^ʕ$	62.2967206	4.27917371	52.70499849 - 69.96637812
Moroccan	s	61.7240059	4.31224706	52.48753399 - 70.84987122
Omani	$s^ʕ$	57.4184892	5.34421769	46.43614647 - 72.21449476
Omani	s	58.2704737	5.54007601	45.77198382 - 74.76649732
Syrian	$s^ʕ$	59.0618631	4.06017966	51.98854165 - 69.47755711
Syrian	s	59.8530003	4.13694516	49.28124692 - 68.23553021
Tunisian	$s^ʕ$	64.8010166	5.12947549	52.69894229 - 76.04887539
Tunisian	s	64.450602	5.33882183	48.22861866 - 74.47904058

9.6 Mean, standard deviation (SD), and range of COG values, grouped by dialect and condition.

dialect	condition	mean_cog	sd_cog	range_cog
Egyptian	sˤ	6673.5417	1307.4655	4035.655937 - 10707.98081
Egyptian	s	7117.92889	1379.3036	3615.680122 - 9904.447537
Iraqi	sˤ	7559.37085	1340.6114	4625.581522 - 10728.92721
Iraqi	s	7615.00287	1419.16188	4656.463007 - 10771.86654
Jordanian	sˤ	7513.45637	1176.62172	4659.107622 - 10210.12669
Jordanian	s	7589.93563	1079.75102	5677.06108 - 10266.49936
Kuwaiti	sˤ	7087.45619	1734.67531	3782.659012 - 10498.11458
Kuwaiti	s	7213.80422	1752.612	1956.821692 - 10932.43082
Moroccan	sˤ	8499.82492	1407.27481	2656.824057 - 11125.03282
Moroccan	s	8402.23517	1221.67143	5046.612882 - 10828.38906
Omani	sˤ	7226.63777	1265.24597	3953.243304 - 9702.74162
Omani	s	7438.6431	1374.93723	4016.141092 - 10657.2332
Syrian	sˤ	6733.78942	1274.5229	3786.422081 - 9437.607705
Syrian	s	7161.40594	1149.10191	4675.948443 - 9592.650797
Tunisian	sˤ	7931.80481	927.161141	5413.617496 - 9563.138972
Tunisian	s	7936.89863	1121.98017	4059.863167 - 10299.31917

9.7 Mean, standard deviation (SD), and range of PeakHz values, grouped by dialect and condition.

dialect	condition	mean_peakHz	sd_peakHz	range_peakHz
Egyptian	sˤ	6371.67473	2303.39519	3173.842928 - 11066.8882
Egyptian	s	7151.59046	2262.08882	3107.75805 - 10947.38346
Iraqi	sˤ	7323.96303	2250.06309	3162.043788 - 14274.62418
Iraqi	s	7334.85107	2408.20934	3722.336096 - 13637.27323
Jordanian	sˤ	8007.70297	2020.67388	2516.759249 - 12588.23808
Jordanian	s	7837.58284	2150.50827	2958.436878 - 14282.08442
Kuwaiti	sˤ	6536.06498	2642.3951	1063.261593 - 12455.63967
Kuwaiti	s	6813.15103	2699.73627	875.9372871 - 13032.17717
Moroccan	sˤ	8561.54153	2213.52208	860.8783492 - 12703.09293
Moroccan	s	8409.26715	2038.98341	3597.644073 - 12174.34732
Omani	sˤ	7237.03282	2198.62741	1372.661663 - 13022.86925
Omani	s	7438.36554	2370.22089	3505.655632 - 13140.54555
Syrian	sˤ	6608.05574	2118.70828	3149.318155 - 9407.188184
Syrian	s	6792.76557	2123.01852	1030.495714 - 9920.808291
Tunisian	sˤ	7927.24414	1857.8998	3614.248107 - 13360.09892
Tunisian	s	7906.77574	1778.14	2899.168804 - 11914.30891

9.8 Mean, standard deviation (SD), and range of Vowel Duration for long and short vowels, grouped by dialect, condition, and vowel label

9.8.1 Long Vowel : /a:/

dialect	condition	vowel_label	mean_vowel_duration	sd_vowel_duration	range_vowel_duration
Egyptian	sˤ	a:	0.15441533	0.05969083	0.08279615 - 0.269423588
Egyptian	s	a:	0.13553086	0.04373662	0.086583389 - 0.23301677
Iraqi	sˤ	a:	0.14175701	0.04083781	0.096586836 - 0.245634877
Iraqi	s	a:	0.13704287	0.037168	0.085635214 - 0.221640467
Jordanian	sˤ	a:	0.14490637	0.03676586	0.100991258 - 0.243163416
Jordanian	s	a:	0.13843237	0.03241129	0.091153245 - 0.212672663
Kuwaiti	sˤ	a:	0.15412889	0.03772636	0.0988961 - 0.232793292
Kuwaiti	s	a:	0.14662682	0.04016522	0.094863147 - 0.273123067
Moroccan	sˤ	a:	0.14792474	0.05126059	0.069460743 - 0.257380632
Moroccan	s	a:	0.14076257	0.04150398	0.078489492 - 0.22
Omani	sˤ	a:	0.17375333	0.0367476	0.106179724 - 0.258579018
Omani	s	a:	0.17015479	0.03784999	0.100909549 - 0.273681406
Syrian	sˤ	a:	0.17068646	0.04828927	0.10379657 - 0.228918964
Syrian	s	a:	0.13954222	0.05270752	0.073455615 - 0.258388366
Tunisian	sˤ	a:	0.13493955	0.04638114	0.075641308 - 0.221650345
Tunisian	s	a:	0.12481217	0.04660119	0.05 - 0.223110878

9.8.2 Long Vowel : /i:/

dialect	condition	vowel_label	mean_vowel duration	sd_vowel duration	range_vowel_duration
Egyptian	sˤ	i:	0.12477923	0.03830899	0.053927371 - 0.2
Egyptian	s	i:	0.12651729	0.03636584	0.064038753 - 0.2
Iraqi	sˤ	i:	0.1208003	0.04850023	0.068730161 - 0.239879035
Iraqi	s	i:	0.1217435	0.05843734	0.04366081 - 0.307688341
Jordanian	sˤ	i:	0.14401118	0.03549709	0.090450908 - 0.192971041
Jordanian	s	i:	0.13808499	0.04110638	0.08 - 0.2
Kuwaiti	sˤ	i:	0.13286053	0.03737986	0.04 - 0.227157448
Kuwaiti	s	i:	0.12519922	0.03931424	0.03222785 - 0.2
Moroccan	sˤ	i:	0.12713727	0.04127982	0.05 - 0.2064942
Moroccan	s	i:	0.11387917	0.03509817	0.050046234 - 0.206660306
Omani	sˤ	i:	0.15489943	0.03778755	0.094897192 - 0.236953598
Omani	s	i:	0.15099373	0.03165106	0.101556406 - 0.222764795
Syrian	sˤ	i:	0.13679918	0.05287987	0.07638349 - 0.241243779
Syrian	s	i:	0.11304821	0.02666785	0.060804472 - 0.162002774
Tunisian	sˤ	i:	0.09675278	0.0296826	0.056079018 - 0.160352219
Tunisian	s	i:	0.10319191	0.02727815	0.056140299 - 0.150942165

9.8.3 Long Vowel : /u:/

dialect	condition	vowel_label	mean_vowel duration	sd_vowel duration	range_vowel_duration
Egyptian	sˤ	u:	0.15932535	0.03798373	0.132466796 - 0.186183906
Egyptian	s	u:	0.16743514	0.05311324	0.060042872 - 0.28
Iraqi	sˤ	u:	0.17225766	0.04219631	0.128379866 - 0.246486214
Iraqi	s	u:	0.14683066	0.05933043	0.061812206 - 0.244552455
Jordanian	sˤ	u:	0.15458325	0.01945919	0.117790644 - 0.175375757
Jordanian	s	u:	0.15062129	0.03393502	0.09 - 0.190747646
Kuwaiti	sˤ	u:	0.17380805	0.03798373	0.132466796 - 0.186183906
Kuwaiti	s	u:	0.15345986	0.04555479	0.103565064 - 0.24
Moroccan	sˤ	u:	0.20818961	0.03542755	0.163416274 - 0.2523692
Moroccan	s	u:	0.15659393	0.05290055	0.07107927 - 0.254561701
Omani	sˤ	u:	0.18977831	0.04789383	0.140987391 - 0.275356128
Omani	s	u:	0.17508275	0.03090549	0.118640215 - 0.21576055
Syrian	sˤ	u:	0.15802484	0.06283712	0.080196537 - 0.23
Syrian	s	u:	0.10397928	0.02177998	0.078133406 - 0.137031355
Tunisian	sˤ	u:	0.13319039	0.0264195	0.10001076 - 0.159085743
Tunisian	s	u:	0.1564463	0.03190923	0.104790896 - 0.194410894

9.8.4 Short Vowel: /a/

dialect	condition	vowel_label	mean_vowel duration	sd_vowel duration	range_vowel_duration
Egyptian	$s^ʕ$	a	0.07061613	0.02277377	0.036555565 - 0.11
Egyptian	s	a	0.06656128	0.02199633	0.03494599 - 0.10731002
Iraqi	$s^ʕ$	a	0.08130617	0.01670089	0.061056406 - 0.110850706
Iraqi	s	a	0.07037789	0.02141764	0.040035138 - 0.101481506
Jordanian	$s^ʕ$	a	0.06556025	0.01468356	0.031866174 - 0.091553968
Jordanian	s	a	0.06213659	0.01648121	0.03 - 0.08533734
Kuwaiti	$s^ʕ$	a	0.08071035	0.0228303	0.037390299 - 0.12528581
Kuwaiti	s	a	0.08000158	0.02414078	0.031375925 - 0.134827463
Moroccan	$s^ʕ$	a	0.05429513	0.01336375	0.038947545 - 0.078010844
Moroccan	s	a	0.05198998	0.01594615	0.036768662 - 0.09
Omani	$s^ʕ$	a	0.07652874	0.02339925	0.046655818 - 0.141250294
Omani	s	a	0.07338033	0.01311938	0.051476126 - 0.096354321
Syrian	$s^ʕ$	a	0.06733179	0.01874914	0.049971839 - 0.085085407
Syrian	s	a	0.06759877	0.02329196	0.039531271 - 0.10586651
Tunisian	$s^ʕ$	a	0.061273	0.01906076	0.03 - 0.086052286
Tunisian	s	a	0.06178387	0.01356215	0.04 - 0.081926026

9.8.5 Short Vowel: /i/

dialect	condition	vowel_label	mean_vowel duration	sd_vowel duration	range_vowel_duration
Egyptian	$s^ʕ$	i	0.04113342	0.01679446	0.02563059 - 0.08
Egyptian	s	i	0.0424594	0.01165424	0.023933675 - 0.064917115
Iraqi	$s^ʕ$	i	0.05154332	0.01579856	0.032070444 - 0.077295898
Iraqi	s	i	0.05403393	0.01430907	0.031353795 - 0.078099361
Jordanian	$s^ʕ$	i	0.04015086	0.0068642	0.029231086 - 0.048974235
Jordanian	s	i	0.03836487	0.0130664	0.02281208 - 0.065419217
Kuwaiti	$s^ʕ$	i	0.03850322	0.00661108	0.028660831 - 0.051394418
Kuwaiti	s	i	0.04427602	0.01407277	0.03 - 0.070994788
Moroccan	$s^ʕ$	i	0.05713143	0.01617518	0.03 - 0.091631052
Omani	$s^ʕ$	i	0.0490205	0.01238974	0.027576496 - 0.070595831
Omani	s	i	0.04541203	0.01044068	0.03 - 0.062115627
Syrian	$s^ʕ$	i	0.0384319	0.01186432	0.022571326 - 0.050904169
Syrian	s	i	0.03685675	0.01004788	0.027506704 - 0.047870545
Tunisian	$s^ʕ$	i	0.04960876	0.01009425	0.039219906 - 0.063289761
Tunisian	s	i	0.04103092	0.01192651	0.024594149 - 0.055152993

9.8.6 Short Vowel: /u/

dialect	condition	vowel_label	mean_vowel duration	sd_vowel duration	range_vowel_duration
Egyptian	$s^ʕ$	u	0.08513688	0.02891596	0.042086374 - 0.16
Egyptian	s	u	0.07025656	0.01770272	0.047921822 - 0.114091796
Iraqi	$s^ʕ$	u	0.06412437	0.01484163	0.044612643 - 0.085135382
Iraqi	s	u	0.07241123	0.00785405	0.061417283 - 0.083669132
Jordanian	$s^ʕ$	u	0.06365672	0.01693031	0.033918964 - 0.086884373
Jordanian	s	u	0.06173264	0.01253885	0.045102892 - 0.085310104
Kuwaiti	$s^ʕ$	u	0.0662665	0.00988232	0.05 - 0.079365837
Kuwaiti	s	u	0.06237489	0.0093247	0.046668964 - 0.07404863
Moroccan	$s^ʕ$	u	0.04440335	0.01536952	0.027392653 - 0.076506052
Moroccan	s	u	0.05199277	0.01266151	0.035446369 - 0.077068468
Omani	$s^ʕ$	u	0.1037455	0.04281711	0.073469233 - 0.134021772
Omani	s	u	0.07956816	0.02279027	0.04965812 - 0.115821284
Syrian	$s^ʕ$	u	0.06333198	0.0190658	0.047227303 - 0.095108272
Syrian	s	u	0.05864091	0.01957381	0.038607095 - 0.076213265
Tunisian	$s^ʕ$	u	0.05582665	0.01270183	0.02912996 - 0.077055733
Tunisian	s	u	0.05735871	0.01566788	0.032135382 - 0.080291863

9.9 Mean, standard deviation (SD), and range of F1 values for long and short vowels, grouped by dialect, condition, and vowel label.

9.9.1 F1 for long vowel: /a:/

dialect	condition	vowel_label	mean_f1	sd_f1	range_f1
Egyptian	sˤ	a:	620.11084	76.365243	492.9968858 - 737.1217983
Egyptian	s	a:	586.018865	79.5766743	446.3575317 - 718.971585
Iraqi	sˤ	a:	780.640088	124.729806	591.1106473 - 1008.69692
Iraqi	s	a:	747.692733	114.896233	563.7364266 - 904.0626843
Jordanian	sˤ	a:	761.741529	120.825903	612.9076727 - 1023.892545
Jordanian	s	a:	743.148468	100.446601	616.9776848 - 946.9818979
Kuwaiti	sˤ	a:	775.05609	113.441628	634.1694551 - 1008.167783
Kuwaiti	s	a:	761.341321	97.0819148	631.4205695 - 893.3210031
Moroccan	sˤ	a:	755.260108	103.127287	611.873318 - 986.6716113
Moroccan	s	a:	685.41723	122.510448	537.1677724 - 969.9921201
Omani	sˤ	a:	767.499305	111.805851	591.4291387 - 926.5168949
Omani	s	a:	748.424777	118.27285	560.0537871 - 909.9326495
Syrian	sˤ	a:	683.179962	107.417358	583.6568541 - 917.1119638
Syrian	s	a:	722.380848	102.924876	610.0726088 - 875.6873616
Tunisian	sˤ	a:	709.725595	121.040987	494.894419 - 912.1811302
Tunisian	s	a:	614.25379	121.403891	423.1142531 - 843.8093487

9.9.2 F1 for long vowel: /i:/

dialect	condition	vowel_label	mean_f1	sd_f1	range_f1
Egyptian	sˤ	i:	348.705488	53.2114006	285.5643097 - 509.904419
Egyptian	s	i:	323.503572	26.4679449	282.1112827 - 369.5781066
Iraqi	sˤ	i:	486.416906	90.3473566	386.3132376 - 754.2540687
Iraqi	s	i:	397.921251	56.9599219	314.2222871 - 518.6281922
Jordanian	sˤ	i:	456.232471	107.729509	321.235708 - 740.258128
Jordanian	s	i:	446.900924	110.894566	318.7397078 - 686.2284378
Kuwaiti	sˤ	i:	443.044057	55.1631246	352.7896944 - 594.0871968
Kuwaiti	s	i:	399.045421	57.8245072	308.7685775 - 512.422545
Moroccan	sˤ	i:	515.716443	76.7530741	402.1741779 - 719.4983573
Moroccan	s	i:	396.642612	78.2394564	286.1399073 - 535.7811101
Omani	sˤ	i:	440.347467	44.2813271	332.1062275 - 494.5670352
Omani	s	i:	392.553937	72.3209752	301.8396114 - 533.3400211
Syrian	sˤ	i:	399.433985	46.7064789	323.7432905 - 478.5177419
Syrian	s	i:	365.502254	47.0291183	278.1695381 - 444.1879632
Tunisian	sˤ	i:	409.463849	91.9646678	308.2969408 - 608.7831401
Tunisian	s	i:	391.326035	71.3087933	308.0955092 - 539.2192134

9.9.3 F1 for long vowel: /u:/

dialect	condition	vowel_label	mean_f1	sd_f1	range_f1
Egyptian	sˤ	u:	454.610715	69.2685782	405.6304335 - 503.5909963
Egyptian	s	u:	409.717223	63.7179052	300.3381877 - 527.5304563
Iraqi	sˤ	u:	491.624426	47.74517	435.1819062 - 565.4997188
Iraqi	s	u:	460.570683	60.7715484	389.8034333 - 593.1049003
Jordanian	sˤ	u:	421.784712	41.0229884	337.8890831 - 479.5210356
Jordanian	s	u:	434.154554	55.3408583	379.3551917 - 587.7810362
Kuwaiti	sˤ	u:	608.263917	NA	608.263917 - 608.263917
Kuwaiti	s	u:	456.801835	45.7839978	371.3430048 - 518.0150453
Moroccan	sˤ	u:	494.512878	39.6891302	459.924784 - 558.4389341
Moroccan	s	u:	542.607845	48.199958	452.1908782 - 613.659878
Omani	sˤ	u:	504.819642	81.2821159	430.7178363 - 641.3439595
Omani	s	u:	455.06045	51.8568971	382.3460408 - 547.0106777
Syrian	sˤ	u:	422.07357	48.6988031	362.4152909 - 474.6649993
Syrian	s	u:	399.472794	39.2093873	349.3019275 - 455.8882109
Tunisian	sˤ	u:	452.078239	46.8886683	407.186909 - 535.3940967
Tunisian	s	u:	422.682044	66.6307435	354.0575627 - 555.0386362

9.9.4 F1 for short vowel: /a/

dialect	condition	vowel_label	mean_f1	sd_f1	range_f1
Egyptian	$s^ʕ$	a	609.402874	90.6270228	513.0194912 - 789.5453467
Egyptian	s	a	599.529501	107.931776	393.2635055 - 798.0415202
Iraqi	$s^ʕ$	a	664.247084	108.002153	517.2174901 - 804.5808332
Iraqi	s	a	575.7501	123.240736	431.3788232 - 775.9236036
Jordanian	$s^ʕ$	a	654.617983	69.6296891	555.7920983 - 786.8522185
Jordanian	s	a	617.240285	104.440477	488.3685915 - 846.0622448
Kuwaiti	$s^ʕ$	a	629.098457	86.6928788	488.2572617 - 776.4144175
Kuwaiti	s	a	614.392176	78.5379757	524.8910606 - 791.1616363
Moroccan	$s^ʕ$	a	565.514659	87.7290745	414.8900612 - 743.3239168
Moroccan	s	a	471.801091	89.5235147	368.9769186 - 652.2540451
Omani	$s^ʕ$	a	644.153772	102.263112	472.4950833 - 780.7736933
Omani	s	a	621.906135	99.5900077	502.676022 - 815.5106848
Syrian	$s^ʕ$	a	571.38475	60.5799285	508.9724771 - 654.1691277
Syrian	s	a	571.920859	77.7219211	499.6830985 - 716.5284707
Tunisian	$s^ʕ$	a	612.038938	116.993749	397.3672204 - 795.9184497
Tunisian	s	a	561.552082	84.1988634	412.1210936 - 692.7989552

9.9.5 F1 for short vowel: /i/

dialect	condition	vowel_label	mean_f1	sd_f1	range_f1
Egyptian	sˤ	i	428.119956	50.9645136	345.8074609 - 496.4036648
Egyptian	s	i	526.373444	376.537513	365.9091703 - 1656.49245
Iraqi	sˤ	i	530.856807	88.2878743	430.3066677 - 742.4351514
Iraqi	s	i	421.860298	67.0308912	328.2753308 - 556.2862688
Jordanian	sˤ	i	513.615187	29.3107189	482.4412658 - 583.5932386
Jordanian	s	i	439.002867	35.0939933	349.1264856 - 473.1474994
Kuwaiti	sˤ	i	509.955473	93.0554204	386.9293098 - 753.5004617
Kuwaiti	s	i	513.069772	207.052391	410.6468622 - 1058.903674
Moroccan	sˤ	i	520.082791	110.623869	413.0179391 - 740.8774965
Omani	sˤ	i	549.597928	68.0288997	432.3850792 - 667.5198305
Omani	s	i	463.099843	31.1832619	431.4064477 - 515.5146928
Syrian	sˤ	i	418.608195	81.6172135	312.8469891 - 488.8913006
Syrian	s	i	435.965656	57.9735132	394.321522 - 518.9311577
Tunisian	sˤ	i	554.302623	70.9174924	503.3726968 - 658.5007404
Tunisian	s	i	419.782228	66.6565032	324.8047499 - 518.0286981

9.9.6 F1 for short vowel: /u/

dialect	condition	vowel_label	mean_f1	sd_f1	range_f1
Egyptian	$s^ʕ$	u	438.775808	65.0131837	338.1969431 - 571.4903326
Egyptian	s	u	430.340458	60.7481794	328.3361573 - 580.0518478
Iraqi	$s^ʕ$	u	571.981274	53.6990629	499.5948301 - 648.6533351
Iraqi	s	u	536.744926	81.0672337	420.5600434 - 665.3943807
Jordanian	$s^ʕ$	u	522.322908	33.0640832	491.9404628 - 585.5212753
Jordanian	s	u	508.269852	79.1164689	414.9281512 - 678.7248332
Kuwaiti	$s^ʕ$	u	537.591996	54.4474685	467.6411296 - 637.9960887
Kuwaiti	s	u	496.376173	47.6609244	449.0036943 - 602.2044994
Moroccan	$s^ʕ$	u	597.528114	71.4217662	500.3895794 - 710.0437895
Moroccan	s	u	453.908573	91.0722585	371.6453283 - 639.251671
Omani	$s^ʕ$	u	646.49378	64.7286267	600.7237294 - 692.2638312
Omani	s	u	528.978203	68.7312809	465.7551765 - 641.5756022
Syrian	$s^ʕ$	u	467.291478	77.3217555	394.4425518 - 593.9903974
Syrian	s	u	464.593854	42.0636492	410.1642451 - 499.6285711
Tunisian	$s^ʕ$	u	529.760787	79.1873036	355.904674 - 614.4185116
Tunisian	s	u	508.465042	71.8931483	374.0888422 - 609.2262035

9.10 Mean, standard deviation (SD), and range of F2 values for long and short vowels, grouped by dialect, condition, and vowel label.

9.10.1 F2 for long vowel: /a:/

dialect	condition	vowel_label	mean_f2	sd_f2	range_f2
Egyptian	sˤ	a:	1158.62594	107.6588659	962.107083 - 1375.327074
Egyptian	s	a:	1614.32141	434.6823933	1045.10067 - 2355.514852
Iraqi	sˤ	a:	1347.40745	195.7262793	1065.145958 - 1730.946942
Iraqi	s	a:	1471.7008	226.2089852	1181.05496 - 1870.571852
Jordanian	sˤ	a:	1245.37324	129.8151596	1053.624166 - 1613.140213
Jordanian	s	a:	1498.159	193.6558132	1146.408982 - 1855.487451
Kuwaiti	sˤ	a:	1320.98997	232.5074331	952.4530264 - 1737.615095
Kuwaiti	s	a:	1493.59552	227.2309771	1088.537864 - 1818.430514
Moroccan	sˤ	a:	1351.42154	197.6636991	1079.495162 - 1732.55953
Moroccan	s	a:	1544.78339	234.8600604	1141.523138 - 2030.909683
Omani	sˤ	a:	1309.25476	136.0299672	1121.651471 - 1526.627243
Omani	s	a:	1505.17006	147.5966231	1277.239528 - 1760.419615
Syrian	sˤ	a:	1208.91769	145.261335	1071.838954 - 1478.665839
Syrian	s	a:	1556.88707	187.7486559	1290.269638 - 1788.918664
Tunisian	sˤ	a:	1385.61098	151.5520228	1099.899019 - 1629.852583
Tunisian	s	a:	1691.73872	307.9434082	1202.547685 - 2257.837681

9.10.2 **F2 for long vowel: /i:/**

dialect	condition	vowel_label	mean_f2	sd_f2	range_f2
Egyptian	$s^ʕ$	i:	2306.94473	277.547884	1537.6182 - 2777.276578
Egyptian	s	i:	2243.70607	319.1369734	1539.400435 - 2823.807996
Iraqi	$s^ʕ$	i:	2083.06269	459.6559868	1287.545599 - 3089.877815
Iraqi	s	i:	2129.24642	593.175452	837.0840625 - 2929.037055
Jordanian	$s^ʕ$	i:	2468.37859	312.95354	1821.749774 - 2860.449149
Jordanian	s	i:	2467.51869	248.7055021	2057.227516 - 2839.649522
Kuwaiti	$s^ʕ$	i:	2130.31813	409.1648998	1197.991045 - 2771.449348
Kuwaiti	s	i:	2324.23766	361.3296172	1709.213198 - 2874.36758
Moroccan	$s^ʕ$	i:	2145.62938	261.477138	1788.034093 - 2628.443855
Moroccan	s	i:	2351.79788	229.2081063	2032.915186 - 2765.367309
Omani	$s^ʕ$	i:	2322.94159	339.2941849	1708.411436 - 2864.817384
Omani	s	i:	2364.98075	322.6036417	1624.312454 - 2900.99903
Syrian	$s^ʕ$	i:	2284.19508	543.8878476	1004.187585 - 2832.643232
Syrian	s	i:	2338.36842	475.0841685	1054.706142 - 2836.455413
Tunisian	$s^ʕ$	i:	2288.0942	328.7230055	1629.875553 - 2908.988093
Tunisian	s	i:	2233.99411	281.6221738	1588.700651 - 2751.594944

9.10.3 **F2 for long vowel: /u:/**

dialect	condition	vowel_label	mean_f2	sd_f2	range_f2
Egyptian	$s^ʕ$	u:	1062.05812	37.37615761	1035.629183 - 1088.487052
Egyptian	s	u:	831.634733	116.4750214	702.3457201 - 1120.228674
Iraqi	$s^ʕ$	u:	972.283739	239.1587628	656.2488345 - 1355.646724
Iraqi	s	u:	1126.14951	223.6282868	902.4141279 - 1518.156404
Jordanian	$s^ʕ$	u:	880.621196	86.55677483	723.8531077 - 984.2825757
Jordanian	s	u:	1021.17564	307.0177975	708.9278387 - 1926.682815
Kuwaiti	$s^ʕ$	u:	1166.29462	NA	1166.294618 - 1166.294618
Kuwaiti	s	u:	1156.2596	239.4185834	866.1154529 - 1707.218199
Moroccan	$s^ʕ$	u:	835.313885	96.19528865	719.9140568 - 942.7759529
Moroccan	s	u:	941.263774	116.5633081	756.166766 - 1125.551331
Omani	$s^ʕ$	u:	964.687425	133.0965269	805.3828155 - 1136.775417
Omani	s	u:	993.193812	66.40998165	920.7099562 - 1098.328722
Syrian	$s^ʕ$	u:	921.828828	134.0348198	793.2287333 - 1057.527367
Syrian	s	u:	1060.76303	165.2433309	858.2031331 - 1274.999798
Tunisian	$s^ʕ$	u:	963.881631	164.1069135	771.0126974 - 1210.885679
Tunisian	s	u:	941.33567	106.9003324	815.6886235 - 1117.084103

9.10.4 **F2 for short vowel: /a/**

dialect	condition	vowel_label	mean_f2	sd_f2	range_f2
Egyptian	sˤ	a	1217.27741	87.4932189	1076.821318 - 1395.100614
Egyptian	s	a	1653.85673	173.305505	1331.382777 - 1961.881827
Iraqi	sˤ	a	1430.1488	243.409168	1113.477614 - 1770.852633
Iraqi	s	a	1680.22664	269.471294	1367.289805 - 2165.143056
Jordanian	sˤ	a	1255.79083	116.613528	1098.422683 - 1501.267048
Jordanian	s	a	1673.54625	149.580328	1441.989323 - 2015.390248
Kuwaiti	sˤ	a	1466.82851	282.474778	1084.697554 - 1924.407142
Kuwaiti	s	a	1722.35333	205.313187	1357.632205 - 1964.543741
Moroccan	sˤ	a	1321.73779	133.603771	1113.601679 - 1531.665559
Moroccan	s	a	1718.16551	113.525108	1572.579509 - 1845.520812
Omani	sˤ	a	1405.24438	165.990779	1146.02709 - 1704.123568
Omani	s	a	1744.01497	136.891499	1470.92276 - 1998.99015
Syrian	sˤ	a	1124.70198	42.3552554	1076.670448 - 1173.379226
Syrian	s	a	1680.3667	175.095352	1512.589981 - 1910.279347
Tunisian	sˤ	a	1418.83561	336.680026	1023.740359 - 2329.268879
Tunisian	s	a	1592.335	199.331256	1293.5041 - 1878.226223

9.10.5 **F2 for short vowel: /i/**

dialect	condition	vowel_label	mean_f2	sd_f2	range_f2
Egyptian	sˤ	i	1490.58026	178.650752	1092.061385 - 1757.882093
Egyptian	s	i	1965.23425	259.196964	1646.937177 - 2436.000118
Iraqi	sˤ	i	1222.48875	203.663073	955.8301981 - 1601.016433
Iraqi	s	i	1838.56847	272.046789	1516.679903 - 2167.633752
Jordanian	sˤ	i	1364.07584	209.857259	1176.163327 - 1718.151818
Jordanian	s	i	1866.12313	154.065974	1626.466772 - 2095.17172
Kuwaiti	sˤ	i	1339.69953	307.243397	1026.981658 - 1974.835082
Kuwaiti	s	i	1735.1557	207.748346	1430.293001 - 1973.994803
Moroccan	sˤ	i	1312.15623	179.427284	1051.576284 - 1587.123637
Omani	sˤ	i	1197.30928	130.490929	1032.102418 - 1430.754618
Omani	s	i	1794.99538	188.856471	1491.487989 - 2036.637725
Syrian	sˤ	i	1136.53507	78.0942555	1023.690117 - 1194.517204
Syrian	s	i	1841.62773	359.89647	1445.942125 - 2218.5109
Tunisian	sˤ	i	1260.52132	125.335528	1161.052389 - 1434.222925
Tunisian	s	i	1787.14724	281.058805	1473.746734 - 2078.671589

9.10.6 **F2 for short vowel: /u/**

dialect	condition	vowel_label	mean_f2	sd_f2	range_f2
Egyptian	sˤ	u	1015.06161	125.100593	855.0026665 - 1276.830001
Egyptian	s	u	1264.35645	203.679216	1021.968222 - 1775.115739
Iraqi	sˤ	u	1135.59273	173.647837	933.3054027 - 1449.493451
Iraqi	s	u	1405.24527	171.959513	1077.518281 - 1589.091185
Jordanian	sˤ	u	1007.83436	53.3861265	950.0992044 - 1111.124549
Jordanian	s	u	1325.51513	160.963965	1114.369572 - 1624.854455
Kuwaiti	sˤ	u	1057.0806	191.071603	817.6973153 - 1318.932604
Kuwaiti	s	u	1383.57955	167.615362	1211.274028 - 1685.062266
Moroccan	sˤ	u	1198.78724	140.894788	1019.846273 - 1500.135402
Moroccan	s	u	1287.30531	201.596837	1061.049535 - 1582.602681
Omani	sˤ	u	1089.24804	47.6325817	1055.566722 - 1122.929365
Omani	s	u	1517.34939	150.397485	1286.444023 - 1738.822613
Syrian	sˤ	u	971.58992	153.157561	778.1254156 - 1142.801338
Syrian	s	u	1366.79463	428.954241	1007.236959 - 1973.515526
Tunisian	sˤ	u	1179.69718	142.227291	961.885432 - 1471.949905
Tunisian	s	u	1254.14688	163.65953	1045.963755 - 1495.786856

9.11 Mean, standard deviation (SD), and range of F3 values for long and short vowels, grouped by dialect, condition, and vowel label.

9.11.1 F3 for long vowel: /a:/

dialect	condition	vowel_label	mean_f3	sd_f3	range_f3
Egyptian	sˤ	a:	2739.42294	207.8695787	2446.809014 - 3180.341886
Egyptian	s	a:	2688.73823	206.8883767	2418.928735 - 3019.284283
Iraqi	sˤ	a:	2622.22042	437.8032156	1609.601416 - 3222.263002
Iraqi	s	a:	2389.26986	445.2955959	1354.482654 - 3098.715136
Jordanian	sˤ	a:	2631.15432	270.3899326	2097.528013 - 3111.992603
Jordanian	s	a:	2548.13155	299.0568251	1892.844631 - 2903.408577
Kuwaiti	sˤ	a:	2658.75572	309.2183502	2093.061802 - 3313.730495
Kuwaiti	s	a:	2665.16576	279.9131578	2122.142981 - 3134.82445
Moroccan	sˤ	a:	2682.22696	254.9907347	2190.090698 - 3213.20731
Moroccan	s	a:	2607.07051	269.7322542	1855.234487 - 3072.607936
Omani	sˤ	a:	2694.54188	342.7804046	1747.884561 - 3200.724176
Omani	s	a:	2658.28384	277.4001698	1863.322582 - 3106.704968
Syrian	sˤ	a:	2744.29156	244.3378153	2332.806319 - 3011.739493
Syrian	s	a:	2643.05902	149.9775993	2364.182114 - 2817.518438
Tunisian	sˤ	a:	2739.31058	310.9346495	1998.619632 - 3117.988853
Tunisian	s	a:	2743.50821	287.9629735	2228.858475 - 3274.526132

9.11.2 **F3 for long vowel: /i:/**

dialect	condition	vowel_label	mean_f3	sd_f3	range_f3
Egyptian	sˤ	i:	2909.15875	316.2246733	2426.057527 - 3681.803185
Egyptian	s	i:	2912.10327	218.9774553	2636.310816 - 3296.595926
Iraqi	sˤ	i:	2819.68321	340.227409	2317.324597 - 3334.750776
Iraqi	s	i:	2872.85468	369.8280556	2256.679877 - 3565.311677
Jordanian	sˤ	i:	3008.64931	264.3723755	2545.694127 - 3461.210238
Jordanian	s	i:	3046.9394	237.6622922	2708.213681 - 3570.266826
Kuwaiti	sˤ	i:	2723.96867	348.6807887	2106.271359 - 3295.329436
Kuwaiti	s	i:	2768.29608	378.0661968	2240.242908 - 3425.988159
Moroccan	sˤ	i:	2859.54237	269.822531	2449.883633 - 3611.285543
Moroccan	s	i:	2948.96651	279.320483	2467.153887 - 3323.706434
Omani	sˤ	i:	2941.87099	268.2382852	2432.656201 - 3359.60073
Omani	s	i:	2950.50948	264.7131864	2546.18431 - 3442.864677
Syrian	sˤ	i:	2962.78057	251.5868662	2539.428982 - 3332.103793
Syrian	s	i:	2898.29185	285.7457641	2461.970776 - 3365.329463
Tunisian	sˤ	i:	2997.96729	399.3756701	2426.064681 - 3660.894649
Tunisian	s	i:	2921.50403	360.1758322	2467.656786 - 3755.829853

9.11.3 **F3 for long vowel: /u:/**

dialect	condition	vowel_label	mean_f3	sd_f3	range_f3
Egyptian	sˁ	u:	2485.68908	84.7208259	2425.782408 - 2545.595749
Egyptian	s	u:	2611.41445	275.2065619	2258.972054 - 2988.819669
Iraqi	sˁ	u:	2843.37512	181.4497583	2644.603373 - 3083.258722
Iraqi	s	u:	2578.30775	277.4874429	2181.062911 - 2904.538136
Jordanian	sˁ	u:	2743.75884	226.3792387	2429.799832 - 3151.745759
Jordanian	s	u:	2709.96683	245.270027	2364.784586 - 3279.4891
Kuwaiti	sˁ	u:	3260.55562	NA	3260.555616 - 3260.555616
Kuwaiti	s	u:	2621.9236	350.0822828	1942.425527 - 3126.394629
Moroccan	sˁ	u:	2755.9143	228.6721607	2496.410293 - 3053.851332
Moroccan	s	u:	2793.48682	259.0419675	2424.775523 - 3305.41019
Omani	sˁ	u:	2754.42252	312.3595148	2193.177885 - 3103.958246
Omani	s	u:	2721.85248	265.1275174	2207.782053 - 3119.840435
Syrian	sˁ	u:	2861.27627	280.4643241	2455.19992 - 3083.29125
Syrian	s	u:	2680.94863	220.6808954	2340.2691 - 2913.565714
Tunisian	sˁ	u:	2744.63011	253.9349378	2405.867557 - 3042.125582
Tunisian	s	u:	2748.12095	227.2903713	2280.170424 - 3100.980279

9.11.4 **F3 for short vowel: /a/**

dialect	condition	vowel_label	mean_f3	sd_f3	range_f3
Egyptian	sˤ	a	2745.962	206.2134571	2473.856749 - 3149.555775
Egyptian	s	a	2716.91078	223.4100292	2385.508467 - 3083.189732
Iraqi	sˤ	a	2667.63158	414.0425496	1880.304791 - 3066.56324
Iraqi	s	a	2547.68049	329.4604094	2027.475529 - 3055.716876
Jordanian	sˤ	a	2703.95756	368.978246	1930.588974 - 3286.509268
Jordanian	s	a	2740.17208	287.8154081	1900.434336 - 3122.97992
Kuwaiti	sˤ	a	2690.76835	416.8718259	1970.024616 - 3373.817282
Kuwaiti	s	a	2671.08875	384.7850183	2099.884251 - 3352.057755
Moroccan	sˤ	a	2890.96641	235.0759114	2462.828643 - 3172.993371
Moroccan	s	a	2807.5123	276.8497438	2351.483258 - 3143.750687
Omani	sˤ	a	2846.29528	265.0000757	2219.006496 - 3386.111942
Omani	s	a	2710.50146	350.4374274	2037.655821 - 3159.102911
Syrian	sˤ	a	2916.00139	518.3002363	2452.438363 - 3489.581114
Syrian	s	a	2665.20586	372.7802005	2166.362351 - 3186.625861
Tunisian	sˤ	a	2870.96384	282.2123693	2285.327413 - 3246.749089
Tunisian	s	a	2598.86405	315.2180298	2002.770161 - 3012.660504

9.11.5 **F3 for short vowel: /i/**

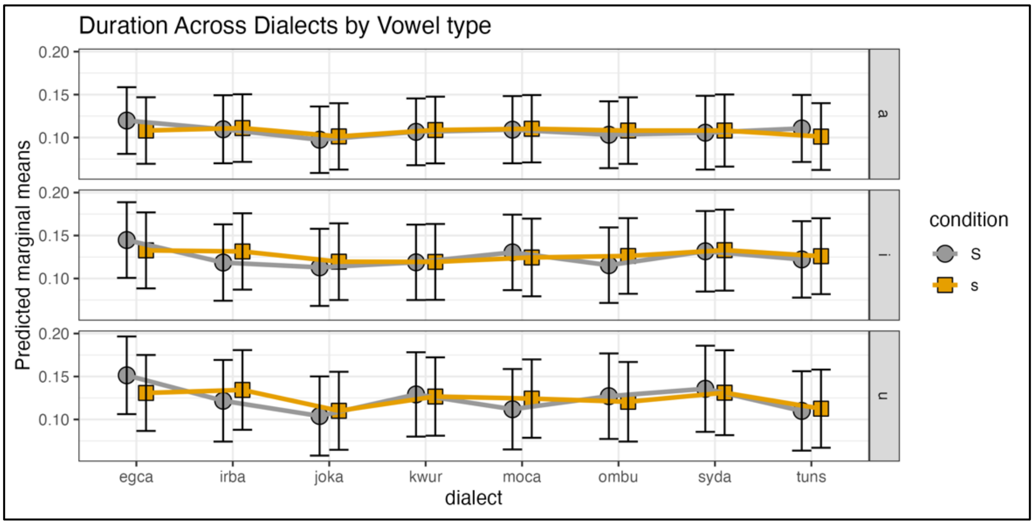
dialect	condition	vowel_label	mean_f3	sd_f3	range_f3
Egyptian	sˤ	i	2763.5714	227.1956487	2396.533431 - 3191.621731
Egyptian	s	i	2782.6993	321.2462151	2421.635238 - 3458.318395
Iraqi	sˤ	i	2828.02196	356.993793	2056.952505 - 3209.913354
Iraqi	s	i	2765.45231	241.9955719	2458.583406 - 3118.734209
Jordanian	sˤ	i	2809.52526	326.4734689	2315.834951 - 3298.832687
Jordanian	s	i	2788.01168	136.302066	2588.577098 - 2989.277216
Kuwaiti	sˤ	i	2818.82985	425.6976923	2090.565258 - 3622.819314
Kuwaiti	s	i	2730.48456	243.629345	2430.014593 - 3101.63407
Moroccan	sˤ	i	2793.23134	331.436762	2188.832032 - 3141.411322
Omani	sˤ	i	2875.60682	168.9006612	2659.684146 - 3134.682716
Omani	s	i	2812.78544	160.9475787	2547.195551 - 3098.066066
Syrian	sˤ	i	2736.52569	582.5199653	2030.491468 - 3352.098134
Syrian	s	i	2742.44569	245.5425118	2449.2328 - 2953.58309
Tunisian	sˤ	i	2687.13017	309.7304049	2276.239137 - 3024.815717
Tunisian	s	i	2779.29674	268.0922971	2307.387944 - 3103.391601

9.11.6 **F3 for short vowel: /u/**

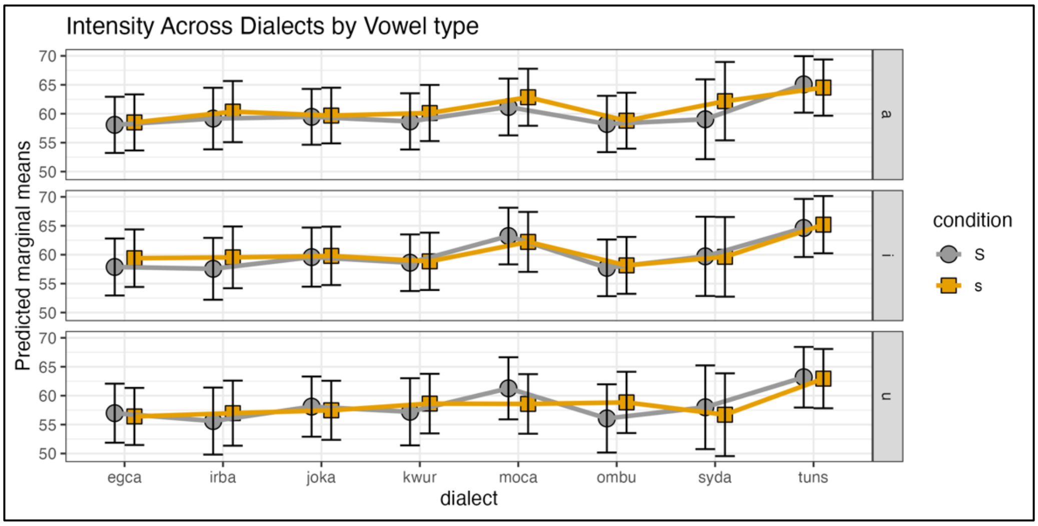
dialect	condition	vowel_label	mean_f3	sd_f3	range_f3
Egyptian	sˤ	u	2799.00678	223.1781279	2328.571702 - 3159.166561
Egyptian	s	u	2618.03384	222.5758386	2303.040132 - 2925.097176
Iraqi	sˤ	u	2782.98447	200.5336954	2437.532679 - 3093.056917
Iraqi	s	u	2627.38487	217.2372791	2275.848085 - 2920.342322
Jordanian	sˤ	u	2749.48878	272.2709595	2433.150972 - 3317.04223
Jordanian	s	u	2621.32672	182.2839961	2391.134433 - 2971.49385
Kuwaiti	sˤ	u	2638.65807	431.1118017	2042.994266 - 3411.092563
Kuwaiti	s	u	2545.33316	300.7917621	2114.986713 - 3153.563915
Moroccan	sˤ	u	2775.72886	200.5370242	2416.633751 - 3012.107616
Moroccan	s	u	2694.89749	168.7779403	2486.928066 - 2933.943042
Omani	sˤ	u	3222.3815	109.909535	3144.663718 - 3300.099273
Omani	s	u	2742.18606	236.1623115	2391.026863 - 3063.35783
Syrian	sˤ	u	2941.37667	249.3624618	2630.527508 - 3209.087231
Syrian	s	u	2643.17281	304.5111246	2365.992574 - 2958.384272
Tunisian	sˤ	u	2779.87896	259.5940388	2235.138704 - 3106.653894
Tunisian	s	u	2655.34289	209.6500308	2316.512484 - 3007.585695

9.12 Appendix shows the predicted mean values for each acoustic measurements by vowel type.

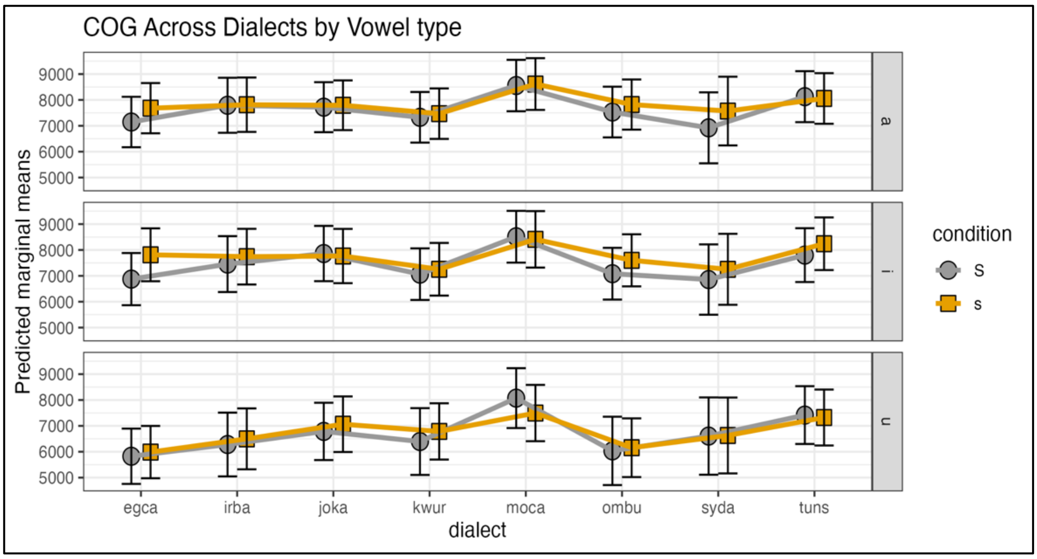
9.12.1 Fricative duration value by vowel type



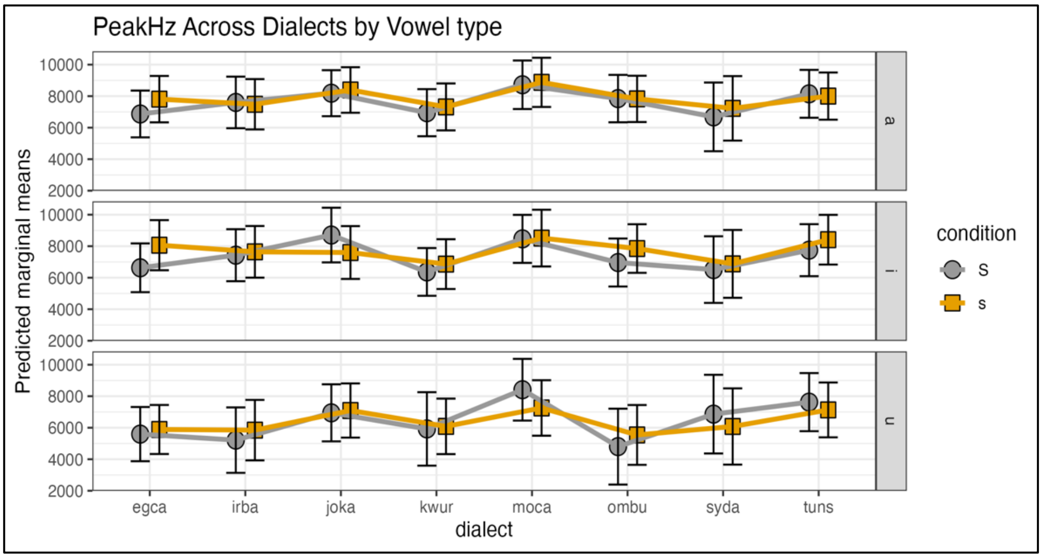
9.12.2 Intensity value by vowel type



9.12.3 COG value by vowel type

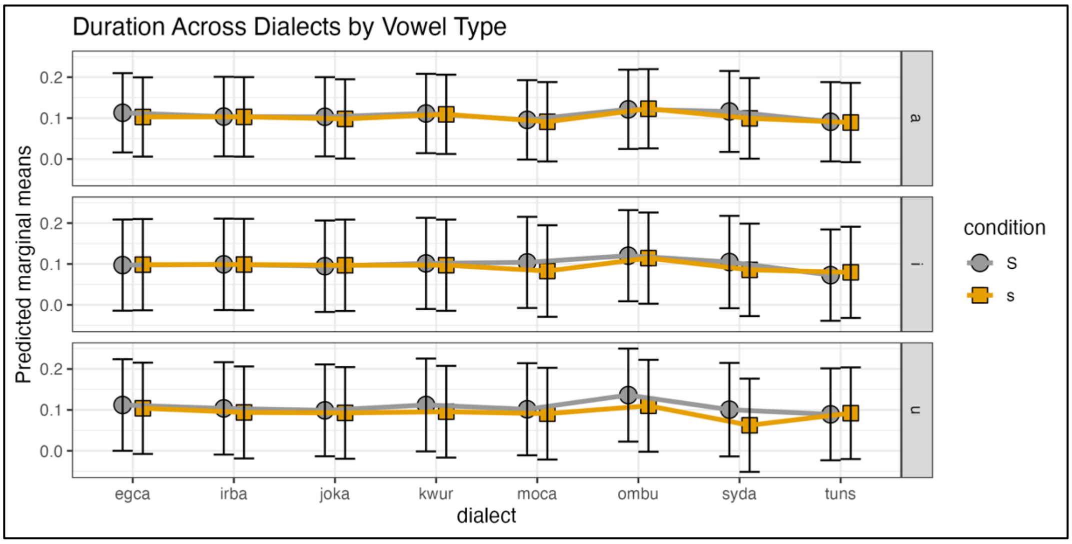


9.12.4 Peak location value by vowel type

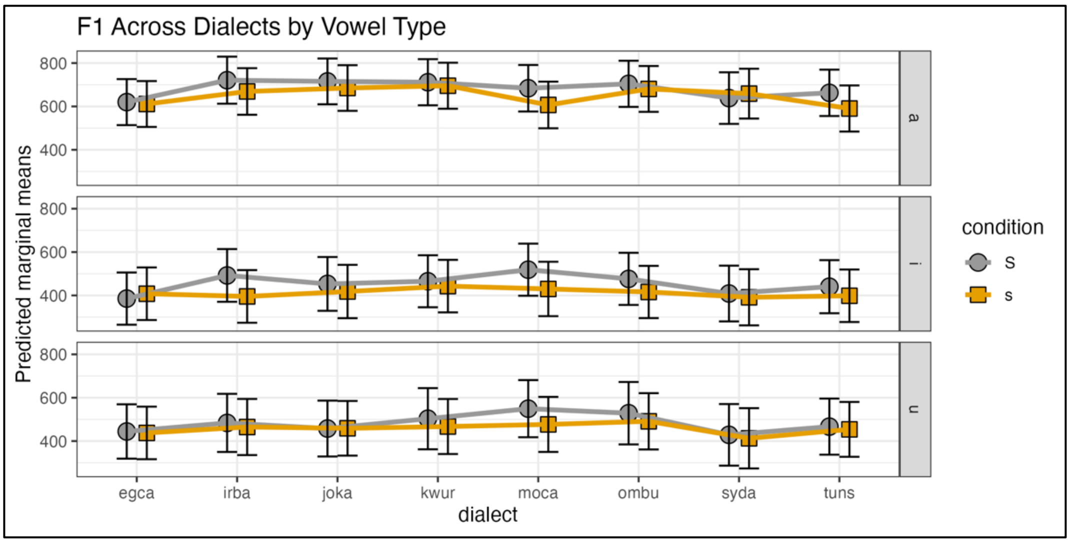


9.13 Appendix shows the predicted mean values for each vowel formant information by vowel type

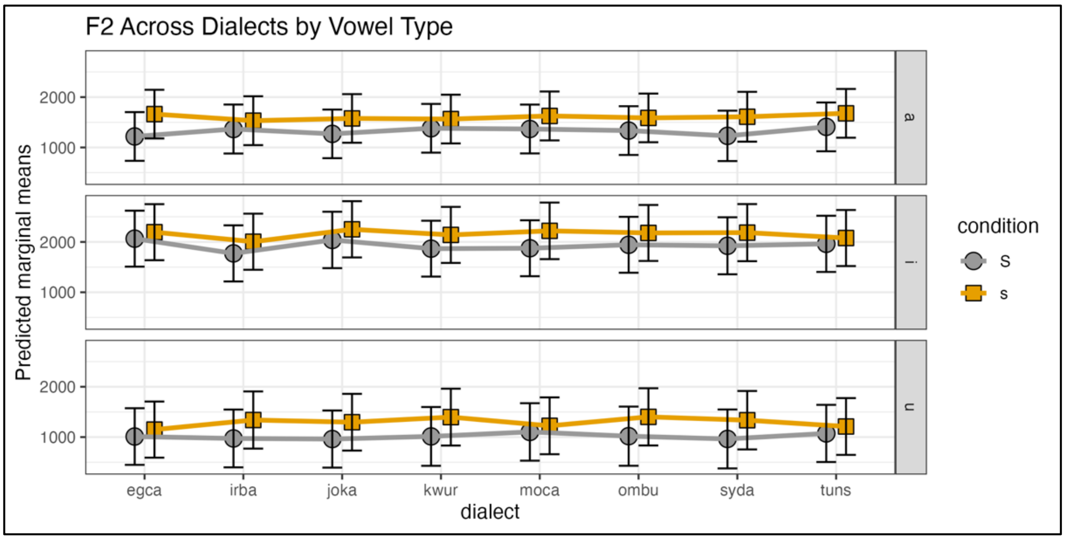
9.13.1 Vowel duration value by vowel type



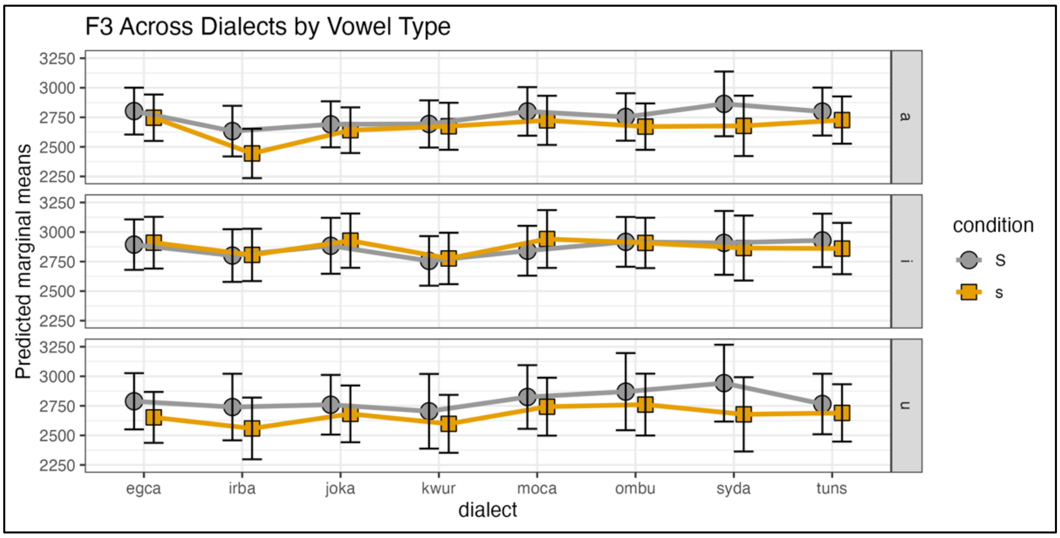
9.13.2 F1 value by vowel type



9.13.3 **F2 value by vowel type**



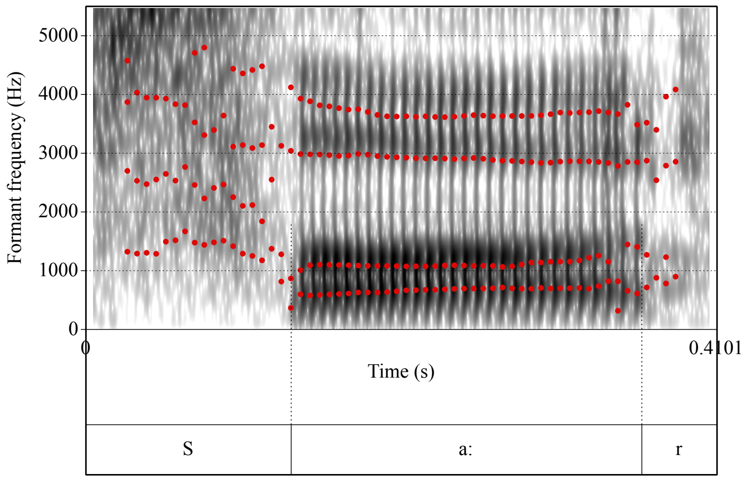
9.13.4 **F3 value by vowel type**



10 Appendix B: (For Perception Study)

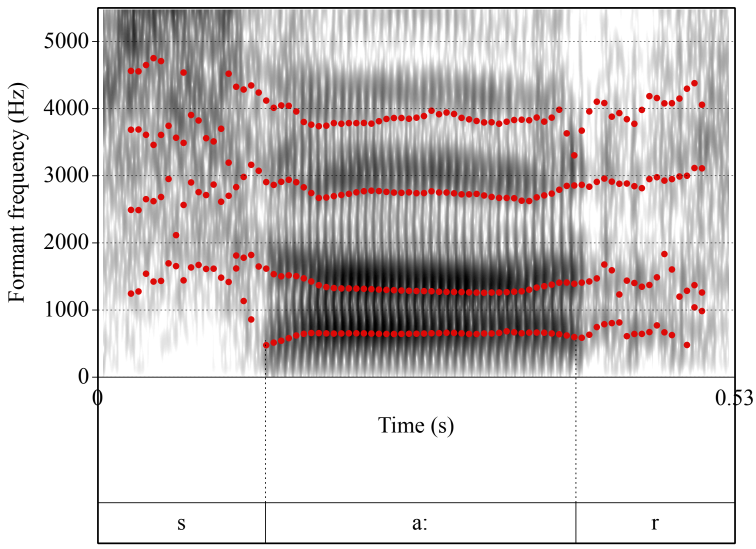
10.1 Raw Spectrograms of F2 for Emphatic and Plain Vowels /a:/ and /i:/ in Syrian Dialect.

10.1.1 Raw Spectrogram of the Emphatic Vowel /a:/ Produced by a Syrian Speaker



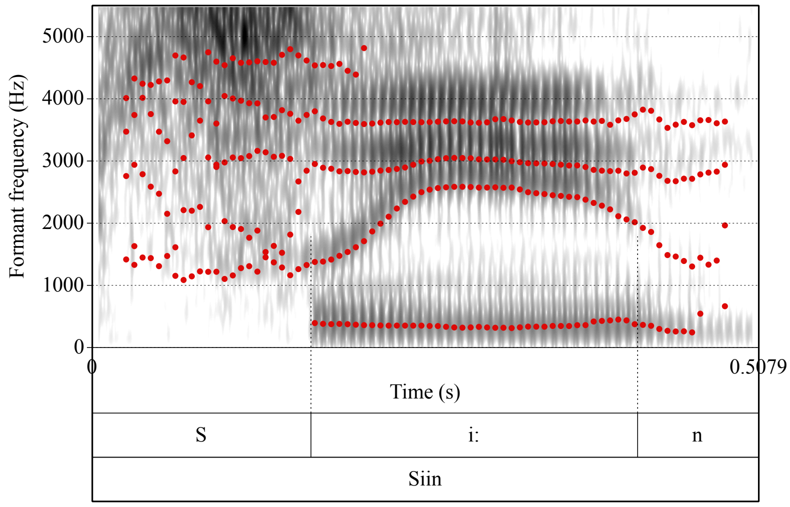
(Speaker ID: syda-ph15-m3)

10.1.2 Raw Spectrogram of the Plain Vowel /a:/ Produced by a Syrian Speaker



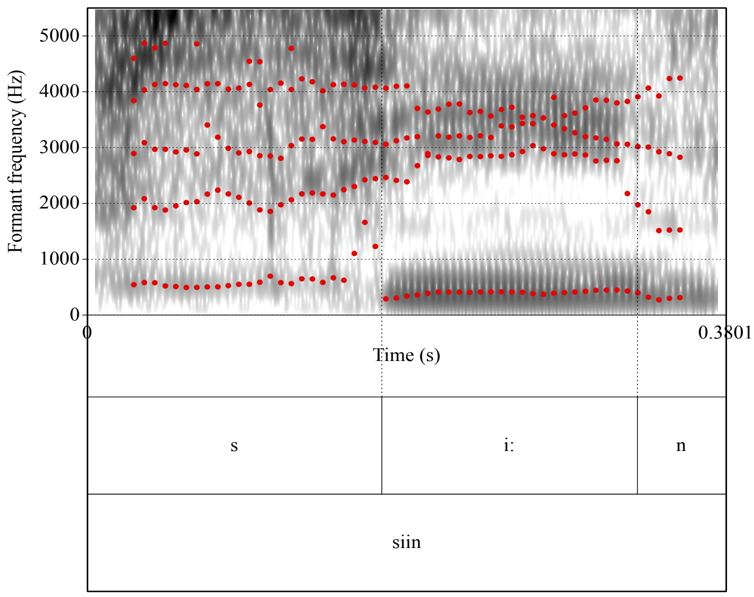
(Speaker ID: syda-pl15-m3)

10.1.3 Raw Spectrogram of the Emphatic Vowel /i:/ Produced by a Syrian Speaker



(Speaker ID: syda-ph13-m4)

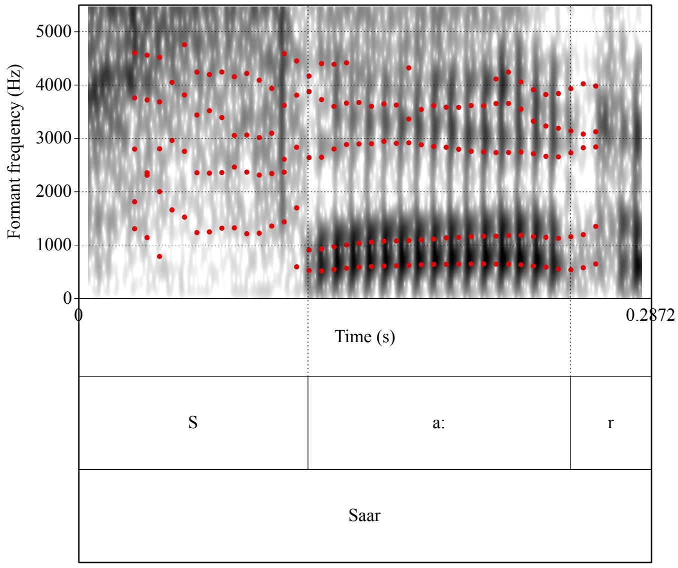
10.1.4 Raw Spectrogram of the Plain Vowel /i:/ Produced by a Syrian Speaker



(Speaker ID: syda-pl13-f2)

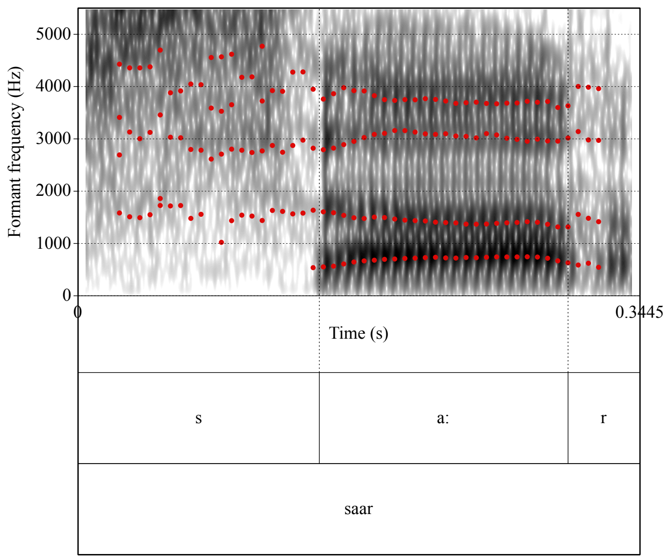
10.2 Raw Spectrograms of F2 for Emphatic and Plain Vowels /a:/ and /i:/ in Iraqi Dialect.

10.2.1 Raw Spectrogram of the Emphatic Vowel /a:/ Produced by an Iraqi Speaker



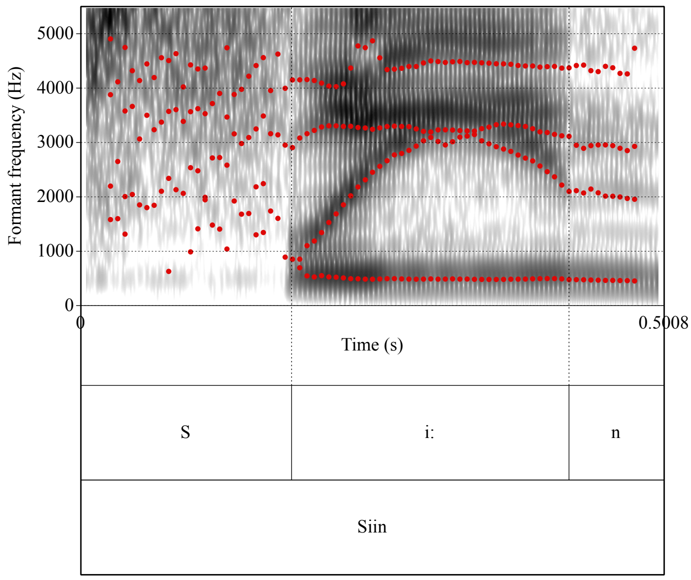
(Speaker ID: irba-ph15-m3)

10.2.2 Raw Spectrogram of the Plain Vowel /a:/ Produced by an Iraqi Speaker



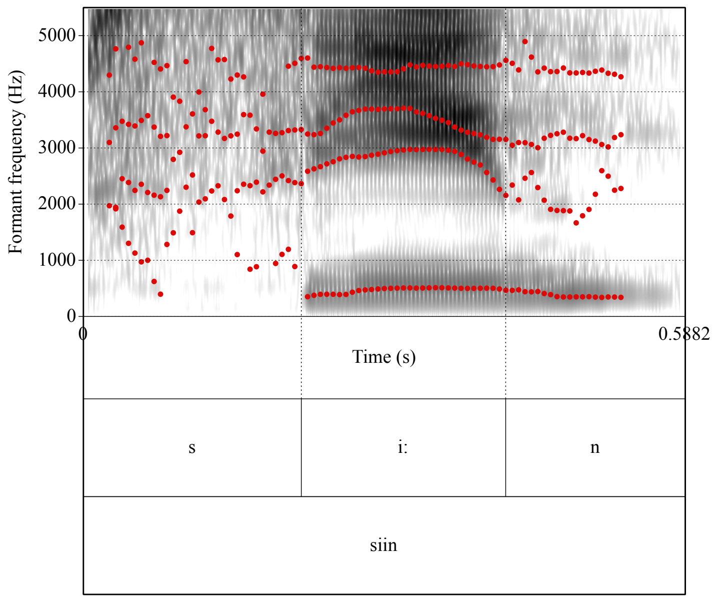
(Speaker ID: irba-ph15-m2)

10.2.3 Raw Spectrogram of the Emphatic Vowel /i:/ Produced by an Iraqi Speaker



(Speaker ID: irba-ph13-f2)

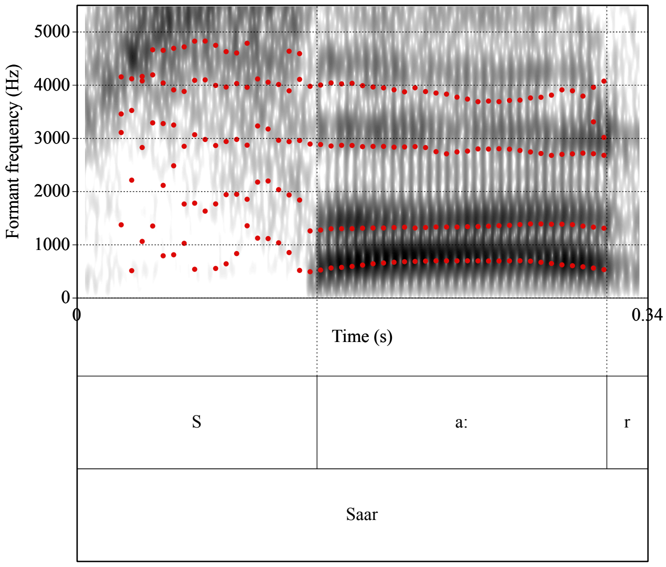
10.2.4 Raw Spectrogram of the Plain Vowel /i:/ Produced by an Iraqi Speaker



(Speaker ID: irba-pl13-f2)

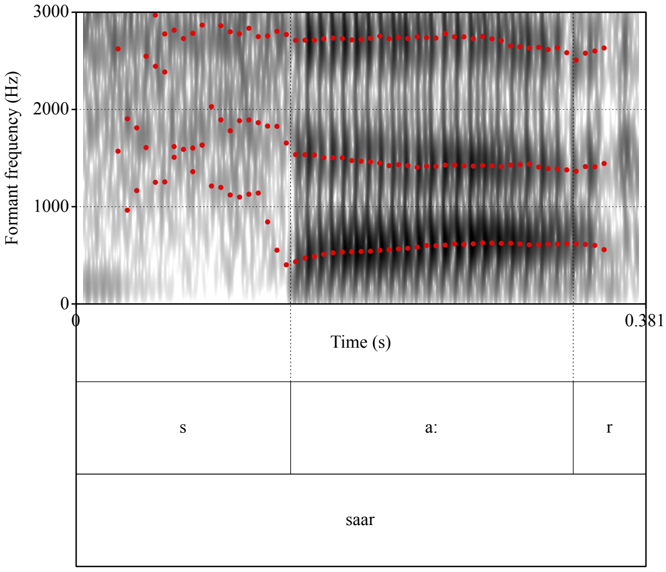
10.3 Raw Spectrograms of F2 for Emphatic and Plain Vowels /a:/ and /i:/ in Moroccan Dialect.

10.3.1 Raw Spectrogram of the Emphatic Vowel /a:/ Produced by a Moroccan Speaker



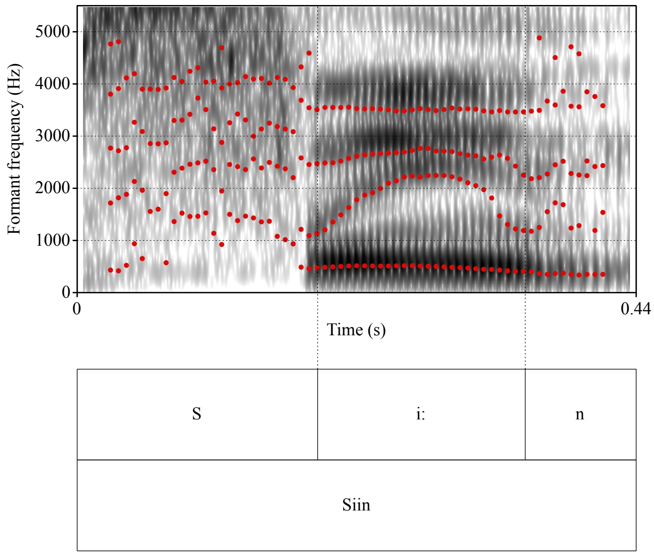
(Speaker ID: moca-ph15-m3)

10.3.2 Raw Spectrogram of the Plain Vowel /a:/ Produced by a Moroccan Speaker



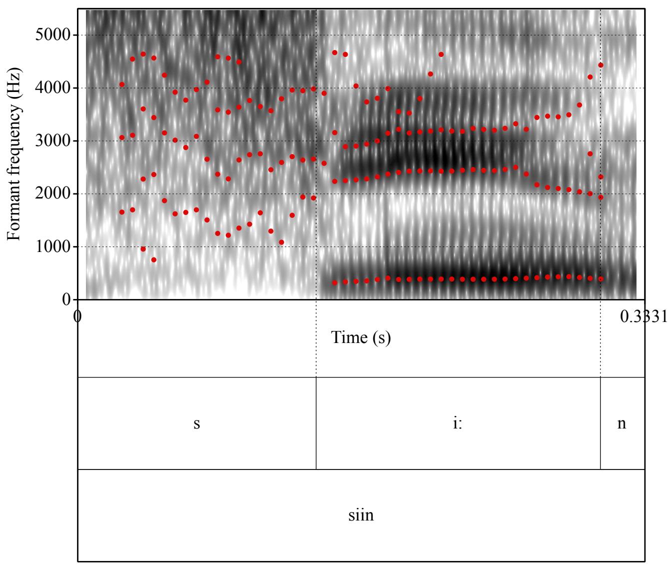
(Speaker ID: moca-pl15-m6)

10.3.3 Raw Spectrogram of the Emphatic Vowel /i:/ Produced by a Moroccan Speaker



(Speaker ID: moca-ph13-m1)

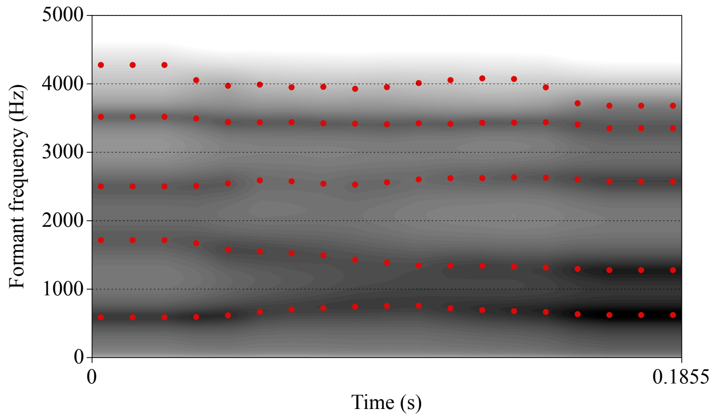
10.3.4 Raw Spectrogram of the Emphatic Vowel /i:/ Produced by a Moroccan Speaker



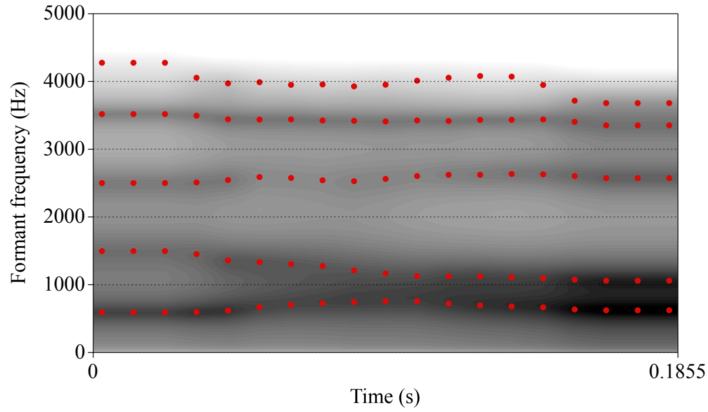
(Speaker ID: moca-pl13-m1)

10.4 Spectrograms of Manipulated F2 Trajectories for the Vowel /a:/ for Three Conditions and the Base Stimuli.

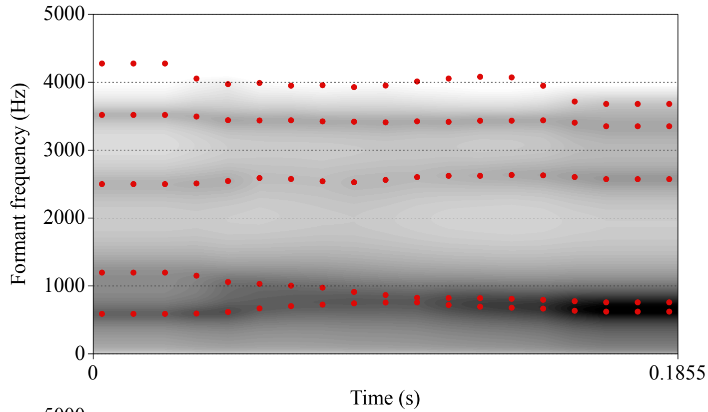
Base Stimuli (plain) for /a:/



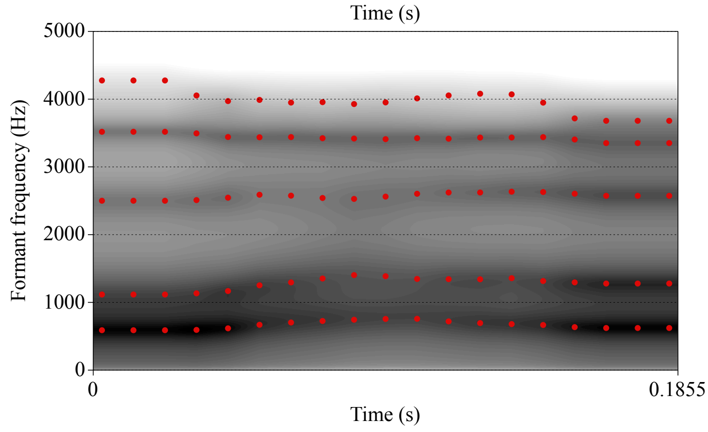
Condition: Medium 100 for /a:/



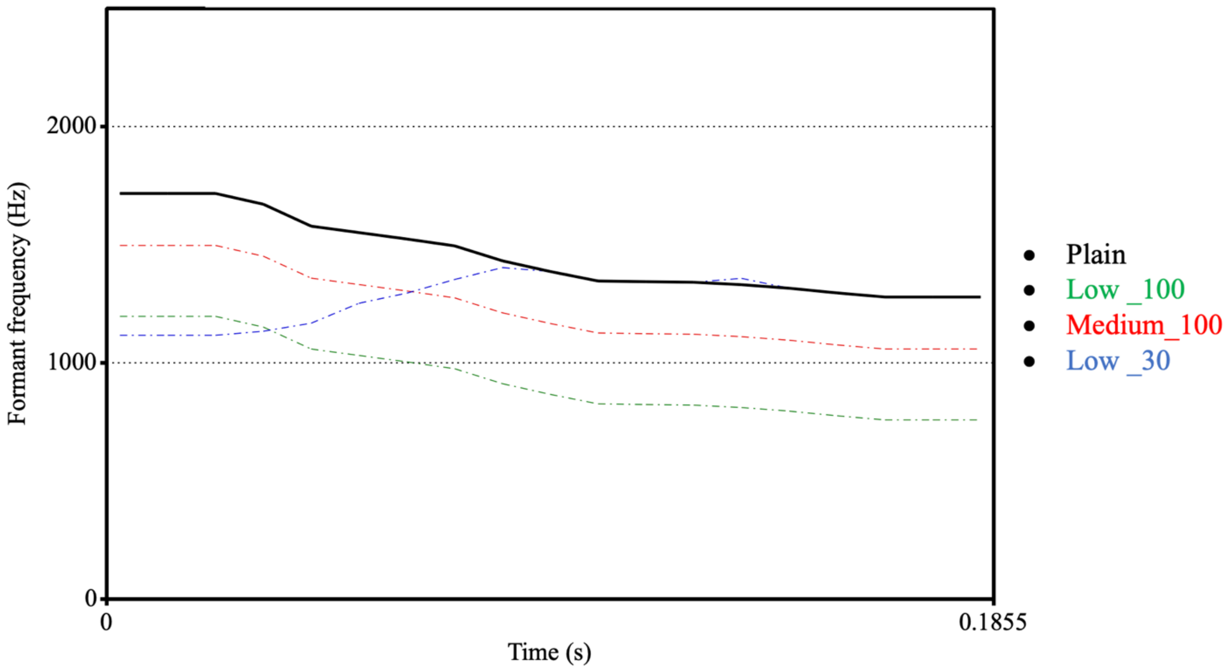
Condition: Low 100 for /a:/



Condition: Low 30 for /a:/

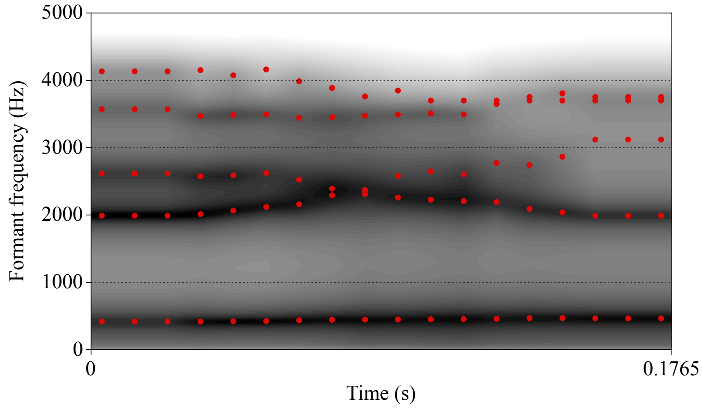


10.5 Plot of Manipulated F2 Trajectories for the Vowel /a:/ for Three Conditions and the Base Stimuli

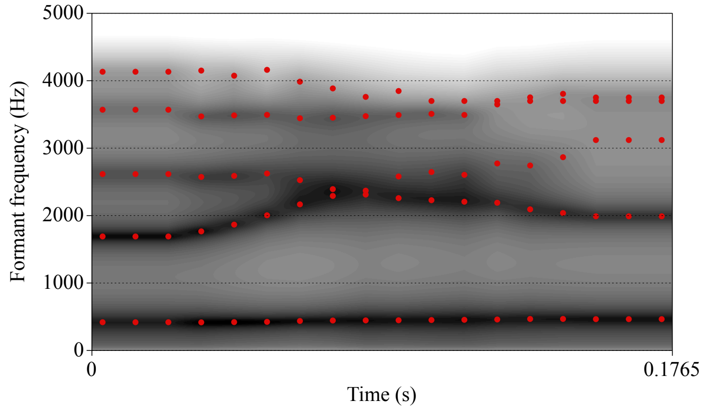


10.6 Spectrograms of Manipulated F2 Trajectories for the Vowel /i:/ for Three Conditions and the Base Stimuli

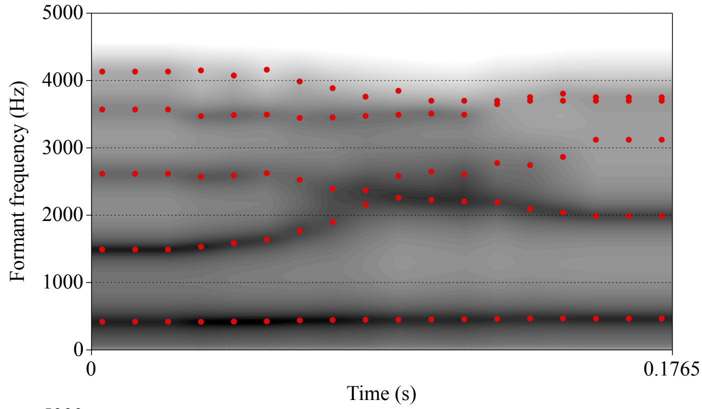
Base Stimuli (plain) for /i:/



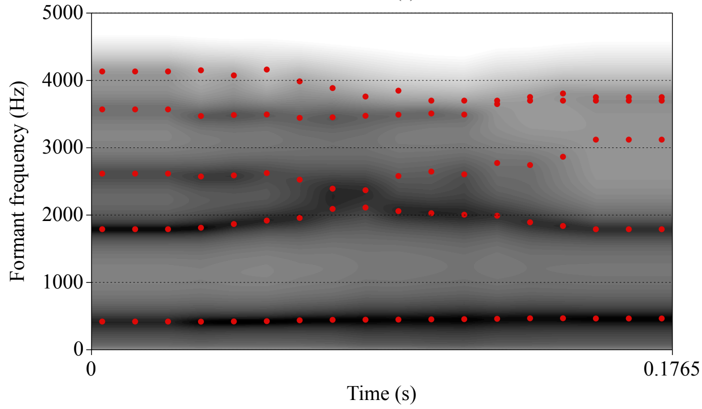
Condition: Low 30 for /i:/



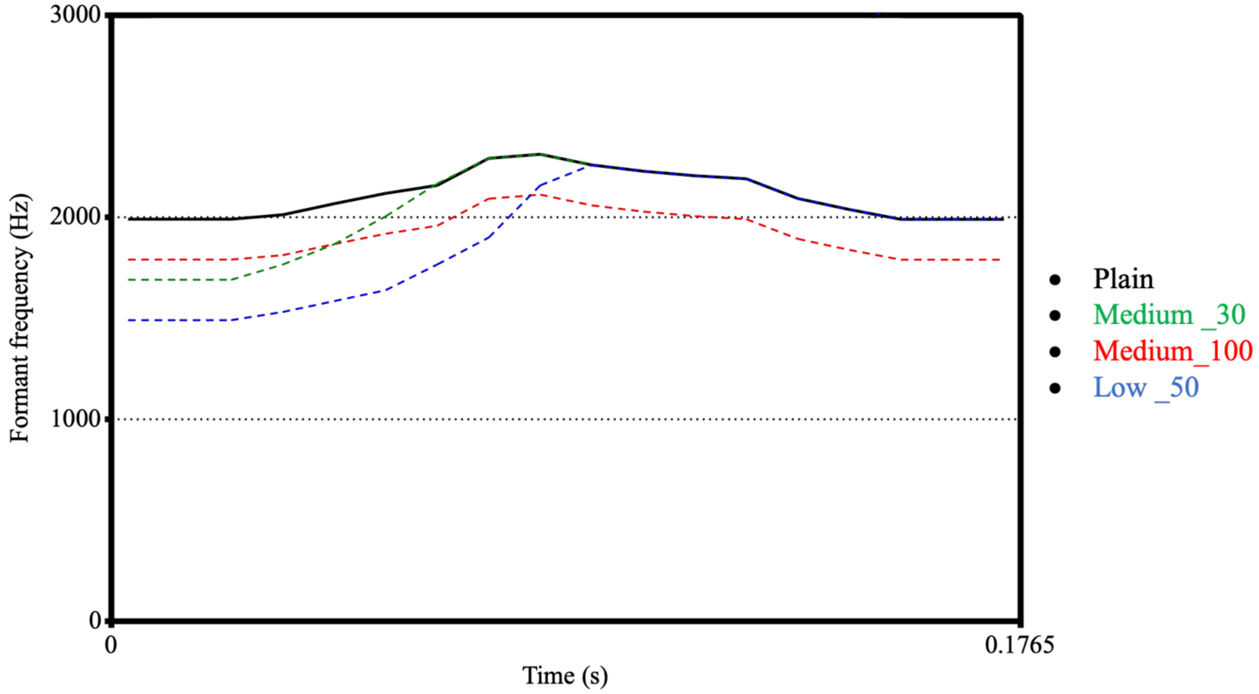
Condition: Low 50 for /i:/



Condition: Medium 100 for /i:/



10.7 Plot of Manipulated F2 Trajectories for the Vowel /i:/ for Three Conditions and the Base Stimuli



10.8 F1 average values across all dialects

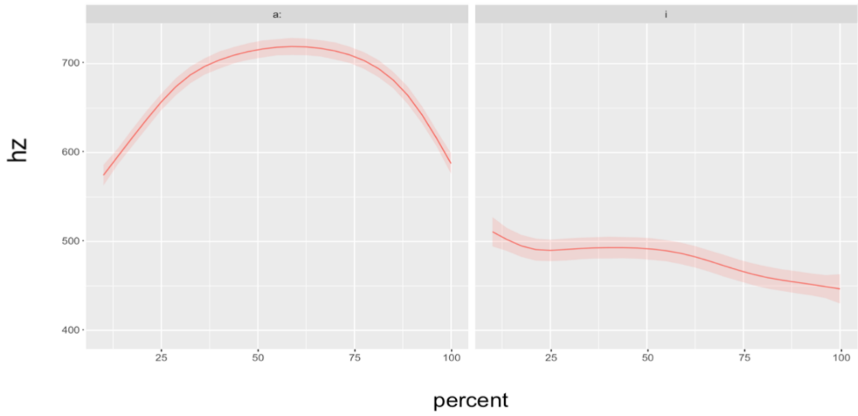


Figure 1 The Average values for F1 in both vowels a: and i:

10.9 F3 average values across all dialects

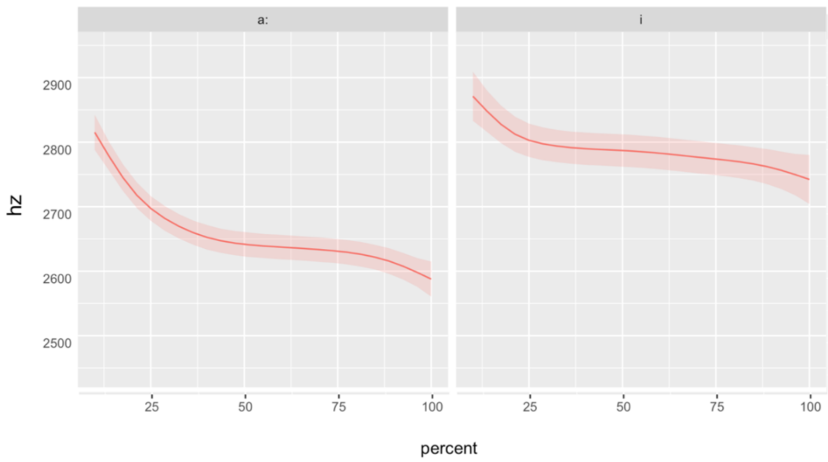


Figure 2 The Average values for F3 in both vowels a: and i:

10.10 Consent Form (with English translation)

Title

<Perception of emphasis by speakers of Arabic dialects>

Study Information

Researcher:

- Name: Mahmoud Abdulelah Alsabhi
- Email: mama507@york.ac.uk
- Mobile: +44 7474142141

What is the research about?

This study will look at the sounds of Arabic fricatives in four different Arabic dialects. It will examine the participant's responses to the target sounds to reach for a typology and see if the results map onto the ones from the production study.

What does the study involve?

The participant will be given a set of stimuli to listen to before being asked to click on the letter (s or S) that appears on the screen.

What will happen to the data I provide?

The data gathered from the participants will be saved on the primary researcher's laptop computer. The laptop will be password-protected. The data will then be analysed and used to support the researcher's thesis.

This study is run in compliance with the University's and will respect all your rights as described therein.

Participant Consent

By clicking the link below, you confirm that you have read and understood the information above, and that you agree to take part in the study.

عنوان الدراسة:

<قياس إدراك متحدثي اللهجات العربية لأصوات التفخيم في اللغة العربية>

معلومات الدراسة

معلومات الباحث:

الاسم: محمود بن عبد الإله السبحي

البريد الالكتروني: mama507@york.ac.uk

رقم التواصل: +44 7474142141

عن ماذا تتحدث الدراسة؟

ستبحث هذه الدراسة الادراكية في أصوات التفخيم العربية في أربع لهجات عربية مختلفة. ستقوم بتحليل ردود واختيارات المشاركين للأصوات المستهدفة للوصول إلى تصنيف معين ومعرفة ما إذا كانت نتائج الردود في هذه الدراسة متشابهة أو مختلفة ولأي مدى سيكون التشابه أو الاختلاف.

ما لذي سوف تتضمنه الدراسة؟

سيستمع المشترك إلى مجموعة من الأصوات، كل صوت في نافذة مختلفة حيث يتحين عليه الاستماع للصوت ومن ثم النقر على الخيار الأقرب لما سمع.

ماذا سوف يحدث للبيانات التي يتم جمعها؟

سيتم حفظ البيانات التي يتم جمعها من المشاركين على الكمبيوتر المحمول الخاص بالباحث الأساسي. سيكون الكمبيوتر المحمول محميًا بكلمة مرور حيث يقوم الباحث بتحليل البيانات واستخدامها لدعم أطروحة الباحث.

سيتم إجراء الدراسة وفقا لإشعار الخصوصية العامة لجامعة يورك في بريطانيا مما يتم فيه احترام وحفظ حقوق المشتركين كما هو موضح في الرابط أدناه:

موافقة المشاركة في الاستبيان

بالنقر على الرابط أدناه، فإنك تؤكد أنك قد قرأت وفهمت المعلومات الواردة أعلاه، وأنك توافق على المشاركة في الدراسة.

10.11 Questionnaire form (with English translation)

Personal Information

المعلومات الشخصية

Sex ما هو جنسك

Please Select...

Age الفئة العمرية

Please Select...

Place of residence بلد الإقامة

Please Select...

Dialect you speak? اللهجة العربية التي تتحدث

Please Select...

When you listen to someone speaking in one of the following dialects, how confident are you in identifying the speaker's dialect?

عند الاستماع لشخص يتحدث بأحد اللهجات التالية مامدى ثقتك بمعرفة لهجة المتحدث؟

(Write the percentage that represents your opinion between 0% and 100%) (اكتب النسبة التي تمثل رأيك مابين 0% إلى 100%)

غير واثق جداً = 0% Not very confident

واثق جداً = 100% Very confident

اللهجة المصرية Egyptian Dialect

اللهجة السورية Syrian Dialect

اللهجة الأردنية Jordanian Dialect

اللهجة المغربية Moroccan Dialect

اللهجة التونسية Tunisian Dialect

اللهجة الاماراتية Emirati Dialect

اللهجة الكويتية Kuwaiti Dialect

اللهجة العراقية Iraqi Dialect

What is the Arabic dialect that you like to listen to?

ماهي اللهجة العربية التي تحب الاستماع إليها؟

(Write the percentage that represents your opinion between 0% and 100%) (اكتب النسبة التي تمثل رأيك مابين 0% إلى 100%)

0% = لا احب الاستماع ابداً I never like to listen

100% = احب الاستماع دائماً I always like to listen

اللهجة المصرية Egyptian Dialect

اللهجة السورية Syrian Dialect

اللهجة الأردنية Jordanian Dialect

اللهجة المغربية Moroccan Dialect

اللهجة التونسية Tunisian Dialect

اللهجة الإماراتية Emirati Dialect

اللهجة الكويتية Kuwaiti Dialect

اللهجة العراقية Iraqi Dialect

What is the percentage of your listening and exposure to speakers of the following dialects?

ماهي نسبة استماعك وتعرضك لمتحدثي اللهجات التالية؟

(Write the percentage that represents your opinion between 0% and 100%) (اكتب النسبة التي تمثل رأيك مابين 0% إلى 100%)

0% = ابداً Never

100% = دائماً Always

اللهجة المصرية Egyptian Dialect

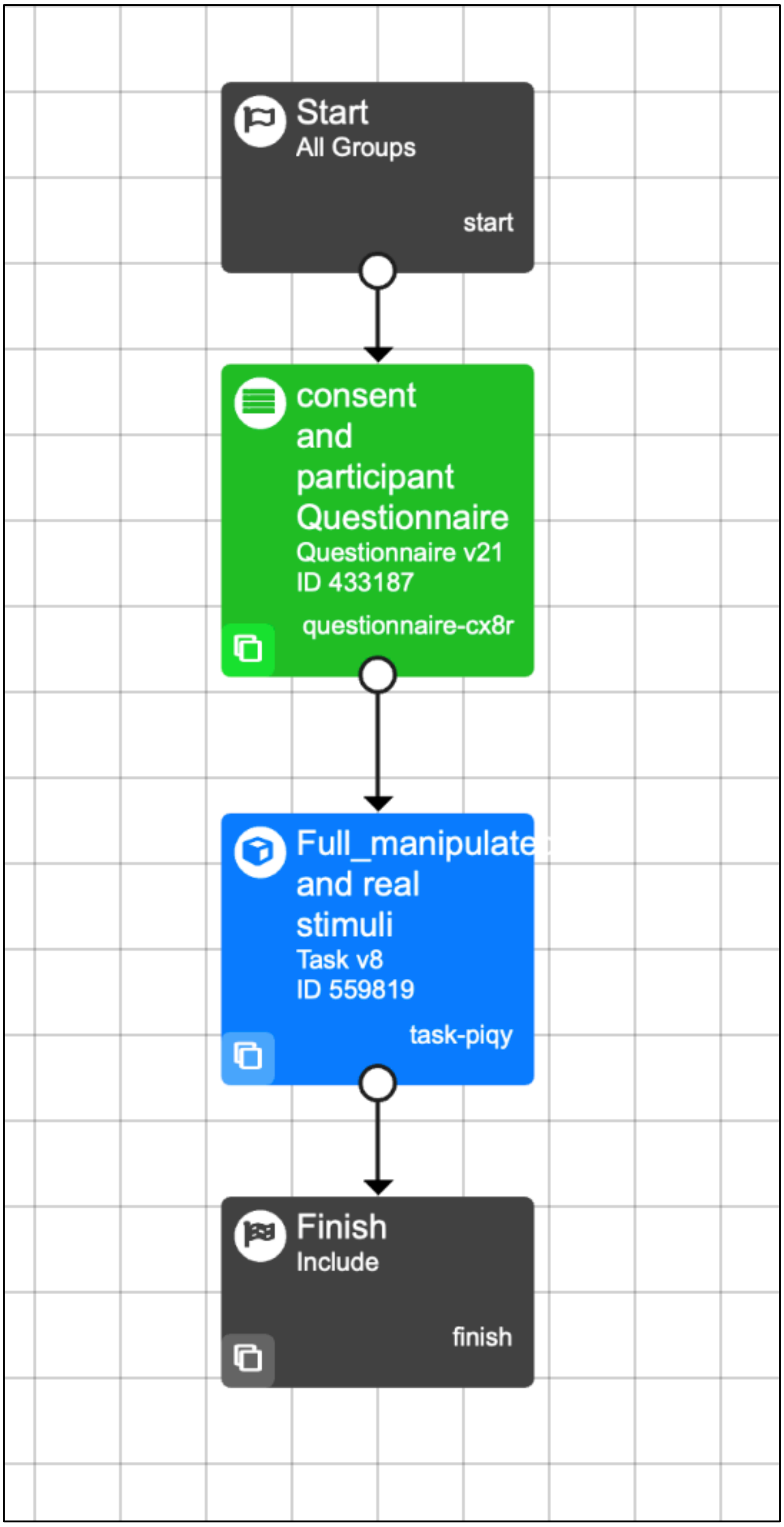
اللهجة السورية Syrian Dialect

اللهجة العراقية Iraqi Dialect

اللهجة المغربية Moroccan Dialect

Next

10.12 Graphical drag-and-drop interface



10.13 Mean and SD values for real stimuli, grouped by listener categories, vowel type, and trajectory.

Dialect Block	Dialect of the listeners	Vowel type	Trajectory	Mean_correct	Sd_correct
Iraqi	Iraqi	a	emphatic	0.9811	0.1374
Iraqi	Moroccan	a	emphatic	0.7333	0.4577
Iraqi	Syrian	a	emphatic	0.9231	0.2700
Moroccan	Iraqi	a	emphatic	0.7925	0.4094
Moroccan	Moroccan	a	emphatic	0.6667	0.4880
Moroccan	Syrian	a	emphatic	0.7179	0.4559
Syrian	Iraqi	a	emphatic	1.0000	0.0000
Syrian	Moroccan	a	emphatic	0.7333	0.4577
Syrian	Syrian	a	emphatic	0.8974	0.3074
Iraqi	Iraqi	i	emphatic	0.9057	0.2951
Iraqi	Moroccan	i	emphatic	0.7333	0.4577
Iraqi	Syrian	i	emphatic	0.8718	0.3387
Moroccan	Iraqi	i	emphatic	0.9811	0.1374
Moroccan	Moroccan	i	emphatic	0.8667	0.3519
Moroccan	Syrian	i	emphatic	0.8974	0.3074
Syrian	Iraqi	i	emphatic	0.9811	0.1374
Syrian	Moroccan	i	emphatic	0.9333	0.2582
Syrian	Syrian	i	emphatic	0.9231	0.2700
Iraqi	Iraqi	a	plain	0.7736	0.4225
Iraqi	Moroccan	a	plain	0.4000	0.5071
Iraqi	Syrian	a	plain	0.7949	0.4091
Moroccan	Iraqi	a	plain	0.2075	0.4094
Moroccan	Moroccan	a	plain	0.6667	0.4880
Moroccan	Syrian	a	plain	0.2308	0.4268
Syrian	Iraqi	a	plain	0.6038	0.4938
Syrian	Moroccan	a	plain	0.8667	0.3519
Syrian	Syrian	a	plain	0.8205	0.3888
Iraqi	Iraqi	i	plain	0.9434	0.2333
Iraqi	Moroccan	i	plain	0.9333	0.2582
Iraqi	Syrian	i	plain	0.8205	0.3888
Moroccan	Iraqi	i	plain	0.9623	0.1924
Moroccan	Moroccan	i	plain	0.9333	0.2582
Moroccan	Syrian	i	plain	0.9231	0.2700
Syrian	Iraqi	i	plain	0.9623	0.1924
Syrian	Moroccan	i	plain	1.0000	0.0000
Syrian	Syrian	i	plain	0.9487	0.2235

10.14 Mean and SD values for manipulated stimuli, grouped by listener categories, vowel type, and trajectory.

Dialect Block	Dialect of the listeners	Vowel type	Trajectory	Mean_correct	Sd_correct
Iraqi	Iraqi	a	emphatic	0.5566	0.4991
Iraqi	Moroccan	a	emphatic	0.3333	0.4795
Iraqi	Syrian	a	emphatic	0.1923	0.3967
Moroccan	Iraqi	a	emphatic	0.1981	0.4005
Moroccan	Moroccan	a	emphatic	0.6667	0.4795
Moroccan	Syrian	a	emphatic	0.1667	0.3751
Syrian	Iraqi	a	emphatic	0.3208	0.4690
Syrian	Moroccan	a	emphatic	0.2667	0.4498
Syrian	Syrian	a	emphatic	0.6667	0.4745
Iraqi	Iraqi	i	emphatic	0.5943	0.4934
Iraqi	Moroccan	i	emphatic	0.5000	0.5085
Iraqi	Syrian	i	emphatic	0.5513	0.5006
Moroccan	Iraqi	i	emphatic	0.1509	0.3597
Moroccan	Moroccan	i	emphatic	0.6333	0.4901
Moroccan	Syrian	i	emphatic	0.4231	0.4972
Syrian	Iraqi	i	emphatic	0.1792	0.3854
Syrian	Moroccan	i	emphatic	0.3000	0.4661
Syrian	Syrian	i	emphatic	0.5769	0.4972
Iraqi	Iraqi	a	plain	0.7736	0.4205
Iraqi	Moroccan	a	plain	0.6667	0.4795
Iraqi	Syrian	a	plain	0.7179	0.4529
Moroccan	Iraqi	a	plain	0.8302	0.3773
Moroccan	Moroccan	a	plain	0.7667	0.4302
Moroccan	Syrian	a	plain	0.7179	0.4529
Syrian	Iraqi	a	plain	0.8113	0.3931
Syrian	Moroccan	a	plain	0.7667	0.4302
Syrian	Syrian	a	plain	0.8846	0.3216
Iraqi	Iraqi	i	plain	0.9528	0.2130
Iraqi	Moroccan	i	plain	0.8667	0.3457
Iraqi	Syrian	i	plain	0.8718	0.3365
Moroccan	Iraqi	i	plain	0.9340	0.2495
Moroccan	Moroccan	i	plain	0.9667	0.1826
Moroccan	Syrian	i	plain	0.8846	0.3216
Syrian	Iraqi	i	plain	0.9057	0.2937
Syrian	Moroccan	i	plain	0.9000	0.3051
Syrian	Syrian	i	plain	0.8974	0.3054

11 Appendix C: (For Articulatory Study)

11.1 Linear regression results for the COG values for the emphatic contrast

subject	**Estimate**	**StdError**	**PValue**
eg01	-164.619	206.925	0.433
eg02	1260.858	253.94	3.68375e-05
hh01	5.817	207.819	0.978
hh02	737.777	208.854	0.002
hh03	-45.83	150.404	0.763
hh04	360.317	252.703	0.166
ir01	-480.636	398.219	0.245
ir02	395.504	160.927	0.022
ir04	247.193	238.651	0.31
mo02	547.305	281.91	0.065
sy01	208.128	150.189	0.178
sy02	266.193	105.067	0.018
sy03	255.058	226.4	0.271
sy04	92.373	523.475	0.861
sy05	88.744	299.707	0.77

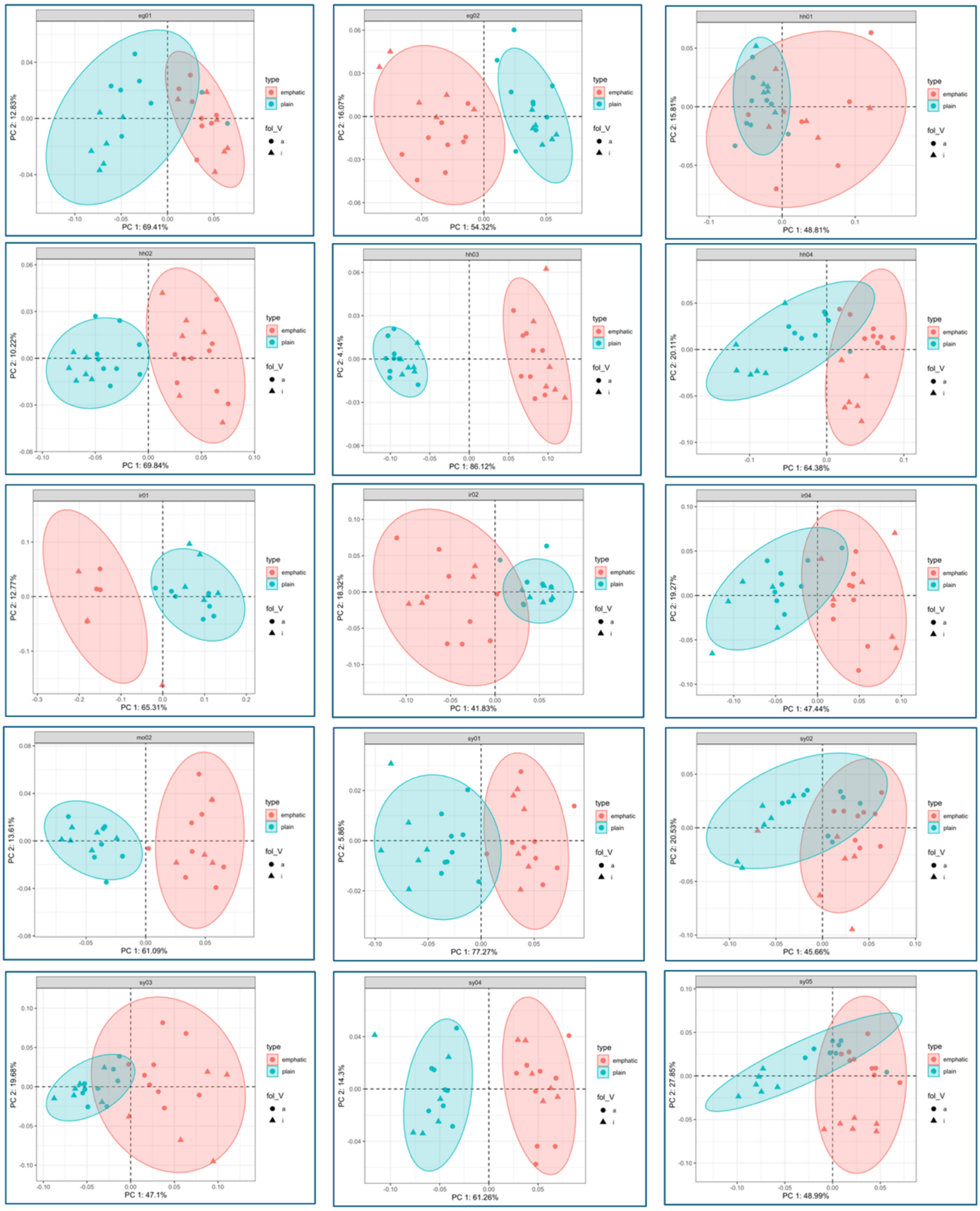
11.2 Linear regression results for the PC1

subject	**Estimate**	**StdError**	**PValue**
eg01	-0.05142	0.013657	> 0.001
eg02	0.067859	0.008676	< 0.001
hh01	-0.07184	0.020592	> 0.002
hh02	-0.08279	0.007801	< 0.001
hh03	-0.17165	0.00569	< 0.001
hh04	-0.07493	0.010822	> 0.079
ir01	0.227433	0.034121	< 0.001
ir02	0.090115	0.013632	< 0.001
ir04	-0.07164	0.013271	> 0.003
mo02	-0.08096	0.008292	< 0.001
sy01	-0.07344	0.007864	< 0.001
sy02	-0.04367	0.012752	> 0.002
sy03	-0.07015	0.01243	< 0.001
sy04	-0.10051	0.008606	< 0.001
sy05	-0.03135	0.0095	> 0.003

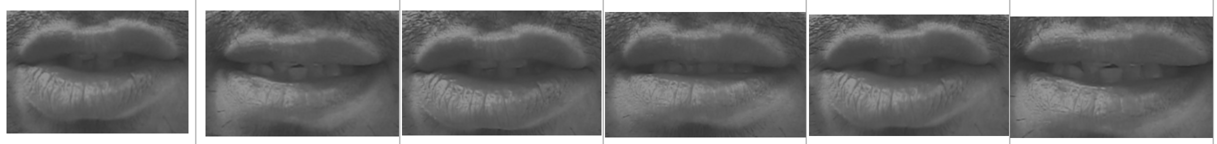
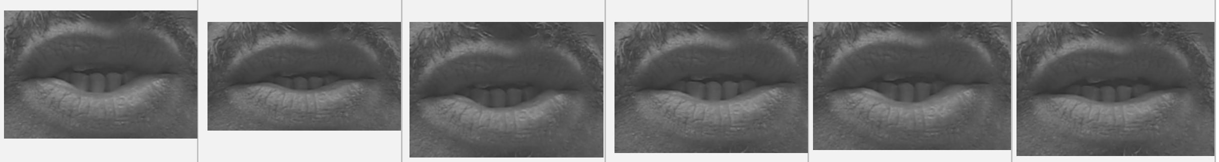
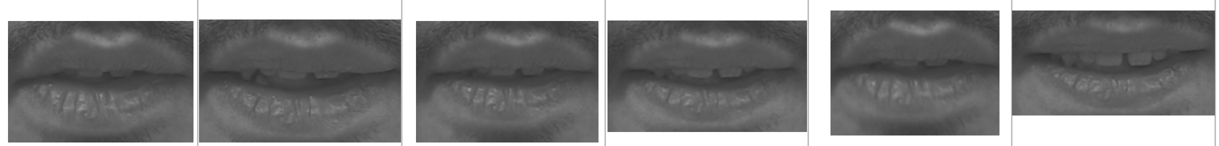
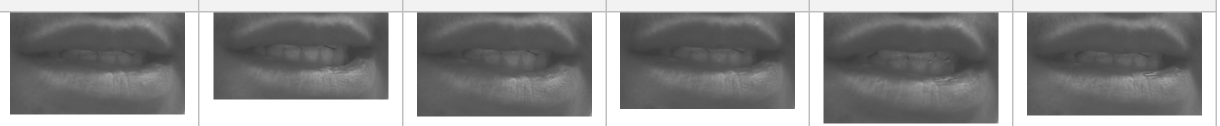
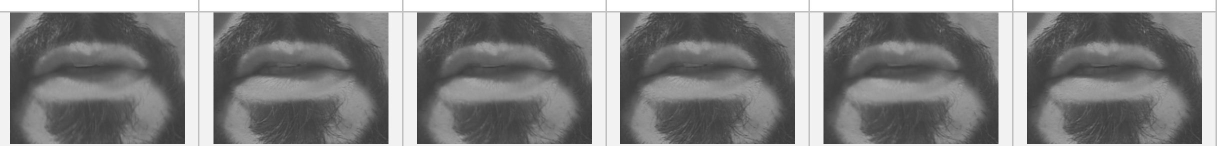
11.3 Linear regression results for the PC2

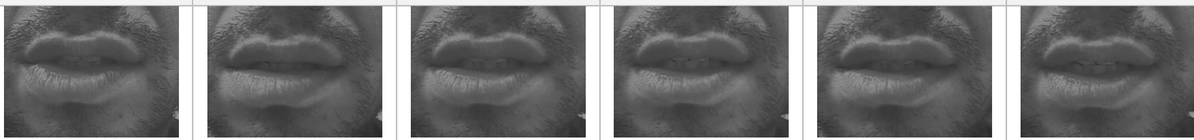
subject	Estimate	StdError	PValue
eg01	0.012646	0.009505	> 0.196
eg02	0.032601	0.008875	< 0.001
hh01	0.01026	0.014537	> 0.487
hh02	0.003627	0.009318	> 0.700
hh03	-0.000054	0.009751	> 0.996
hh04	0.001336	0.009701	> 0.892
ir01	-0.016999	0.030993	> 0.043
ir02	0.023494	0.018434	> 0.215
ir04	0.027194	0.017745	> 0.138
mo02	-0.005391	0.012001	> 0.658
sy01	-0.001433	0.0068	> 0.835
sy02	0.013582	0.01131	< 0.002
sy03	-0.019031	0.015006	> 0.216
sy04	0.008129	0.01385	> 0.563
sy05	0.011357	0.005847	< 0.001

11.4 PC1 and PC2 scores for tongue shapes for the emphatics and non-emphatics for all subjects



11.5 The associations of the lip's movement during the articulation of the emphatics and non-emphatics

Subject	Saar	saar	Saam	saam	Saab	saab
Eg01						
eg02						
hh01						
hh02						
Hh03						
Hh04						
Ir01						
Ir02						

Ir04	
Mo0 2	
Sy01	
Sy02	
Sy03	
Sy04	
Sy05