Université Paris Cité
Laboratoire de Linguistique Formelle (UMR 7110)

On the contribution of "Fine Phonetic Detail" (FPD) to Gradience in Phonology: Acoustic, Articulatory, Perceptual and Automatic Methods

Jalal Al-Tamimi PhD

Presentation of Research Activities to obtain the Habilitation à Diriger des Recherches (HDR)
in Phonetics and Phonology

Presented and defended publicly on 24 November 2023

Jury members:

President	Mme Martine Adda-Decker	Laboratoire de Phonétique et Phonologie, CNRS, France
Rapporteurs	Mr James Scobbie	Queen Margaret University, Edinburgh, UK
	Mr Rachid Ridouane	Laboratoire de Phonétique et Phonologie, CNRS, France
	Mme Ioana Chitoran (garante)	CLILLAC-Arp & Université Paris Cité, France
Examiners	Mme Ghada Khattab	Newcastle University, Newcastle upon Tyne, UK
	Mr John Esling	University of Victoria, British Colombia, Canada

Abstract

This HDR thesis summarises the various research activities I engaged in during my time at Lyon, Newcastle and Paris. The central idea of this thesis is that Fine-Phonetic-Detail (FPD) is an important aspect to consider for speech production, perception, articulation and in automated methods. FPD is defined as speaker-specific details, which are stored in the mental representation and used to identify specific categories and/or the speaker/dialect, following exemplar-based approaches to speech production and perception. Using FPD, I show its role in defining gradiency in Phonology.

My work spanned various areas. Starting from dynamic specification of vowels and the role of FPD (within the vowel and/or due to coarticulation) in identifying vowel categories, in production, perception and learning, in Arabic, French, English, and Mandarin. Then we move to the importance of the phonological feature [±Tense] in specifying geminate consonants in Lebanese Arabic, in both adult and child speech. We look at how adults systematically use this feature in their productions and how it influences the production of the first words in L1 acquisition. We look then at the Voicing contrast in Arabic in interaction with gemination, claiming that the phonological features [Long], [Tense] and [Voice] are all active, following a gradient privative features account. The specification of the Voicing contrast in other dialects shows variable behaviour for the Voiceless set, where it can either receive an [SG] or a [Tense] feature depending on the dialect. The next major research activity looks at the role of the epilarynx in the production of pharyngealisation and of gutturals, using acoustic, Ultrasound Tongue Imaging and ElectroGlottoGraphy. The combination of techniques allows to evaluate the articulatory to acoustic mapping to identify primary and secondary correlates of a combined supra-laryngeal and laryngeal gesture. Using extensions to this area, we look at dialectal and speaker identification in the context of gutturals and evaluating nasalisation in the production of pharyngeal consonants in Iraqi Arabic. We then travel to a new area of research I have engaged in on the use of automatic signal processing algorithms and machine learning to support diagnosis of cognitive disorders, followed by the development of the Arabic WebMAUS and WebMINNI services for forced-alignment. Other collaborative work is presented, such as, relationship between Adult-Directed-Speech (ADS) and Child-Directed-Speech (CDS), acoustics of the fifth liquid in Malayalam, impact of exposure on perceptual adaptations to various English accents, research on L2 phonology, the impact of native language on the preferred tongue postures, use of GAMMs (Generalised Additive Mixed-effects Models) on UTI data and comparison with SSANOVAs, and finally, our contribution to the Many Speech Analyses project, before finishing with the Open Source material that I put to the service of the research community.

We then move to a presentation of my significant activities as a PI, research management and co-supervision, and ending with my future research projects, where I aim to continue working in the same areas, but by expanding methodologies and performing cross-linguistic comparisons.

Résumé

Cette thèse d'HDR présente mes principales activités de recherches durant mon temps à Lyon, Newcastle et Paris. L'idée centrale de cette thèse est que le détail phonétique fin "Fine-Phonetic-Detail (FPD)" est un aspect important à considérer dans la production, perception, articulation et dans les méthodes automatisées. FPD est défini comme les détails spécifiques au locuteur qui sont stockés dans la représentation mentale et sont utilisés pour identifier des catégories spécifiques et/ou le locuteur/dialecte même, suivant les approches à exemplaires en production et perception de la parole. En utilisant le FPD, je démontre son rôle en définissant la Phonologie gradiente.

Ma recherche s'inscrit dans plusieurs domaines. En commençant par la spécification dynamique des voyelles et le rôle du FPD (à l'intérieure d'une voyelle et/ou dû à la coarticulation) en définissant les catégories vocaliques, en production, perception et en apprentissage, en arabe, français, anglais, et mandarin. Ensuite, on explore l'importance du trait phonologique [±Tense] dans la description des consonnes géminées en arabe libanais, chez l'adulte et chez l'enfant. On évalue comment les adultes utilisent ce trait systématiquement dans leurs productions et comment celles-ci influencent la production des premiers mots en acquisition de la L1. On explore ensuite le contraste du voisement en arabe en interaction avec la gémination en postulant que les traits phonologiques [Long], [Tense] et [Voice] sont tous les trois actifs, en suivant une approche privative gradiente. La spécification du contraste du voisement dans d'autres dialectes est variable pour les consonnes Non-Voisées, où elles reçoivent le trait [SG] ou [Tense] en fonction du dialecte. L'activité de recherche suivante met l'accent sur le rôle de l'épilarynx dans la production de la pharyngalisation et des consonnes gutturales, en utilisant des analyses acoustiques, de l'échographie de la langue et de l'ÉlectroGlottoGraphie. La combinaison des techniques permet d'évaluer le lien entre articulation et acoustique afin d'identifier les corrélats primaires et secondaire du geste combiné supra-laryngal et laryngal. Des extensions à ce domaine explorent l'identification dialectal et du locuteur dans le contexte des gutturales et évaluent le degré d'expansion de nasalité dans la production des consonnes pharyngales en arabe irakien. Ensuite, nous voyageons dans un nouveau domaine de recherche dans lequel je me suis engagé sur l'utilisation des méthodes automatisées du traitement du signal et de l'apprentissage automatique afin de soutenir le diagnostic des pathologies cognitives, suivi par le développement du système l'alignement forcé : Arabic WebMAUS et WebMINNI. D'autres activités en collaboration sont présentées, comme par exemple, la relation entre la Parole-Adressée-à-l'Adulte (ADS) et la Parole-Adressée-à-l'Enfant (CDS), l'analyse acoustique de la 5e consonne liquide en Malayalam, l'impact de l'exposition sur l'adaptation perceptive aux différences accentuelles de l'anglais, la recherche en phonologie de la L2, l'impact de la langue native sur les formes préférées des contours de la langue, l'utilisation des GAMMs (Régressions Généralisées Additives à Effets Mixtes) sur les données de l'échographie de la langue et comparaison avec les SSANOVAs, et finalement, notre contribution au projet "Many Speech Analyses", avant de terminer avec les ressources en sciences ouvertes que j'ai mis à disposition de la communauté de la recherche.

Ensuite, nous continuons avec une présentation de mes activités les plus marquantes en tant que PI, mes capacités à coordonner des activités de recherche et d'encadrement en terminant avec mes futurs projets de recherche, où je vais continuer à travailler dans les mêmes domaines en explorant de nouvelles méthodologies et en effectuant des comparaisons cross-linguistiques.

Acknowledgments

This HDR thesis summarises my research activities since 2001, when I first got introduced to the research community at the Laboratoire Dynamique du Langage (DDL) in Lyon during my masters and PhD. Since then, I engaged in various activities from both teaching, research and administrative duties, in Lyon, Newcastle and now in Paris. I am grateful to all colleagues and friends, whom I have known during these 22 years. Without their support, I would not have arrived to where I am today. The various exchanges I have had with you were important to shape my ideas and to allow me to expand on my knowledge. Thank you all! In the next few lines, I'll thank specific colleagues who had a major impact on my career as a person and as a researcher.

I start by thanking my committee members, without the feedback of whom this work will not have appeared. I am indebted to Jim Scobbie and Rachid Ridouane for critically evaluating my HDR, Ioana Chitoran for being my garante and critically evaluating this HDR, for Martine Adda-Decker for being the president of the committee and an examiner, for John Esling, for accepting to be part of the committee and to examine my HDR and finally, to my friend and colleague, Ghada Khattab for accepting to examine my work.

During my time in Lyon, I have had the chance to know many colleagues and I am deeply indebted to the following for various reasons. Lolke Van Der Veen for introducing me to the beauties of acoustic phonetics during my L3 studies at the Université Lyon 2 and supervising my Maîtrise (M1). Jean-Marie Hombert for guiding me during my DEA (M2) and PhD journeys. Special thanks go for René Carré for guiding me towards my PhD and for all the fruitful exchanges we have had on the role of dynamics in vowel specification. For François Pellegrino and Egido Marsico for the continued support during my PhD and after that. For the late John Ohala for his availability and guidance when I discussed my results with him. For my PhD committee whose discussions and insights shaped my future ideas. Special thanks go for Björn Lindblom for being one of the external examiners on my PhD, who pushed me to think outside of the box! And of course to all my colleagues and friends at DDL.

Then, after moving to Newcastle and staying there for slightly over 13 years and 9 months, my research and teaching careers have had a major turn. That is where I started challenging the accepted wisdom and explored the role of secondary correlates in specifying categories (for the features [±Tense] and [±CET]). I always got excited at discovering new tools, either from articulatory equipment, to latest techniques in signal processing and statistical analyses. 2017 was a major turn of event in how I approached data: it coincided with the time I first got to use the software R and got to know Bodo Winter, who initiated me to the beauties of Predictive Modelling and Random Forests! Many thanks indeed for this. That's when I also started using HPC (High Power Computing) and would like to thank the support teams at Newcastle University, UK and in France at the CNRS/TGIR HUMA-NUM, IN2P3 and at GENCI-IDRIS on using Jean-Zay. Many thanks to Jim Scobbie and Alan Wrench for introducing me to Ultrasound Tongue Imaging.

Of course, special thanks go to Ghada Khattab, with whom I worked (and continue!) almost 14 years in various teaching, research and supervision activities. You have introduced me to auditory analysis and ear training, which is an important step for any impressionistic transcriptions. You are a close friend and colleague and I cannot but thank you for believing in me! Special thanks for Gerry Docherty who introduced me to ear training and to sociophonetics. You guided me throughout my career. To Paul Foulkes for all the discussions and interests in sociophonetics, speaker-specific details, exemplar theory and forensics. Being immersed in sociophonetics from my first days there, I have had the chance to learn the importance of auditory analysis and ear training and to see how sociophonetics is employed, which I aim to expand on here in Paris. To Nicole Lallini (and Fabio Lallini) for being close friends and for the various discussions on employing Ultrasound Tongue Imaging to investigate speech initiation and planning in Apraxia of Speech. Special thanks of course go to all my previous colleagues at Newcastle University, your support and guidance helped me a lot during my time there.

During my time at Newcastle University, I have had the chance to meet various researchers who had a major impact on my research, including: Bronwen Evans, Tamar Keren-Portnoy, Chris Carignan, Danielle Turton, Barry Heselwood, Janet Watson, Alex Bellem, Bodo Winter, Timo Roettger, Stefano Coretta, Míša Hejná, Jim Scobbie, Alan Wrench, Pertti Palo, Jason Shaw, Cathy Best, Jen Hay, Márton Sóskuthy, Eleanor Chodroff, Scott Moisik, John Esling, Will Johns, Wolfram Hinzen, Matthias Heyne, Donald Derrick, Sam Hellmuth, Sam Kirkham, Niamh Kelly, Rory Turnbull, Alfie Herrero de Haro, Jane Stuart-Smith. (I am sure to have forgotten many!).

Since moving to the Université Paris Cité in October 2021, I have had the support from various colleagues: Ioana Chitoran, Hiyon Yoo, Giusy Turco, Barbara Hemforth, Anne Abeillé, Claire Saillard, Guillaume Wisniewski, Timothée Bernard, Olivier Bonami, Benoît Crabbé, Philippe Martin, Heather Burnett, Patrick Caudal, Ira Novek, Achille Falaise, Doriane Gras, Alexandre Roulois, Sophie Lempérière, Ali Dhibi, Maxime Hebert, Christel Préterre, Gabriel Thiberge..... Of course, during my time in Lyon, Newcastle and now Paris, I have had the chance to know and work directly with a few people. Special thanks go for: Rachid Ridouane, Martine Adda-Decker, Jacqueline Vassière, Cédric Gendrot, Anne Hermes, Nicolas Audibert, Claire Pillot-Loiseau, Cécile Fougeron, Maud Pélissier, Djidji Amazouz, Olivier Crouzet, and from elsewhere in the world: Wolfram Hinzen, Mary Lofgren, Rui He, Mohammad Abuoudeh, Josh Penny, Florian Schiel, Georgia Zellou, Martine Grice, etc.. (again I forgot many!!!).

I'd also like to thank my friend and colleague Emmanuel Ferragne. I have known you since 2003, first sharing our office while working towards our PhDs, then during my time at Newcastle and now as a colleague at the same university. What the odds are that we start our career together at the same lab and university and to be working for the same institution. What a small world! I'd like to express my profound gratitude to you Manu for all the support you have provided me with during these years, always being there responding to my queries, and for all your support.

The work summarised in the following pages and my subsequent research will not have seen the day without the support of the 2350 participants who participated in all my research activities (my own and collaborative) and for the additional 2000 participants I am planning to work on their productions and results from perception! Without your acceptance to be part of the various research activities, my work will not have seen the day. I'd also like to thank all the students with whom I have worked towards their PhDs (or Master degrees/internships): Wasan, Ying, Na, Dan, Yiling, Jasmine, Hana, Hajar, Saleh, Nief, Wael, Knight, Abdulrahman, Sarah, Bilal, Chenyu, Ziqi, Dan, Amina; also, Caihong, Zifeng, Fang, Dorotea, and all those who I have forgotten.

Lastly, I'd like to express my deepest thank-yous to my family: My late mum, my late dad and my late sister; my sister Meena, my brother Jamal, my sister in law Sooma, my niece Maisa, my nephew Osama, and little Ziad. My deepest thank-yous to my wife, Loana, who stood next to me during all of these years (easy and difficult) and to my lovely daughter, Léa. These last few years were full of challenges, but the future is full of surprises.. I love you baba and I love you Loana!

Part I

Curriculum Vitae, List of publications and PhD diploma

Chapter 1

Curriculum Vitae

Dr Jalal Al-Tamimi, BA, MA, PhD, UK-AFHEA

Linguistics department, Université Paris Cité, Laboratoire de Linguistique Formelle (LLF)
Associate Professor (MCF) of Experimental Psycholinguistics and Phonology
Homepage; GoogleScholar; Researchgate; GitHub; OSF; GlobalR; ORCID; HAL

Main research interests and research programme

My research falls within the Laboratory Phonology framework where I look at the role of fine-phonetic detail and its role in defining phonological categories and impact on production, perception and learning/acquisition. My current research interests are looking at the role of 1) secondary correlates in specification of pharyngealisation and gemination in Arabic; 2) dynamic specification of vowels in production, perception and in the learners' variety in Arabic, and Mandarin; 3) automated methods in prosodic and voice quality analyses in supporting diagnosis of schizophrenia and dementia; 4) Automated analyses of large-scale dataset and use of machine learning; 5) development of a variety independent romanization and forced-alignment system for Arabic.

Education/Qualifications

2020-2024	**Qualification for Professor level**, section 7 (Language Sciences), France (with HDR exemption)
2014-2015	**Certificate** of Newcastle Teaching Award. UKPSF: Descriptor 1, Newcastle University, UK
2012-2016	**Qualification for MCF (Associate Professor) level**, section 7 (Language Sciences), France
2011-2012	**Certificate** of "Teaching and Learning in Higher Education", Newcastle University, UK
2002-2007	**PhD** in Language Sciences (Phonetics and Phonology) University of Lyon 2, France. Committee: Jean-Marie Hombert and René Carré (Directors); Björn Lindblom and Salem Ghazali (external examiners; rapporteurs); Thami Benkirane and Willy Serniclaes (examiners)
2001-2002	**MA2 (DEA)** in Language Sciences (Phonetics and Phonology), University of Lyon 2, France
2000-2001	**MA1 (Maîtrise)** in Language Sciences (Phonetics and Phonology), University of Lyon 2, France
1999-2000	**BA** in Language Sciences, University of Lyon 2, France
1994-1998	**BA** in French Language and Literature, minor English, Yarmouk University, Irbid, Jordan

Appointments

2021-ongoing	**Associate Professor (MCF) of Experimental Psycholinguistics and Phonology**, Linguistics department, Université Paris Cité, France
2019-2021	**Senior Lecturer in Phonetics**, Speech and Language Sciences, Newcastle University, UK
2013-2019	**Lecturer in Phonetics**, Speech and Language Sciences, Newcastle University, UK
2012-2013	**Lecturer in Phonetics (50%), and Research Associate (50%)**, Australian Research Council (ARC) project, Newcastle University, UK
2011-2014	**Member** of the research network "NorPhLex" funded by the European Union Research Network NordForsk.
2010-2012	**Research Associate**, ESRC project, Newcastle University, UK
2007-2010	**Research Associate**, ESRC project, Newcastle University, UK
2002-2007	**Assistant Lecturer, contracted and ATER**, École d'Orthophonie (Speech and Language Sciences), Université Lyon 1, France
2002-2007	**Research Assistant**, Laboratoire Dynamique du Langage, Université Lyon 2, France

Memberships of professional bodies

2023-ongoing	Member of the "Statistics for Linguistics" Network
2023-ongoing	Member of the International DISCOURSE in Psychosis consortium
2022-ongoing	Member of the DementiaBank research group
2022-ongoing	Member of SIGMORPHON: "Special Interest Group on Computational Morphology and Phonology"
2022-ongoing	Member of European Language Resources Association (ELRA)
2021-ongoing	Member of the Labex (Laboratoire d'Excellence) "Empirical Foundations of Linguistics" Strand: 1: Complexity in Phonetics and Phonology and Strand 2: "Experimental Grammar in a cross-linguistic perspective"
2020-ongoing	Member of the International Phonetic Association (previously 2015-2016)
2020-ongoing	Member of the International Speech Communication Association, ISCA (previously 2002-2010)
2020-ongoing	Member of the Association Francophone de la Communication Parlée, AFCP (previously 2002-2010)
2017-ongoing	Member of the Association for Laboratory Phonology
2011-2014	Member of the International Congress for the Study of Child Language, IASCL
2008-ongoing	Member of the British Association of Academic Phoneticians, BAAP

Teaching activities

2021-ongoing	**Module leader and teacher** of 1) Experimental Methods and Psycholinguistics (BA 3^{rd} year); 2) French Phonetics and Phonology (BA 3^{rd} year; minor FLE); 3) Phonology (MA 1^{st} year); 4) Experimental Phonology (MA 1^{st} year); 5) Experimental design (MA 1^{st} year); 6) Technical Skills in Experimental and Computational Phonology (MA 2^{nd} year); Linguistics department, Université Paris Cité, France
2013-ongoing	**Member of PhD defence jury** - 10 PhD candidates: 6 PhD as external examiner (1 UK, 3 France, 1 India, 1 Pakistan); 4 internal examiner at Newcastle University, UK

Teaching activities (Cont.)

2008-ongoing **Member of PhD committees** - 11 PhD candidates at Newcastle University, UK; 1 PhD candidate at Université Paris Cité, France

2013-2021 **Module leader and teacher** of: Laboratory and experimental Phonology (IPhD in Phonetics and Phonology), Speech and Language Sciences, Newcastle University, UK

2011-2021 **Module leader and teacher** of: Articulatory, acoustic and clinical phonetics (BSc, MSpeech, MSc in Speech and language Pathology), Speech and Language Sciences, Newcastle University, UK

2008-2021 **Additional teaching and workshops** - experimental phonetics and phonology: e-learning material, experimental phonology, experimental phonetics: Praat scripting, normalisation, Voice Quality, quantitative methods and statistical techniques in phonetics and linguistics; Speech and Language Sciences, Newcastle University, UK

2005-2007 **Module leader and teacher** of: General linguistics – BSc, MA, École d'Orthophonie, Université Lyon 1, France

2002-2007 **Module leader and teacher** of: Experimental and clinical phonetics – BSc, MA, École d'Orthophonie, Université Lyon 1, France

International Collaborations and grants

Successful funding applications

2023-ongoing **Project title:** "GeSTeCEMA": Acquiring a new EMA (ElectroMagneticArticulatograph) machine
Team: Jalal Al-Tamimi (France), with Emmanuel Ferragne and Giusy Turco
Funding: AAP Petits & Moyens Équipements 2023 (€39.743k), LLF (€5k), CLILLAC-ARP (€13k) and Labex EFL Strand 1 (€7k); total secured €64.743k
Role: Lead on an internal funding application to acquire a new EMA (ElectroMagneticArticulatograph) machine

2023-ongoing **Project title:** "Système d'alignement forcé et de transcription automatisée de l'arabe dialectal"
Team: Jalal Al-Tamimi (France)
Funding: AAP Aide à la recherche BRIO « Budget Restreint, Impact Optimal » - ExcellencES funding within the "Humanities and Social Sciences Atrium", ATRIUM S&H, Université Paris Cité, €5k ; from the LLF, €1.8k; Settling-in funding, €2.5k
Role: Continue the development of the WebMAUS Forced-Alignment system and proposing an automated transcription system of dialectal Arabic

2022-ongoing **Project title:** "SDJAD: Speech Database of Jordanian Arabic Dialects"
Team: Mohammad Abuoudeh (Jordan), **Jalal Al-Tamimi (France)** and Olivier Crouzet (France)
Funding: Al Hussein Bin Talal University, Jordan, €10k
Role: Consultant on project design, recording, transcription and forced-alignment

International Collaborations and grants (Cont.)

Successful funding applications

2022-ongoing **Project title:** "Automatic methods in support of diagnosis of Alzheimer vs controls"
Team: Jalal Al-Tamimi (France)
Funding: Labex-EFL, €5.6k
Role: Project lead seeking funding and recruitment of an intern from the MA in Computation Linguistics

2022 **Project title:** Acquiring a new Ultrasound Tongue Imaging (UTI) and Licences for the software Articulate Assistant Advanced
Team: Jalal Al-Tamimi (France), LLF and CLILLAC-ARP
Funding: Settling-in funding (€6k), Laboratoire de Linguistique Formelle, €2.5k) and CLILLAC-ARP (€2.5k); €11k
Role: Project lead seeking funding from various sources

2021 **Project title:** Acquiring a new portable Electroglottography (EGG) and software VoceVistaVideo Pro
Team: Jalal Al-Tamimi (France)
Funding: Laboratoire de Linguistique Formelle €1.5k
Role: Project lead seeking funding

2019 **Project title:** "Role of the epilarynx in speaker and dialect classification of Levantine Arabic"
Team: Jalal Al-Tamimi (UK)
Funding: Faculty Bid Preparation Fund – HASS (Humanities and Social Sciences) Faculty, Newcastle University, UK; £5k
Role: Lead on project with networking and discussions with various research to prepare a large-scale ESRC funding application

2018-2019 **Project title:** "From articulation to speech recognition in investigating the Arabic sound system"
Team: Jalal Al-Tamimi (UK)
Funding: International Academic Fellowship, The Leverhulme Trust, UK; £28k
Role: Lead on project; recruitment of an assistant professor; visit to the Laboratoire de Phonétique et Phonologie (LPP), Université Paris 3, France.

2017-2021 **Project title:** "Forced alignment system for Arabic – Integration with MAUS system"
Team: Jalal Al-Tamimi (UK/France) and Florian Schiel (Germany)
Funding: Research account Newcastle University, UK, €5k
Role: Lead on project; development of an Arabic-variety independent romanization and forced-alignment system; data preparation, sharing with MAUS team

2017-2019 **Project title:** "Articulatory to acoustic mapping of the Epilarynx in the production of back consonants in Arabic"
Team: Jalal Al-Tamimi (UK)
Funding: British Academy/Leverhulme Small Research Grants; £10k
Role: Lead on project; Data acquisition (acoustic, Ultrasound Tongue Imaging, Electroglottography) of back consonants in Levantine Arabic

International Collaborations and grants (Cont.)

Successful funding applications

2015-2016 **Project title:** "Arabic Forced Alignment using PraatAlign and Maus automatic segmentation systems"
Team: Jalal Al-Tamimi (UK)
Funding: Faculty Research Fund, Newcastle University, UK; £4k
Role: Lead on project; adaptation of forced alignment systems to Arabic

2011-2021 **Project title:** Applictions to participate in the conferences: 17^{th} ICPhS, BAAP2014, 2^{nd} PaPE
Team: Jalal Al-Tamimi (UK)
Funding: School Research Committee (SRC), Newcastle University, UK, £6k; British Academy Overseas Conference Grant (OCG 2011), £650
Role: Lead on applications from various funding sources

Non-funded projects

2015-2016 **Project title:** "Investigating typical voice development in Tyneside children aged 4 – 11"
Team: PIs: **Jalal Al-Tamimi (UK)**, Nick Miller, Lindsay Pennington; students: Myah Ashkenazi, Sophie Taylor, Rebekah Drennan, Olga Mytili
Funding: Non-funded project
Role: Co-lead on project, ethical approval, data collection and processing; supervisor of a BSc dissertation

2018-ongoing **Project title:** "Automatic Prosody profile in Schizophrenia and dementia"
Team: Jalal Al-Tamimi (France), Will Jones (UK), Wolfram Hinzen (Spain)
Funding: Non-funded project
Role: Consultant on automated methods of prosody, voice quality and long-term measures; use of machine learning algorithms

Member of funded projects

2012-2015 **Project title:** "You came TO DIE?! Perceptual adaptation to regional accents as a new lens on the puzzle of spoken word recognition"
Team: Catherine Best (Australia), Jen Hay (New Zealand), Bronwen Evans (UK), Jason Shaw (USA), Paul Foulkes (UK) and Gerry Docherty (Australia)
Funding: Australian Research Council (ARC), Australia
Role: Post-doctoral research associate. Data recordings for English (Newcastle and York), Newcastle University, UK; Manager of finances between the two institutions, £100k

2011-2014 **Project title:** "NorPhLex: Phonological and lexical acquisition in mono- and bilingual children in the Nordic and the Baltic states"
Team: Kristian Kristoffersen (Norway), with other teams
Funding: NordForsk (European Union)
Role: Member of the UK team/research associate

International Collaborations and grants (Cont.)

Member of funded projects

2009-2013 **Project title:** "Psychological Significance of Production Templates in Phonological and Lexical Advance: A Cross-Linguistic Study"
Team: Marilyn Vihman (UK), Tamar Keren-Portnoy (UK), Rory DePaolis (USA), Ghada Khattab (UK) and Sophie Wauquier (France)
Funding: Economic and Social Research Council (ESRC), UK
Role: Member of the UK team/research associate

2007-2010 **Project title:** "Phonological acquisition in multilingual settings: the case of Lebanese Arabic"
Team: Ghada Khattab (UK)
Funding: Economic and Social Research Council (ESRC), UK
Role: Post-doctoral research associate

2005-2008 **Project title:** "Complexité, Langage et Langues – CL^2"
Team: François Pellegrino (France)
Funding: Agence Nationale de la Recherche, France
Role: pre-doctoral research associate

2001-2003 **Project title:** "Cognitique (P3) : Variabilité phonétique en production et perception de parole : rôle et limites des stratégies individuelles"
Team: René Carré (France)
Funding: Ministère de la Recherche et de l'Enseignement Supérieur, France
Role: pre-doctoral research associate

Unsuccessful funding applications

2020-2022 **Project title:** "Contrast reduction and loss in the sound systems of Arabic, Australian English, Japanese, Kalasha, Korean, and Spanish: production, perception and phonological theory"
Team: Alfredo Herrero de Haro (Australia), Ksenia Gnevsheva (Australia), Satoshi Nambu (Australia), Joshua Penney (Australia), Qandeel Hussain (Canada), **Jalal Al-Tamimi (France)**, and Yong-cheol Lee (Korea)
Funding: Discovery project, Australian Research Council; AUD $216k
Role: Lead on the Arabic side, on phonetics-phonology interface and statistical designs

2018-2019 **Project title:** "Second Language Acquisition: Diagnosis and remediation"
Team: Frédéric Isel, Emmanuel Ferragne, Cédric Gendrot; consultant: **Jalal Al-Tamimi (UK)**
Funding: Agence National de la Recherche (ANR); €428k
Role: Support on the Arabic phonetics and phonology front

2018-2019 **Project title:** "A cross-dialectal articulatory and instrumental investigation of the Arabic sound system"
Team: PI: Amel Issa; Mentor: **Jalal Al-Tamimi (UK)**
Funding: British Academy Postdoctoral Fellowship; £250k
Role: Mentor of a postdoctoral applicant; support on the Arabic phonetics and phonology and gutturals; advice on data collection and processing

International Collaborations and grants (Cont.)
Member of funded projects

2015-2016 **Project title:** "ARMPHLEX: Unravelling the intricate role of Arabic morphology, phonology and the lexicon in language development"
Team: Ghada Khattab (lead PI), **Jalal Al-Tamimi** (co-PI), Kai Alter and clinical teams
Funding: European Research Council (ERC), ERC-Consolidator Grant; €1509k
Role: Co-PI and co-lead on workpackage 1 "Phonetic properties of child-directed speech"; work on characteristics of IDS (Infant-Directed Speech), in comparison with ADS (Adult-Directed Speech) on a pre-existing corpus of over 300hour spontaneous mother-child interactions

---

Leadership
Administrative positions

2023-ongoing **Director of the Master in Language Sciences** - Unité de Formation et Recherches - Linguistique (UFR-L), Université Paris Cité (from September 2023)

2023-ongoing **Invited permanent member of the Board of Studies, Faculty of Humanities and Social Sciences** - Representing the Unité de Formation et Recherches - Linguistique (UFR-L), Université Paris Cité (from September 2023)

2023-ongoing **Deputy Representative at the Scientific Council** of the Unité de Formation et Recherches - Linguistique (UFR-L) representing the Laboratoire de Linguistique Formelle, Université Paris Cité (from January 2023)

2022-ongoing **Co-Director of the Strand 1 "Complexity in Phonetics and Phonology"** of the Labex (Laboratoire d'Excellence) "Empirical Foundations in Linguistics"

2022-ongoing **Co-Director of the Experimental Linguistics Research Strand**, Laboratoire de Linguistique Formelle; Leading on writing up of the Lab's 2025-2029 project assessed by the HCÉRES

2020-2021 **External examiner** "MSc in Forensic Speech Science" and "MA in Phonetics and Phonology", Language and Linguistics department, York University, UK

2019-2021 **Degree Programme Director** for the IPhD in Phonetics and Phonology, Newcastle University, UK

2017-2021 **Admissions tutor** for BSc Speech and Language Therapy and Integrated Master in Speech and Language Sciences, ECLS, Newcastle University, UK

2014-2021 **Chair and member** of the Speech and Language Sciences Ethics Committee, Newcastle University, UK

Service to the community

2022-ongoing **Member of the Board of Examiners** for the Theoretical and Experimental Linguistics (L3, M1, and M2) - Unité de Formation et Recherches - Linguistique (UFR-L), Université Paris Cité

2022-ongoing **Representative** of the Phonetics and Phonology (Phi&Phi) Master's program; Paris Graduate School of Linguistics (PGSL), Université Paris Cité, France

2022-ongoing **Founding Member** of the LPRG: "Laboratory Phonology Research Group", Université Paris Cité, France

Leadership (Cont.)

Service to the community

2021-ongoing	**Member** of Paris Graduate School of Linguistics (PGSL) Research Ethics Committee, Linguistics Department, Université Paris Cité, France
2021-ongoing	**Member** of the pedagogic team of the Master in Phonetics and Phonology, Paris Graduate School of Linguistics (PGSL), Université Paris Cité, France
2020-ongoing	**Co-Founding member** of GlobalR meetup group; Support from the R foundation; regular meetings and online training events
2018-2020	**Founding member** of Adventures in R meetup group; Support from the R foundation; regular meetings and training events, Newcastle University, UK
2018-2021	**Member** of the HASS Faculty Library committee, Newcastle University, UK
2018-2021	**Independent Chair** for PhD examinations across the three faculties, , Newcastle University, UK
2018-2019	**Mentor** for a Teaching Fellow in Phonetics, Newcastle University, UK
2017-2021	**Member** of ECLS ethics committee, representative of SLS Section, Newcastle University, UK
2017-2021	**Member** of the "Athena SWAN Self-Assessment Team" at ECLS coordinating submission for a Bronze Athena SWAN award, Newcastle University, UK
2016-2020	**Mentor** for a first time PhD supervisor, Newcastle University, UK
2014-2021	**Leading role** in redesign of modules in the IPhD in Phonetics and Phonology; candidate selection and ways to improve programme, Newcastle University, UK
2014-2021	**Member of interviewing board** 3 teaching assistants in phonetics and phonology; one Senior Lecturer in Quantitative methods, Newcastle University, UK
2014-2017	**Coordinator** for the peer mentoring scheme, Newcastle University, UK
2014-2017	**Member** of cross faculty steering group for the peer mentoring scheme
2014-2017	**Coordinator** of the peer dialogue scheme, Newcastle University, UK
2014-2017	**Coordinator** of the NSS survey. Meeting with students and advertising dates, Newcastle University, UK
2008-2013	**RA representative** at the CRiLLS (Centre for Research in Language and Linguistics, Newcastle University) executive committee, Newcastle University, UK

BSc, MA, Internship and PhD Supervision

Year	Name	Split (order)	Topic/Programme
BSc level (4)			
2018-2019	Jess Cooper	80% (first) With Jones	"The prosodic features of controlled speech of patients with a diagnosis of Schizophrenia"; BSc (Hons) Speech and Language Sciences, Newcastle University, UK

BSc, MA, Internship and PhD Supervision

Year	Name	Split (order)	Topic/Programme
BSc level (4)			
2018-2019	Jennifer Booth	80% (first) With Jones	"A comparative acoustic analysis of prosody in spontaneous speech between individuals with Schizophrenia and Neuro-Typically healthy controls"; BSc (Hons) Speech and Language Sciences, Newcastle University, UK
2015-2016	Myah Ashkenazi	100%	"Investigating norms in vowel formant frequency variation within typical voice development: A comparison of sustained and single word production in Typically developing boys and girls in the North East of England, aged 4-11 years"; BSc (Hons) Speech and Language Sciences, Newcastle University, UK
2007-2008	Anna Lalot & Sylvain Reyt	100%	"Étude comparative de la production des voyelles entre des locuteurs sains et pathologiques : effet du contexte et de la composante de souffle"; Mémoire de fin d'étude en Orthophonie, Université Lyon 1, France
MA Internship (3)			
2022-2023	Eric Jordan	100%	"Guiding Alzheimer's diagnosis using Speech processing techniques"; 6-months Internship funded by the Labex-EFL, MA1 Computational Linguistics and Automatic Speech Recognition; Laboratoire de Linguistique Formelle, Université Paris Cité, France
2021-2022	Eric Jordan	100%	"On the use of OpenSMILE to guide diagnosis in Alzheimer disease"; 3-months Internship funded by the Labex-EFL, MA1 Computational Linguistics and Automatic Speech Recognition; Laboratoire de Linguistique Formelle, Université Paris Cité, France
2021-2022	Amina Djafri	100%	"Arabic Forced-alignment system: continue developments of system"; 4-months Internship funded by the LLF, MA1 INALCO, Computational Linguistics and Automatic Speech Recognition; Laboratoire de Linguistique Formelle, Université Paris Cité, France
MA level (6)			
2022-2023	Ziqi Zhou	80% (first) With Yoo	"What makes an utterance ironic? A predictive modelling approach to perceptual judgments of L2 learners of French"; Masters in Phonetics and Phonology, Linguistics department, Université Paris Cité, France, funded by the Labex-EFL

BSc, MA, Internship and PhD Supervision (Cont.)

Year	Name	Split (order)	Topic/Programme
MA level (6)			
2021-2022	Chenyu Li	80% (first) With Wu	"Le rôle des facteurs influençant la réalisation des diphtongues en mandarin standard : la modélisation dynamique avec les GAMMs"; Masters in Phonetics and Phonology, Linguistics department, Université Paris Cité, France, funded by the Labex-EFL
2017-2018	Mahmoud Alsabhi	50% (second) With Khattab	"The production of Emphatic Consonants in Hijazi and Eastern Dialects of Arabic"; Summer project for the IPhD in Phonetics and Phonology, Newcastle University, UK, funded by the Saudi Government
2017-2018	Abdulmajeed Al-Jehani	50% (second) With Khattab	"The Perception of Initial English Consonant Clusters by Urban and Bedouin Hijazi Speakers"; Summer project for the IPhD in Phonetics and Phonology, Newcastle University, UK, funded by the Saudi Government
2016-2017	Nief Al-Gamdi	50% (second) With Khattab	"Voicing contrast in Najdi Arabic initial stops"; Summer project for the IPhD in Phonetics and Phonology, Newcastle University, UK, funded by the Saudi Government
2016-2017	Wael Almurashi	50% (second) With Khattab	"Static and Dynamic Cues in Vowel Production in Hijazi Arabic"; Summer project for the IPhD in Phonetics and Phonology, Newcastle University, UK, funded by the Saudi Government
PhD level *Current (4)*			
2023-2027	Ziqi Zhou	50% (second) With Chitoran	"Cross-Linguistic Influence on Verbal Irony in L3 French: An Acoustic and Perceptual Analysis of Mandarin-English Bilinguals"; PhD in Phonetics & Phonology, Linguistics department, Université Paris Cité, France, funded by the Chinese Scholarship Council (CSC)
2022-2026	Chenyu Li	100%	"The Role of Factors Influencing the Realization of Diphthongs in Mandarin: A Sociophonetic and Phonological Perspective"; PhD in Phonetics & Phonology, Linguistics department, Université Paris Cité, France, funded by the Chinese Scholarship Council (CSC)

BSc, MA, Internship and PhD Supervision (Cont.)

Year	Name	Split (order)	Topic/Programme
PhD level			
Current (4)			
2019-2023	Bilal Alsharif	30% (second) With Khattab & Turnbull	"The Articulatory Dynamics of Postvelar Production and Coarticulation in Levantine Arabic"; IPhD Linguistics, SELLL, Newcastle University, UK, self-funded, submitted
2019-2023	Sarah Alghabban	50% (second) With Khattab	"Improving the Perception and Production of English Vowels among Arabic EFL learners in Saudi Arabia"; PhD in Speech Sciences, ECLS, Newcastle University, UK, funded by the Saudi Government
Completed (13)			
2019-2023	Abdulrahman Dallak	50% (second) With Khattab	"Arabic Obstruents: Laryngeal Contrast and Representation"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, funded by the Saudi Government
2018-2022	Knight Kamphikul	50% (second) With Young-Scholten	"The Intonation Patterns of Thai Non-Native Speakers of English: Forms and Functions towards Communication"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, funded by the Thai Government
2017-2022	Wael Almurashi	50% (first) With Khattab	"Dynamic specification of vowels in Hijazi Arabic"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, funded by the Saudi Government
2017-2022	Nief Al-Gamdi	50% (first) With Khattab	"Acoustic properties of the voicing contrast in Najdi Arabic"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, funded by the Saudi Government
2018-2021	Saleh Ghadanfari	30% (second) With White & Khattab	"Syllable weight and perceptual energy: a case study of Kuwaiti Arabic"; PhD in Speech Sciences, ECLS, Newcastle University, UK, funded by the Kuwaiti Government
2016-2020	Hajar Moussa	90% (first) With White	"The phonetics and phonology of Jeddah Arabic prosody"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK
2016-2020	Hana Ehbara	40% (second) With Young-Scholten	"The Impact of Computer-Assisted Pronunciation Teaching on Libyan Child Learners of English"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, funded by the Libyan Government

BSc, MA, Internship and PhD Supervision (Cont.)

Year	Name	Split (order)	Topic/Programme
PhD level			
Completed (13)			
2018-2020	Jasmine Warburton	30% (second) With Turton	"Vowel Mergers in Tyneside English: how does sound change affect the relationship between production and perception?"; PhD Linguistics, SELLL, NINEDTP, Newcastle University, UK, funded by the NINEDTP-ESRC
2014-2018	Yiling Chen	50% (second) With Khattab	"The effect of explicit instruction and auditory/audio-visual training on Chinese EFL learners' acquisition of intonation"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, self-funded
2014-2018	Daniel Mccarthy	50% (first) With Khattab	"The acoustics of place of articulation in English plosives"; PhD in Speech Sciences, ECLS, NINEDTP, Newcastle University, UK, funded by the NINEDTP-ESRC
2013-2016	Patchanok Kitikanan	50% (first) With Khattab	"L2 English Fricative Production by Thai Learners"; IPhD Phonetics & Phonology, ECLS, Newcastle University, UK, funded by the Thai government
2012-2015	Ying Li	40% (second) With Hannahs	"Audio-visual training effect on L2 perception and production of English /θ/-/s/ and /ð/-/z/ by Mandarin speakers"; PhD Linguistics, SELLL, Newcastle University, UK, self-funded
2009-2013	Wasan AlSiraih	50% (second) With Khattab	"Voice Quality Features in the Production of Pharyngeal Consonants by Iraqi Arabic Speakers"; PhD in Speech Sciences, ECLS, Newcastle University, UK, funded by the Iraqi government

PhD Examination (10)

Year	Name	Title	Examination; University
2022	Natalja Ulrich	Variation, Change and Complexity in Linguistic and Health-Related Bahaviour	External examiner (pre-viva report, examination, post-viva report); PhD in Language Sciences, option Phonetics and Phonology, Université Lumière Lyon 2, Lyon, France
2020	Azza Al-Kendi	Foreigner-Directed-Speech and L2 Speech Learning Model in An Understudied Interactional Setting: The Case of Foreign-Domestic Helpers in Oman	Internal examiner (pre-viva report, examination, post-viva report); IPhD in Linguistics and English Language, SELLL, Newcastle University, UK

PhD Examination (Cont.)

Year	Name	Title	Examination; University
2019	Maha JASIM	Tafxi:m in the Qəltu and Gilit dialects of Iraqi Arabic	Internal examiner(pre-viva report, examination, post-viva report); IPhD in Phonetics and Phonology, ECLS, Newcastle University, UK
2018	Ammar ZEINEB	Perception et production des voyelles orales françaises par des enfants tunisiens	External examiner (examination, post-viva report); PhD in Phonetics and Phonology, Université Sorbonne Nouvelle - Paris 3, Paris, France
2018	Mohammad ABUOUDEH	De l'impact des variations temporelles sur les transitions formantiques	External examiner (examination, post-viva report); PhD in Phonetics and Phonology, Université de Nantes, Nantes, France
2018	Aisha MOHDAR	Vowel-to-Vowel coarticulation in Yemeni Arabic Dialects: Impact of vowel space and pharyngealized consonants	External examiner (pre-viva report, examination, post-viva report); PhD in Phonetics and Phonology, University of Hyderabad, India
2016	Amin ALSHANGITI	Lexical borrowings in immigrant speech: A sociolinguistic study of Hassaniyya Arabic speakers in Medina (Saudi Arabia	External examiner (pre-viva report, examination, post-viva report); PhD in Phonetics and Phonology, Durham University, UK
2015	Hisham ALKADI	English speakers' common orthographic errors in Arabic as L2 writing system: An analytical case study	Internal examiner (pre-viva report, examination, post-viva report); PhD in Applied Linguistics, ECLS, Newcastle University, UK
2013	Isao HARA	An Acoustic Analysis of Vowel Sequences in Japanese	Internal examiner (pre-viva report, examination, post-viva report); PhD in Speech Sciences, ECLS, Newcastle University, UK
2012	Muhammad KHAN	Pashto Phonology: The relationship between syllable structure and word order	External examiner (pre-viva report, examination, post-viva report); PhD in Phonetics and Phonology, The University of Azad Jammu & Kashmir, Muzaffarabad, Pakistan

Teaching esteem indicator and recognition

2021-ongoing	**Member of various committees** within the UFR-L, the Laboratoire de Linguistique Formelle PGSL, and the Labex EFL, Université Paris Cité, France
2020-2021	**Appointed as an External examiner** for two Master programmes at the Language and Linguistics department, University of York, UK
2019	**Independent evaluator** of an IPhD programme in Linguistics at the University of York, UK
2017-2021	**Praise for proposing evidence-based assessments** in Speech and Language Sciences (by MSc external examiner), Speech and Language sciences, Newcastle University, UK
2017-2021	**Praise for the state-of-the-art teaching in experimental phonetics and phonology** judged as one of the best in the UK (by IPhD in Phonetics and Phonology external examiner), Speech and Language sciences, Newcastle University, UK
2014-2021	**Member** of team working on revisions of the IPhD in Phonetics and Phonology, Newcastle University, UK
2014-2021	**Recruitment** of five ESRC funded PhD candidates, Speech and Language sciences, Newcastle University, UK
2010-2011	**Member** of team developing the new IPhD in Phonetics and Phonology, Newcastle University, UK

Research esteem indicator and recognition

2023	**Invited** to stand for election - Member of the Scientific Council - International Phonetic Association
2022	**Invited** to stand for election - Deputy representative at the Scientific Council of the Unité de Formation et Recherches Linguistique (UFR-L), representing the Laboratoire de Linguistique Formelle
2022	**Invited** to stand for election - Co-Director of the research strand: Experimental Linguistics, Laboratoire de Linguistique Formelle
2022	**Invited** to stand for election - Councillor at large - Association for Laboratory Phonology
2019	**Nomination** – "Research Director of the year" by students as part of the Newcastle University "Teaching Excellence Awards, Newcastle University, UK
2018	**Received** the Peter Ladefoged prize for the presentation most in the spirit of the work of the late Peter Ladefoged, British Association for Academic Phoneticians.
2018	**Nomination** – "Research Director of the year" by students as part of the Newcastle University "Teaching Excellence Awards, Newcastle University, UK
2012-2013	**Manager** of an Australian Research Council (ARC) Discovery project. Management of finances and recordings at Newcastle and York Universities
2008-ongoing	**Grant reviewer** for ESRC (4), QNRF (Qatar National Research Fund)
2008-ongoing	**Peer reviewer** for 28 International journals and 9 international conferences (full list below)

Research esteem indicator and recognition (Cont.)

Invited speaker

2023 **Research Seminar** - LingLunch, Laboratoire de Linguistique Formelle, Université Paris Cité, France

2022 **Research Seminar** - Phonetics and Phonology (SRPP), Laboratoire de Phonétique et Phonologie (LPP), Université Paris 3, France

2022 **Research seminar** - Symposium Labex EFL (Empirical Foundations of Linguistics), France

2021 **Invited Workshop lead** - Introduction to Random Forests - Phonetics and Phonology Research Group at Aarhus University, Denmark

2020 **Research seminar** - Speech Science Forum, Hearing and Phonetic Sciences – University College London, UK

2020 **Invited Workshop co-lead** LingLab LingLab Methods Fair, delivered to Researchers and Doctoral Students in Linguistics and Phonetic Sciences, Newcastle University, UK

2019 **Research Seminar** - Traitement Du Langage Parlé (TLP), Laboratoire d'Informatique Pour La Mécanique et Les Sciences de l'Ingénieur (LIMSI) - Université Paris-Sud, France

2019 **Research Seminar** - Phonetics and Phonology (SRPP), Laboratoire de Phonétique et Phonologie (LPP), Université Paris 3, France

2018 **Research Seminar** - Presentation of current work at the Joint Adult and Child Language Research Groups at the Speech and Language Sciences section, Newcastle University, UK

2018 **Workshop** - Temporalité et séquentialité dans les formes sonores, Nantes, France.

2016 **Invited speaker Conference** - Issues in experimental designs in Arabic research, Arabic Linguistics Forum, York University, UK

2012 **Research Seminar** of the Language and Linguistic department, University of York, UK

2010 **Research Seminar** of the CRiLLS-Heidelberg meeting, Newcastle University, UK

2010 **Research Seminar** of Laboratoire Dynamique du Langage, University of Lyon 2, France

Organiser of the following research events

2026 **Conference**: Member of the organising committee of the 12th Ultrafest Conference, Université Paris Cité, France, 2026

2024 **Conference**: Member of the organising committee of the 13th International Seminar on Speech Production (th ISSP), Autrans, France, 13-17 May 2024

2023 **Outreach/Teaching**: Member of the LingFest, summe school in Linguistics, leading on the session: Introduction to Phonetics, PGSL, Université Paris Cité, Paris, France, 4-8September 2023

2023 **Outreach**: Member of the organising committee of the "Discovering the IPA" within the LinguaFest, Paris, France, 29 September-1 October 2023

2023 **Workshop**: Coordinator of an Eye-Tracking workshop, Labex-EFL, 2023

2018-2021 **Four Two-day workshop** on advanced statistics in HASS faculty training for PGR (June 2018, 2019, 2020 and 2021)

2016 **Three-day workshop** on Regression and Mixed Effects Modelling, Newcastle University, 25-27 January 2016

2010 **CRiLLS/IoN Research Away Day**, Newcastle University, 25th June 2010

Research esteem indicator and recognition (Cont.)

Organiser of the following research events

2009	**Workshop** on Pharyngeals and Pharyngealisation, Newcastle University, 26-27 March 2009
2007	**Workshop** "Coarticulation: Cues, Direction & Representation", University of Montpellier 3, France, 7th December 2007

Scientific dissemination

2020	**Invitation** to review a book proposal for Wiley Blackwell: "Phonetics Experiments: A Guide for Speech Scientists", by Henning Reetz and Allard Jongman
2014	**Invitation** to review a book proposal for Palgrave: "Acoustic Phonetics: A practical introduction" By Frank Herrmann
2013	**Invitation** to review a book "Articulatory Phonetics" (see J. Al-Tamimi, 2013)
2011	**Invitation** to translate an article from English to French (see J. Al-Tamimi, 2011)

Member of the peer review committee for

International Journals: 1) The Journal of the Acoustical Society of America (JASA), 2) Journal of the International Phonetic Association (JIPA), 3) Journal of Phonetics, 4) Phonetica, 5) Laboratory Phonology, 6) Speech, Language and Hearing, 7) Applied Psycholinguistics, 8) Journal of Sociolinguistic Studies, 9) International Journal of Arabic-English Studies, 10) Folia Linguistica, 11) Acta Linguistica, 12) Studia Linguistica, 13) Revue Canadienne de Linguistique, 14) Open Computer science, 15) Linguistics Vangard, 16) Language and Speech, 17) Brill Journal Afroasiatic Languages and Linguistics, 18) Ampersand, 19) Second Language Research, 20) Sociolinguistic Studies, 21) Journal of Speech Language and Hearing Research, 22) PLoS ONE, 23) Speech Communication, 24) Computer Speech & Language, 25) Radical: A Journal of Phonology, 26) SKASE, 27) Perspectives on Arabic Linguistics XXXIV, 28) Aphasiology, 29) Espaces linguistiques;

Conferences : 1) LabPhon Conference, 2) Interspeech, 3) International Congress in Phonetic Sciences (ICPhS), 4) British Association of Academic Phoneticians (BAAP) Colloquium, 5) Newcastle working papers in Linguistics, 6) York working papers in Linguistics, 7) workshop "pharyngeal and pharyngealization", 8) Workshop coarticulation et 9) Phonetics and Phonology in Europe (PaPE)

Chapter 2

Publication List

All my publications under 2.1, 2.2 and 2.3 are accessible as open access publications either at Researchgate or at HAL. For 2.8, all my open source notebooks are accessible from my GitHub

2.1 Articles (27): "in preparation" (9; refs. 1-9), "under revisions" (1; ref. 10), "under review" (3; refs. 11-13), and "Published" in Peer Reviewed Journals (16; refs. 14-26)

1. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (In preparation). Voicing Contrast and Voicing Assimilation in Najdi Arabic. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.
2. **Al-Tamimi, J.** (In preparation). Acoustic properties of the laryngeal and supra-laryngeal effects of epilaryngeal constrictions in Levantine Arabic guttural consonants. *The Journal of the Acoustical Society of America*.
3. **Al-Tamimi, J.**, Jones, W., & Hinzen, W. (In preparation). Automatic acoustic profiling of prosody in schizophrenia. *The Journal of the Acoustical Society of America, Express Letters*.
4. **Al-Tamimi, J.**, Lofgren, M., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (In preparation). Automated prosodic and voice quality profiles to distinguish between groups at elevated risk of Alzheimer's disease. *Speech Communication*.
5. **Al-Tamimi, J.**, & Palo, P. (In preparation). Gradient epilaryngeal constriction of back consonants in Levantine Arabic: A generalised additive mixed modelling of Ultrasound Tongue Imaging. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.
6. Almurashi, W., Khattab, G., & **Al-Tamimi, J.** (In preparation). Acoustic Analysis of English Vowel Production by Hijazi Arabic L2 Learners. *The Journal of the Acoustical Society of America*.
7. Li, C., & **Al-Tamimi, J.** (In preparation). The Impact of the Tonal Factor on Diphthong Realizations in Standard Mandarin: Explanations within the Articulatory Phonology Framework modelled via Generalized Additive Mixed Models. *Laboratory Phonology:*

Journal of the Association for Laboratory Phonology.

8. Derrick, D., **Al-Tamimi, J.**, & Heyne, M. (In preparation). Accounting for within and between-subject variation using generalized additive mixed models on ultrasound tongue contours. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.
9. Moussa, H., & **Al-Tamimi, J.** (In preparation). Marking Focus in Jeddah Arabic. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.
10. Almurashi, W., **Al-Tamimi, J.**, & Khattab, G. (Under Revisions). Dynamic specification of vowels in Hijazi Arabic. *Phonetica*.
11. Almurashi, W., **Al-Tamimi, J.**, & Khattab, G. (Under Review). Acoustic Cues in Production of Southern British English. *Speech Communication*.
12. He, R., **Al-Tamimi, J.**, Sánchez-Benavides, G., Montaña-Valverde, G., Gispert, J. D., Grau-Rivera, O., Suárez-Calvet, M., Minguillon, C., Fauria, K., Navarro, A., & Hinzen, W. (Under Review). Atypical cortical hierarchy in Aβ-positive older adults in the context of atypical speech prosody. *Neurobiology of Aging*.
13. Lofgren, M., **Al-Tamimi, J.**, Chapin, K., García-Gutiérrez, F., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J., Sanabaria, A., Alegret, M., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (Under Review). An automated acoustic-prosodic analysis pipeline for identifying early cognitive decline and Alzheimer's disease. *Journal of Psychiatric Research*.
14. Coretta, S., Casillas, J. V., Roessig, S., Franke, M., Ahn, B., Al-Hoorie, A. H., **Al-Tamimi, J.**, Alotaibi, N. E., AlShakhori, M. K., Altmiller, R. M., Arantes, P., Athanasopoulou, A., Baese-Berk, M. M., Bailey, G., Sangma, C. B. A., Beier, E. J., Benavides, G. M., Benker, N., BensonMeyer, E. P., Benway, N. R., Berry, G. M., Bing, L., Bjorndahl, C., Bolyanatz, M., Braver, A., Brown, V. A., Brown, A. M., Brugos, A., Buchanan, E. M., Butlin, T., Buxó-Lugo, A., Caillol, C., Cangemi, F., Carignan, C., Carraturo, S., Caudrelier, T., Chodroff, E., Cohn, M., Cronenberg, J., Crouzet, O., Dagar, E. L., Dawson, C., Diantoro, C. A., Dokovova, M., Drake, S., Du, F., Dubuis, M., Duême, F., Durward, M., Egurtzegi, A., Elsherif, M. M., Esser, J., Ferragne, E., Ferreira, F., Fink, L. K., Finley, S., Foster, K., Foulkes, P., Franzke, R., Frazer-McKee, G., Fromont, R., García, C., Geller, J., Grasso, C. L., Greca, P., Grice, M., Grose-Hodge, M. S., Gully, A. J., Halfacre, C., Hauser, I., Hay, J., Haywood, R., Hellmuth, S., Hilger, A. I., Holliday, N., Hoogland, D., Huang, Y., Hughes, V., Icardo Isasa, A., Ilchovska, Z. G., Jeon, H.-S., Jones, J., Junges, M. N., Kaefer, S., Kaland, C., Kelley, M. C., Kelly, N. E., Kettig, T., Khattab, G., Koolen, R., Krahmer, E., Krajewska, D., Krug, A., Kumar, A. A., Lander, A., Lentz, T. O., Li, W., Li, Y., Lialiou, M., Lima, R. M., Lo, J. J. H., Lopez Otero, J. C., Mackay, B., MacLeod, B., Mallard, M., McConnellogue, C.-A. M., Moroz, G., Murali, M., Nalborczyk, L., Nenadić, F., Nieder, J., Nikolić, D., Nogueira, F. G. S., Offerman, H. M., Passoni, E., Pélissier, M., Perry, S. J., Pfiffner, A. M., Proctor, M., Rhodes, R., Rodríguez, N., Roepke, E., Röer, J. P., Sbacco, L., Scarborough, R., Schaeffler, F., Schleef, E., Schmitz, D., Shiryaev, A., Sóskuthy, M., Spaniol, M., Stanley, J. A., Strickler, A., Tavano, A., Tomaschek, F., Tucker, B. V., Turnbull, R., Ugwuanyi, K. O., Urrestarazu-Porta, I., Van De Vijver, R., Van Engen, K. J., Van Miltenburg, E., Wang, B. X., Warner, N., Wehrle, S., Westerbeek, H., Wiener, S., Winters, S., Wong, S. G.-J., Wood, A., Wottawa, J., Xu, C., Zárate-Sández, G., Zellou, G., Zhang, C., Zhu, J., & Roettger, T. B.

(2023). Multidimensional Signals and Analytic Flexibility: Estimating Degrees of Freedom in Human-Speech Analyses. *Advances in Methods and Practices in Psychological Science*, *6*(3), 25152459231162567. doi: 10.1177/25152459231162567

15. He, R., Chapin, K., **Al-Tamimi, J.**, Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (2023). Automated classification of cognitive decline and probable Alzheimer's dementia across multiple speech and language domains. *American Journal of Speech-Language Pathology*, *32*(5), 2075–2086. doi: 10.1044/2023_AJSLP-22-00403
16. **Al-Tamimi, J.** (2023). Discussion in: Lara, Andres Felipe & Pillot-Loiseau, Claire (auth.) "Phonetic convergence and vowel overlap". *Radical: A Journal of Phonology*, *3*, 270–279. Retrieved from https://radical.cnrs.fr/wp-content/uploads/2023/05/LaraPillot-Loiseau_2023.pdf
17. 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: Special Issue on the "Phonetics of Under-Documented Languages"*, *147*(4), 2917–2927. doi: 10.1121/10.0001004
18. Heyne, M., Derrick, D., & **Al-Tamimi, J.** (2019). Native language influence on brass instrument performance: An application of generalized additive mixed models (GAMMs) to midsagittal ultrasound images of the tongue. *Frontiers in Psychology*, *10*(2597), 1–26. doi: 10.3389/fpsyg.2019.02597
19. **Al-Tamimi, J.**, & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics, Invited manuscript for the special issue of Journal of Phonetics, "Marking 50 Years of Research on Voice Onset Time and the Voicing Contrast in the World's Languages" (eds., T. Cho, G. Docherty & D. Whalen)*, *71*, 306–325. doi: 10.1016/j.wocn.2018.09.010
20. Khattab, G., **Al-Tamimi, J.**, & Alsiraih, W. (2018). Nasalisation in the production of Iraqi Arabic pharyngeals. *Phonetica*, *75*(4), 310–348. doi: 10.1159/000487806
21. **Al-Tamimi, J.** (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1), 1–40. doi: 10.5334/labphon.19
22. **Al-Tamimi, J.**, & Khattab, G. (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. *The Journal of the Acoustical Society of America*, *138*(1), 344–360. doi: 10.1121/1.4922514
23. Khattab, G., & **Al-Tamimi, J.** (2014). Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *5*(2), 231–269. doi: 10.1515/lp-2014-0009
24. Punnoose, R., Khattab, G., & **Al-Tamimi, J.** (2013). The contested fifth liquid in Malayalam: A window into the lateral-rhotic relationship in dravidian languages. *Phonetica*, *70*(4), 274–297. doi: 10.1159/000356359
25. **Al-Tamimi, J.** (2013). Articulatory Phonetics Edited by Bryan Gick, Ian Wilson Donald Derrick (Malden, MA: Wiley Blackwell, 2013) [Pp. 272] ISBN 978-1-4051-9321-4. £60.00 (hbk), ISBN: 978-1-4051-9320-7. £24.99 (pbk). *International Journal of Language &*

Communication Disorders, *48*(5), 597–598. doi: 10.1111/1460-6984.12029

26. Kitikanan, P., & **Al-Tamimi, J.** (2012). The earliest stage of voiceless fricative acquisition among Thai learners of Mandarin Chinese. *Annual Review of Education, Communication & Language Sciences*, *9*, 91–114. Retrieved from https://www.researchgate.net/publication/234036658_THE_EARLIEST_STAGE_OF_VOICELESS_FRICATIVE_ACQUISITION_AMONG_THAI_LEARNERS_OF_MANDARIN_CHINESE
27. Khattab, G., & **Al-Tamimi, J.** (2008). Durational cues for gemination in Lebanese Arabic. *Language and Linguistics*, *22*, 39–55. Retrieved from https://www.researchgate.net/publication/256295296_Durational_Cues_for_Gemination_in_Lebanese_Arabic

2.2 Chapters Published in Peer Reviewed Books (4)

1. Ehbara, H., Young-Scholten, M., & **Al-Tamimi, J.** (2021). The role of delayed output on Second/Foreign language pronunciation in children. In C. N. Giannikas (Ed.), *Teaching practices and equitable learning in children's language education* (pp. 23–44). IGI Global. doi: 10.4018/978-1-7998-6487-5.ch002 Retrieved from https://www.researchgate.net/publication/349360413_The_Role_of_Delayed_Output_on_SecondForeign_Language_Pronunciation_in_Children
2. Khattab, G., & **Al-Tamimi, J.** (2013). Influence of geminate structure on early Arabic templatic patterns. In M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic Evidence* (pp. 374–414). Cambridge University Press. Retrieved from https://www.researchgate.net/publication/259889502_Influence_of_geminate_structure_on_early_Arabic_templatic_patterns
3. **Al-Tamimi, J.** (2011). La palatalisation en coréen comme exemple de coarticulation : Données ciné-IRM stroboscopiques et acoustiques des mouvements graduels de la langue, de " (translation to French from: ", Hyunsoon KIM, Korean Palatalization as Coarticulation: Stroboscopic cine-MRI and acoustic data on gradual tongue movements). In M. Embarki & C. Dodane (Eds.), *La coarticulation. Des indices à la représentation* (pp. 253–270). L'Harmattan.
4. **Al-Tamimi, J.**, & Barkat-Defradas, M. (2002). Inter-dialectal and inter-individual variability in production and perception: A preliminary study in Jordanian and Moroccan Arabic. In I. Ferrando & J. J. S. Sanchez (Eds.), *Actes des 5èmes rencontres de l'Association internationale de dialectologie arabe (AIDA), cadiz, espagne* (pp. 171–186). Retrieved from https://www.researchgate.net/publication/256295295_Inter-dialectal_and_inter-individual_variability_in_production_and_perception_a_preliminary_study_in_Jordanian_and_Moroccan_Arabic

2.3 Articles Published in Peer Reviewed Conference Proceedings (24)

1. **Al-Tamimi, J.**, & Palo, P. (2023). Dynamics of the tongue contour in the production of guttural consonants in Levantine Arabic. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 2095–2099). Prague, Czech Republic (7-11 August 2023): Guarant International.
2. Almurashi, W., **Al-Tamimi, J.**, & Khattab, G. (2023). Dynamic and static features of English monophthongal vowels production by Hijazi Arabic L2 learners. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 2562–2566). Prague, Czech Republic (7-11 August 2023): Guarant International.
3. Dallak, A., Khattab, G., & **Al-Tamimi, J.** (2023). Obstruent Voicing and Laryngeal Features in Arabic. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 2154–2158). Prague, Czech Republic (7-11 August 2023): Guarant International.
4. Li, C., **Al-Tamimi, J.**, & Wu, Y. (2023). Tone as a factor influencing the dynamics of diphthong realizations in Standard Mandarin. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 1876–1880). Prague, Czech Republic (7-11 August 2023): Guarant International.
5. **Al-Tamimi, J.**, Schiel, F., Khattab, G., Sokhey, N., Amazouz, D., Dallak, A., & Moussa, H. (2022). A Romanization System and WebMAUS Aligner for Arabic Varieties. In *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), © European Language Resources Association (ELRA), licensed under CC-BY-NC-4.0* (pp. 7269–7276). Marseille, 20-25 June 2022. Retrieved from http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.789.pdf
6. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (2019). The acoustic properties of laryngeal contrast in Najdi Arabic initial stops. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 2051–2055). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc. Retrieved from https://www.researchgate.net/publication/334965210_The_Acoustic_Properties_of_Laryngeal_Contrast_in_Najdi_Arabic_Initial_Stops
7. Almurashi, W., **Al-Tamimi, J.**, & Khattab, G. (2019). Static and dynamic cues in vowel production in Hijazi Arabic. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 3468–3472). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc. Retrieved from https://www.researchgate.net/publication/335017406_Static_and_dynamic_cues_in_vowel_production_in_Hijazi_Arabic
8. Chen, Y., Khattab, G., & **Al-Tamimi, J.** (2019). The effect of explicit instruction and auditory/audio-visual training on Chinese learners' acquisition of English intonation.

In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 2286–2290). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc. Retrieved from https://www.researchgate.net/publication/335016177_The_effect_of_explicit_instruction_and_auditoryaudio-visual_training_on_Chinese_learners'_acquisition_of_English_intonation

9. McCarthy, D., & **Al-Tamimi, J.** (2019). F2R: A Technique for collapsing F2 onset and F2 mid into a single acoustic attribute. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 3777–3781). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc. Retrieved from https://www.researchgate.net/publication/335018542_F2R_A_Technique_for_Collapsing_F2onset_and_F2mid_into_a_Single_Acoustic_Attribute
10. **Al-Tamimi, J.** (2015). Spectral tilt as an acoustic correlate to pharyngealisation in Jordanian and Moroccan Arabic (Article: 0436). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015. Retrieved from https://www.researchgate.net/publication/280715904_Spectral_tilt_as_an_acoustic_correlate_to_pharyngealisation_in_Jordanian_and_Moroccan_Arabic
11. Best, C., Shaw, J., Docherty, G., Evans, B., Foulkes, P., Hay, J., **Al-Tamimi, J.**, Mair, K., Mulak, K., & Wood, S. (2015). From newcastle MOUTH to Aussie ears: Australians' perceptual assimilation and adaptation for Newcastle UK vowels. In *Proceedings of the Annual Conference of the International Speech Communication Association (ISCA), INTERSPEECH* (pp. 1932–1936). Dresden, Germany (6-10 September 2015). Retrieved from https://www.researchgate.net/publication/299340499_From_Newcastle_MOUTH_to_Aussie_ears_Australians'_perceptual_assimilation_and_adaptation_for_Newcastle_UK_vowels
12. Best, C., Shaw, J., Mulak, K., Docherty, G., Evans, B., Foulkes, P., Hay, J., **Al-Tamimi, J.**, Mair, K., & Wood, S. (2015). Perceiving and adapting to regional accent differences among vowel subsystems (Article 0964). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015. Retrieved from https://www.researchgate.net/publication/280716113_Perceiving_and_adapting_to_regional_accent_differences_among_vowel_subsystems
13. Khattab, G., & **Al-Tamimi, J.** (2015). The acquisition of gemination in Lebanese Arabic children (article: 0870). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015. Retrieved from https://www.researchgate.net/publication/280716101_The_acquisition_of_gemination_in_Lebanese_Arabic_Children

14. Kitikanan, P., **Al-Tamimi, J.**, & Khattab, G. (2015). An acoustic investigation of the production of English/s/by L2 Thai learners (Article: 0655). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015. Retrieved from https://www.researchgate.net/publication/280716178_The_production_of_English_s_by_L2_Thai_learners
15. Shaw, J. A., Best, C. T., Mulak, K. E., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., **Al-Tamimi, J.**, Mair, K., & Peek, M. (2014). Effects of short-term exposure to unfamiliar regional accents: Australians' categorization of London and Yorkshire English consonants. In *Proceedings of the 15th Australasian International Speech Science and Technology conference.* (pp. 72–75). Christchurch, New Zealand (6-9 December 2014). Retrieved from https://www.researchgate.net/publication/270897401_Effects_of_short-term_exposure_to_unfamiliar_regional_accents_Australians'_categorization_of_London_and_Yorkshire_English_consonants
16. **Al-Tamimi, J.**, & Khattab, G. (2011). Multiple cues for the singleton-geminate contrast in Lebanese Arabic: Acoustic investigation of stops and fricatives. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)* (pp. 212–215). Hong Kong, China (17-21 August 2011). Retrieved from https://www.researchgate.net/publication/256295420_Multiple_cues_for_the_singleton-geminate_contrast_in_Lebanese_Arabic_acoustic_investigation_of_stops_and_fricatives
17. **Al-Tamimi, J.** (2008). L'espace vocalique perceptif dépend de la densité des systèmes vocaliques: Étude translinguistique en arabe marocain, en arabe jordanien et en français. In *Actes des 27èmes journées d'Études sur la Parole (JEP)* (pp. 1–4). Avignon, France (9-13 June 2008). Retrieved from https://www.researchgate.net/publication/256295426_L'espace_vocalique_perceptif_depend_de_la_densite_des_systemes_vocaliques_Etude_translinguistique_en_arabe_marocain_en_arabe_jordanien_et_en_francais
18. **Al-Tamimi, J.** (2007). Static and dynamic cues in vowel production: A cross dialectal study in Jordanian and Moroccan Arabic. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)* (pp. 541–544). Saarbrücken, Germany (6-10 August 2007). Retrieved from https://www.researchgate.net/publication/256295422_Static_and_Dynamic_cues_in_vowel_production_a_cross_dialectal_study_in_Jordanian_and_Moroccan_Arabic?
19. Dodane, C., & **Al-Tamimi, J.** (2007). An acoustic comparison of vowel systems in adult-directed-speech and child-directed-speech: Evidence from french, english & japanese. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)* (pp. 1573–1576). Saarbrücken, Germany (6-10 August 2007). Retrieved from https://www.researchgate.net/publication/228593556_An_acoustic_comparison_of_vowel_systems_in_adult-directed_speech_and_child-directed_speech_Evidence_from_French_English_Japanese
20. **Al-Tamimi, J.** (2006). Analyse dynamique de la réduction vocalique en contexte CV à partir

des pentes formantiques en arabe dialectal et en français. In *Actes des 26èmes journées d'Études sur la Parole (JEP)* (pp. 357–360). Dinard, France (12-16 June 2006). Retrieved from https://www.researchgate.net/publication/237341111_Analyse_dynamique_de_la_reduction_vocalique_en_contexte_CV_a_partir_des_pentes_formantiques_en_arabe_dialectal_et_en_francais

21. **Al-Tamimi, J.**, & Ferragne, E. (2005). Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. In *Proceedings of the 9th Eurospeech (European Conference on Speech Communication and Technology)* (pp. 2465–2468). Lisbon, Portugal (4_8 September 2005). Retrieved from https://www.researchgate.net/publication/221489776_Does_vowel_space_size_depend_on_language_vowel_inventories_Evidence_from_two_Arabic_dialects_and_French
22. **Al-Tamimi, J.** (2004). L'équation du locus comme mesure de la coarticulation VC et CV: Étude préliminaire en Arabe Dialectal Jordanien. In *Actes des 25èmes journées d'Études sur la Parole (JEP)* (pp. 9–12). Fès, Morocco (19-22 April 2004). Retrieved from https://www.researchgate.net/publication/256295418_L'equation_du_locus_comme_mesure_de_la_coarticulation_VC_et_CV_Etude_preliminaire_en_Arabe_Dialectal_Jordanien
23. Barkat-Defradas, M., **Al-Tamimi, J.**, & Benkirane, T. (2003). Phonetic variation in production and perception of speech: A comparative study of two Arabic dialects. In *Prroceedings of the 15th International Congress of Phonetic Sciences (ICPhS).* Barcelona, Spain (3-9 August 2003). Retrieved from (https://www.researchgate.net/publication/253927754_Phonetic_variation_in_production_and_perception_of_speech_a_comparative_study_of_two_Arabic_dialects
24. **Al-Tamimi, J.**, Girard, M., & Marsico, E. (2002). Variabilité inter-langue et inter-individuelle en production et en perception: étude préliminaire en arabe dialectal et en français. In *Actes des 24èmes Journées d'Études sur la Parole (JEP)* (pp. 169–172). Nancy, France (June 2002). Retrieved from https://www.researchgate.net/publication/256295189_Variabilite_inter-langue_et_inter-individuelle_en_production_et_en_perception_Etude_preliminaire_en_arabe_dialectal_et_en_francais

2.4 Presentations in Peer Reviewed Conferences (54)

1. Ghadanfari, S., White, L., **Al-Tamimi, J.**, & Khattab, G. (2023). Effects underlying hierarchical speech timing in Kuwaiti Arabic dialects. In *Proceedings of SPaB2023: Speech Prosody and Beyond.* Seoul, South Korea, (20-22 June 2023).
2. **Al-Tamimi, J.** (2023). Différents stades d'Alzheimer : Rôle des méthodes automatisées de traitement de la prosodie et de la qualité de la voix. In *Actes de les 9èmes Journées de Phonétique Clinique (JPC2023): "Prendre la mesure de la Parole".* Toulouse 15-17 juin 2023.
3. Jordan, E., & **Al-Tamimi, J.** (2023). Classification d'Alzheimer à partir de paramètres acoustiques et prosodique avec de l'apprentissage automatique. In *Actes de les 9èmes Journées de*

Phonétique Clinique (JPC2023): "Prendre la mesure de la Parole". Toulouse 15-17 juin 2023.

4. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (2022b). Voicing contrast in Najdi Arabic stops: Implications for Laryngeal realism. In *Proceedings of the 18th LabPhon (Laboratory Phonology) Conferene.* Online conference (23-25 June 2022).
5. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (2022a). Voicing contrast in Najdi Arabic stops: Implications for Laryngeal realism. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, UK (4-8 April 2022).
6. Ghadanfari, S., White, L., **Al-Tamimi, J.**, & Khattab, G. (2022). Hierarchical timing in two varieties of Kuwaiti Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, UK (4-8 April 2022).
7. **Al-Tamimi, J.** (2021). Acoustic correlates of epilaryngeal constriction in Levantine Arabic guttural consonants. In *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).
8. Almurashi, W., **Al-Tamimi, J.**, & Khattab, G. (2021). Acoustic analysis of the acquisition of english vowels by Hijazi Arabic learners. In *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).
9. Jones, W., **Al-Tamimi, J.**, & Hinzen, W. (2021). The role of automatic acoustic profiling of prosody in schizophrenia. In *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).
10. **Al-Tamimi, J.**, Heyne, M., & Derrick, D. (2020b). From SS-ANOVA to GAMMs: Accounting for within and between-subject variation using generalized additive mixed models on ultrasound tongue contours. In *Proceedings of the 12th International Seminar on Speech Production (ISSP).* Online Virtual Conference (14-18 December 2020).
11. **Al-Tamimi, J.**, & Palo, P. (2020). Quantifying gradience of epilaryngeal constriction in Levantine Arabic "gutturals": A Generalized Additive Modelling approach to ultrasound tongue contours. In *Proceedings of the 12th International Seminar on Speech Production (ISSP).* Online Virtual Conference (14-18 December 2020).
12. **Al-Tamimi, J.**, & Ferragne, E. (2020b). The phonetic basis of the guttural natural class in Levantine Arabic: Evidence from coarticulation and energy components using Deep Learning and Random Forests. In *Proceedings of the 17th LabPhon (Laboratory Phonology) Conference.* Vancouver, Canada (5-9 July 2020).
13. **Al-Tamimi, J.**, & Ferragne, E. (2020a). Deep Learning and Random Forests show acoustic-phonetic evidence supporting the guttural natural class in Levantine Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).
14. **Al-Tamimi, J.**, Heyne, M., & Derrick, D. (2020a). Accounting for within and between-subject variation using Generalized Additive Mixed Models: Application on Ultrasound Tongue contours with and without normalization. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).
15. Almurashi, W., **Al-Tamimi, J.**, & Khattab, G. (2020). Acoustic cues in production of english vowels by Hijazi Arabic learners. In *Proceedings of BAAP (The British Association of*

Academic Phoneticians). York University, York, UK (1-3 April 2020; postponed due to Covid).

16. Jones, W., & **Al-Tamimi, J.** (2020). Using Prosogram's automatic prosodic profiling in diagnosing schizophrenia. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).
17. Turton, D., Khattab, G., Alsharif, B., & **Al-Tamimi, J.** (2020). Syllabic and emphatic conditioning of /l/ in Levantine Arabic: An auditory, acoustic and articulatory analysis. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).
18. Heyne, M., **Al-Tamimi, J.**, & Derrick, D. (2019). Using Generalized Additive (Mixed) Models to uncover novel patterns of speech production and account for by-group random patterns of variation. In *Proceedings of the Boston Speech Motor Control Symposium.* Boston University, Boston, USA (21 June 2019).
19. **Al-Tamimi, J.** (2018a). A Generalised Additive Modelling approach to ultrasound tongue surface: Quantifying retraction in Levantine Arabic back consonants. In *Proceedings of the 16th LabPhon (laboratory phonology) satellite workshop: New developments in speech sensing and imaging.* Lisbon, Portugal (23 June 2018).
20. Khattab, G., & **Al-Tamimi, J.** (2018). When lenis and fortis properties collide: The case of voiced geminates in Lebanese Arabic. In *Proceedings of the 16th LabPhon (laboratory phonology).* Lisbon, Portugal (20-22 June 2018).
21. Al-Gamdi, N., Khattab, G., & **Al-Tamimi, J.** (2018). Voicing contrast in Najdi Arabic initial stops. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).
22. **Al-Tamimi, J.** (2018b). Gradient epilaryngeal constriction in Levantine Arabic "gutturals": A Generalised Additive Modelling approach to ultrasound tongue surface. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).
23. **Al-Tamimi, J.**, & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate plosives. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).
24. Almurashi, W., Khattab, G., & **Al-Tamimi, J.** (2018). Static and dynamic cues in vowel production in Hijazi Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).
25. Ehbara, H., & **Al-Tamimi, J.** (2018). The impact of computer assisted pronunciation training on Libyan child learners of english. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).
26. McCarthy, D., & **Al-Tamimi, J.** (2018). The role of frequency and magnitude scales for identifying plosives' place of articulation. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).
27. Moussa, H., & **Al-Tamimi, J.** (2018). Focus and F0 patterns in Jeddah Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).

28. Khattab, G., & **Al-Tamimi, J.** (2017). Contact-induced constraints on the acquisition of phonological length contrast by Lebanese Arabic-speaking children. In *Proceedings of the IASCL (international association for the study of child language).* Lyon, France (17-21 July 2017).
29. **Al-Tamimi, J.** (2017). Acoustics of constricted epilarynx: The case of voiced coronal pharyngealized consonants in Jordanian and Moroccan Arabic. In *Proceedings of the 2nd PaPE (Phonetics and Phonology in Europe).* Cologne, Germany (12-14 June 2017).
30. McCarthy, D., & **Al-Tamimi, J.** (2017). Place of articulation in plosives: A light on the mapping between phonetics and phonology. In *Proceedings of the 2nd PaPE (Phonetics and Phonology in Europe).* Cologne, Germany (12-14 June 2017).
31. **Al-Tamimi, J.** (2016). Is pharyngealization associated with an epilaryngeal constriction? An acoustic investigation in Jordanian and Moroccan Arabic. In *Proceedings of the Arabic linguistic forum (ALiF).* York University, UK (12-14 December 2016).
32. Chen, Y., Khattab, G., & **Al-Tamimi, J.** (2016a). Auditory or audio-visual feedback in EFL intonation training. In *Proceedings of the New Sounds 2016: 8th international conference on second-language speech. Aarhus university, denmark.* Aarhus University, Denmark (10-12 June 2016).
33. Chen, Y., Khattab, G., & **Al-Tamimi, J.** (2016b). Comprehension of english intonation by chinese EFL learners. In *Proceedings of BAAP (The British Association of Academic Phoneticians) colloquium.* Lancaster University, Lancaster, UK, (31 March-2nd April 2016).
34. Khattab, G., & **Al-Tamimi, J.** (2015). Spectral properties of vowels surrounding singleton and geminate sonorants in Lebanese Arabic. In *In Proceedings of GemCon2015 - ICPhS 2015 satellite workshop on 'Geminate consonants across the world'.* Glasgow University, Glasgow, UK (12 August 2015).
35. Khattab, G., Vihman, M., **Al-Tamimi, J.**, Nakai, S., & Kunnari, S. (2014). A comparison of long consonant acquisition in Arabic, Finnish, Japanese and Welsh. In *Proceedings of the 14th Laboratory Phonology Conference (LabPhon14).* Tokyo, Japan (25-27 July 2014).
36. **Al-Tamimi, J.** (2014). An acoustic investigation of pharyngealisation effects on vowels: A cross-dialectal investigation in Jordanian and Moroccan Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* Oxford University, Oxford, UK (7-9 April 2014).
37. Khattab, G., & **Al-Tamimi, J.** (2014). Disentangling phonetic from phonological length: The effect of multilingual input on Lebanese-Arabic-learning children's production. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* Oxford University, Oxford, UK (7-9 April 2014).
38. Khattab, G., & **Al-Tamimi, J.** (2013). The role of French loan word adaptation in the acquisition of consonant length in Lebanese Arabic. In *Proceedings of the International Child Phonology Conference (ICPC).* Radboud University, Nijmegen, Netherlands (10-12 June 2013).
39. Kitikanan, P., **Al-Tamimi, J.**, & Young-Scholten, M. (2012). The earliest stage of fricative acquisition among Thai learners of Mandarin Chinese. In *Proceedings of the Second Language Acquisition of Phonology Conference.* York University, York, UK (6 July 2012).
40. Khattab, G., & **Al-Tamimi, J.** (2012b). Learning to control phonetic timing for phonological

purposes: The acquisition of gemination by Lebanese-speaking children. In *Proceedings of the International Child Phonology Conference (ICPC).* Minneapolis, USA (04-06 June 2012).

41. Khattab, G., & **Al-Tamimi, J.** (2012a). Learning to control phonetic timing for phonological purposes: The acquisition of gemination by Lebanese-speaking children. In *Proceedings of BAAP (The British Association of Academic Phoneticians Colloquium,.* Leeds University, Leeds, UK (26-28 March 2012).
42. Khattab, G., & **Al-Tamimi, J.** (2011b). Phonological templates in lebanese arabic. In *Proceedings of the 12th International Association of the Study of Child Language conference (IASCL),.* Montreal, Canada (19-23 July 2011).
43. Khattab, G., & **Al-Tamimi, J.** (2011c). The role of geminates in shaping early word patterns by Lebanese-Arabic speaking children. In *Proceeings of the International Child Phonology Conference.* York University, UK (15-18 June 2011).
44. Khattab, G., & **Al-Tamimi, J.** (2011a). Phonological patterns in early word production by Lebanese-Arabic speaking children. In *Proceedings of the Child Language Seminar.* Newcastle University, UK (13-14 June 2011).
45. **Al-Tamimi, J.** (2010). Static and dynamic cues in vowel production and perception: Cross-linguistic investigation of moroccan arabic, jordanian arabic and french. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* London, UK (29-31 March 2010).
46. **Al-Tamimi, J.**, & Khattab, G. (2010). Multiple cues for the singleton-geminate contrast in Lebanese Arabic: The role of non-temporal characteristics. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* London, UK (29-31 March 2010).
47. **Al-Tamimi, J.** (2009). Effect of pharyngealisation on vowels revisited: Static and Dynamic analyses of vowels in Moroccan and Jordanian Arabic. In *Proceedings of the International Workshop on Pharyngeals & Pharyngealisation.* Newcastle University, UK (26-27 Mars 2009).
48. Khattab, G., & **Al-Tamimi, J.** (2009). Phonetic cues to gemination in Lebanese Arabic. In *Proceedings of the 17th Manchester Phonology Meeting,.* Manchester, UK (28-30 May 2009).
49. Khattab, G., & **Al-Tamimi, J.** (2008a). Phonetic cues to gemination in Lebanese Arabic. In *Proceedings of the 11th LabPhon (Laboratory Phonology) conference,.* Wellington, New Zealand (30 June-2 July 2008).
50. Khattab, G., & **Al-Tamimi, J.** (2008b). Phonetic patterns of gemination in Lebanese Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium,.* Sheffield, UK (31 March-2 April 2008).
51. **Al-Tamimi, J.** (2007b). Rôle des indices statiques et dynamiques dans la classification des voyelles en arabe marocain et arabe jordanien : étude acoustico-perceptive. In *Actes de la conférence "Typologie des parlers arabes modernes",.* Montpellier, France (14-15 May 2007).
52. **Al-Tamimi, J.** (2007a). Production et perception des voyelles en arabe marocain et arabe jordanien : Pertinence des indices statiques et dynamiques. In *Actes de la 1ères Journées des Sciences de la Parole.* Charleroi, Belgium (30-31 March 2007).
53. **Al-Tamimi, J.** (2006). Indices dynamiques et perception des voyelles : Étude translinguistique

en arabe dialectal et en français. In *Actes de l'école d'été : Voix, parole & langues,*. Cargèse, France (4-9 June 2006).

54. **Al-Tamimi, J.**, Carré, R., & Marsico, E. (2004). The status of vowels in Jordanian and Moroccan Arabic: Insights from production and perception. In *Proceedings of the 148th meeting of the Acoustical Society of America (ASA); JASA, vol. 116, no. 4, p. 2629,*. San Diego, USA (15-19 November 2004).

2.5 Thesis (3)

1. **Al-Tamimi, J.** (2007). *Indices dynamiques et perception des voyelles: étude translinguistique en arabe dialectal et en français* (PhD Thesis in Language Sciences - Phonetics and Phonology, Université Lyon 2). Retrieved from http://theses.univ-lyon2.fr/documents/lyon2/2007/al-tamimi_je
2. **Al-Tamimi, J.** (2002). *Variabilité phonétique en production et en perception de la parole : Le cas de l'arabe jordano-palestinien* (Master Thesis - DEA in Language Sciences - Phonetics and Phonology). Université Lyon 2, Lyon.
3. **Al-Tamimi, J.** (2001). *Description linguistique de 6 parlers jordaniens : étude des interférences en production entre l'arabe dialectal jordanien et l'arabe classique ; une approche phonético-phonologique et acoustique* (Master Thesis - Maîtrise in Language Sciences - Phonetics and Phonology). Université Lyon 2, Lyon.

2.6 Invited speaker (15)

1. **Al-Tamimi, J.** (2023). *Alzheimer and pre-AD: Can automated speech analyses serve as potential bio-markers?* [LingLunch Seminar, Laboratoire de Linguistique Formelle, Université Paris Cité (12 January 2023)].
2. **Al-Tamimi, J.** (2022a). *A whole tongue approach to gutturals in Levantine Arabic using Generalized Additive Mixed Modelling of Tongue surfaces* [Séminaire de Recherche En Phonétique et Phonologie, Laboratoire de Phonétique et Phonologie (SRPP-LPP) - Paris 3 (7th October 2022)].
3. **Al-Tamimi, J.** (2022b). *On the contribution of automatic methods at quantifying the impact of pathologies on speech* [Symposium Labex EFL (Emperical Foundations of Linguistics)]. Paris, France (June 2022).
4. **Al-Tamimi, J.** (2020a). *LingLab Methods Fair: Ultrasound Tongue Imaging, Electroglottography, Electropalatography, and acoustics* [Workshop Delivred to Researchers and Doctoral Students in Linguistics and Phonetic Sciences, Newcastle University (January 2020)].
5. **Al-Tamimi, J.** (2020b). *The role of the epilarynx in the production of guttural consonants in Levantine Arabic* [Speech Science Forum, Hearing and Phonetic Sciences – University College London, UK (November 2020)]. Retrieved from https://www.ucl.ac.uk/pals/events/2020/nov/19th-november-ssf-jalal-al-tamimi
6. **Al-Tamimi, J.** (2019a). *The role of the epilarynx in male/female and dialect discrimination in dialectal Arabic/Le rôle de l'épilarynx dans la discrimination entre locuteurs*

hommes/femmes et entre dialectes arabes [Séminaire de Recherche : Traitement Du Langage Parlé (TLP), Laboratoire d'Informatique Pour La Mécanique et Les Sciences de l'Ingénieur (LIMSI) - Université Paris-Sud, France (May 2019)]. Retrieved from https://www.limsi.fr/fr/component/content/article/75-recherche/equipes-de-recherche/tlp-categorie/tlp-seminaires/854-jalal-al-tamimi-newcastle-university?Itemid=106

7. **Al-Tamimi, J.** (2019b). *The role of the epilarynx in the production of gutturals in Levantine Arabic/Rôle de l'épilarynx dans la production des gutturales en Arabe Levantin* [Séminaire de Recherche En Phonétique et Phonologie, Laboratoire de Phonétique et Phonologie (SRPP-LPP) - Paris 3 (March 2019)]. Retrieved from https://roland.ilpga.fr/evenement/le-role-de-lepilarynx-dans-la-production-des-gutturales-en-arabe-levantin/
8. **Al-Tamimi, J.** (2018a). *Gradient epilaryngeal constriction in Levantine Arabic "gutturals": A Generalised Additive Modelling approach to ultrasound tongue surface* [Atelier « Temporalité et Séquentialité Dans Les Formes Sonores », Nantes, France (June 2018)]. Retrieved from https://framaforms.org/atelier-temporalite-et-sequentialite-dans-les-formes-sonores-27-juin-2018-nantes-france-1527251630
9. **Al-Tamimi, J.** (2018b). *A multifaceted approach to Child and Adult phonology: From Auditory, articulatory, acoustic, and perceptual approaches to speech recognition* [Research Seminar - A Joint Child-Adult Research Seminar, Speech and Language Sciences, Newcastle University, UK (January 2018)].
10. **Al-Tamimi, J.** (2016). *Experimental issues surrounding data collection/analyses in Arabic and statistical designs: Tutorial* [Arabic Linguistic Forum (ALiF), York University, UK (13 December 2016)]. Retrieved from http://arabiclinguisticsforum.com/york-2016/
11. **Al-Tamimi, J.** (2012). *The role of multiple acoustic cues in the perception, production and learning of language-specific phonological patterns* [Research Seminar in Linguistics, York University, UK (April 2012)].
12. **Al-Tamimi, J.** (2010a). *Aspects phonétiques et phonologiques de la gémination en Arabe Libanais* [Séminaire Du Laboratoire Dynamique Du Language (DDL), Lyon, France (May 2010)].
13. **Al-Tamimi, J.** (2010b). *Phonetic and Phonological Aspects of Gemination in Lebanese Arabic* [The CRiLLS-Heidlberg Research Seminar, Newcastle University, UK (June 2010)].
14. **Al-Tamimi, J.** (2007). *Rôle des indices dynamiques dans la caractérisation des voyelles en arabe dialectal et en français : De la production à la perception* [Séminaire de Recherche : Traitement Du Langage Parlé (TLP), Laboratoire d'Informatique Pour La Mécanique et Les Sciences de l'Ingénieur (LIMSI), September 2007)].
15. **Al-Tamimi, J.** (2005). *The Locus Equation : Un indice du lieu d'articulation et/ou du degré de coarticulation* [Séminaire de Recherche de l'équipe Identification, Laboratoire Dynamique Du Langage (DDL), Lyon (May 2005)].

2.7 Workshops and training events (24)

1. **Al-Tamimi, J.** (2022). *R training. From basics to advanced topics (Introduction, tidyverse, visualisation, inferential statistics (LM, GLM, CLM, LMM), PCA, decision trees, random forests)* [Training Delivered as Part of the LRPG (Laboraotry Phonology Research Group), Université Paris Cité]. Paris, France. Retrieved from https://jalalal-tamimi.github.io/R-Training/
2. **Al-Tamimi, J.** (2021a). *Introduction to Random Forests* [Workshop Delivered to the Phonetics and Phonology Research Group at Aarhus University (June 2021)]. Retrieved from https://jalalal-tamimi.github.io/Intro-Random-Forests/
3. **Al-Tamimi, J.** (2021b). *Session 4 : Techniques in data analyses* [Workshop Adventure in R: Training in Using R Delivered to PhD Students in Humanities and Sociale Sciences, Newcastle University (June 2021)]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_3-AnalysingData2021.nb.html
4. **Al-Tamimi, J.** (2020a). *Introduction to High Power Computing (HPC)* [Workshop Delivered to the Phonetics and Phonology Research Group at Newcastle University (November 2020)].
5. **Al-Tamimi, J.** (2020b). *Introduction to Praat: Voice analysis* [Workshop Delivered to the Voice Forum, Newcastle University (January 2020)].
6. **Al-Tamimi, J.** (2020c). *Session 4 : Techniques in data analyses* [Workshop Adventure in R: Training in Using R Delivered to PhD Students in Humanities and Sociale Sciences, Newcastle University (June 2020)]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_3-AnalysingData2020.nb.html
7. **Al-Tamimi, J.** (2019a). *Introduction to R* [Workshop Adventure in R: Training in Using R Delivered to PhD Students in Humanities and Sociale Sciences, Newcastle University (June 2019)]. Retrieved from https://jalalal-tamimi.github.io/R-Introduction-to-R/
8. **Al-Tamimi, J.** (2019b). *Introduction to VoiceSauce: A program for voice analysis* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University (November 2019)].
9. **Al-Tamimi, J.** (2019c). *Praat scripting* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University (November 2019)].
10. **Al-Tamimi, J.** (2019d). *Session 4 : Techniques in data analyses* [Workshop Adventure in R: Training in Using R Delivered to PhD Students in Humanities and Sociale Sciences, Newcastle University (June 2019)]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062019.nb.html
11. **Al-Tamimi, J.** (2018a). *Praat scripting - video tutorials* [Video Tutorials Delivered to PhD Students on the IPhD in Phonetics and Phonology, Newcastle University].
12. **Al-Tamimi, J.** (2018b). *Session 4 : Techniques in data analyses* [Workshop Adventure in R: Training in Using R Delivered to PhD Students in Humanities and

Sociale Sciences, Newcastle University (June 2018)]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062018.nb.html

13. Riches, N., & **Al-Tamimi, J.** (2017). *Brief introduction to LaTeX* [Seminar of the "Centre for Research in Linguistics and Language Sciences (CRiLLS)", Newcastle University (December 2017)].
14. Riches, N., & **Al-Tamimi, J.** (2014). *Brief introduction to LaTeX* [Seminar of the "Centre for Research in Linguistics and Language Sciences (CRiLLS)", Newcastle University (December 2014)].
15. **Al-Tamimi, J.** (2013a). *An introduction to acoustic analyses with Praat - video tutorials* [Video Tutorials Delivered to UG and PG Students, Newcastle University].
16. **Al-Tamimi, J.** (2013b). *Praat scripting* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University (Various Sessions)].
17. Riches, N., & **Al-Tamimi, J.** (2013). *Brief introduction to LaTeX* [Seminar of the "Centre for Research in Linguistics and Language Sciences (CRiLLS)", Newcastle University (December 2013)].
18. **Al-Tamimi, J.** (2012a). *Introduction to Praat and Phon for non-specialists* [Workshop Delivered to PGR Students as Part of the Training Offered by the Centre for Research in Linguistics and Language Sciences (CRiLLS), Newcastle University].
19. **Al-Tamimi, J.** (2012b). *Praat scripting* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University].
20. **Al-Tamimi, J.** (2011a). *Introduction to acoustic phonetics and praat for beginners* [Workshop Delivered to the Researchers and PGR Students, York University].
21. **Al-Tamimi, J.** (2011b). *Normalisation techniques: Vowels, nasalisation and phonation measures* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University].
22. **Al-Tamimi, J.** (2009). *From field recordings to statistical analyses - various sessions* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University].
23. **Al-Tamimi, J.** (2008a). *Praat labelling tutorial: How to label data using praat* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University].
24. **Al-Tamimi, J.** (2008b). *Praat Scripting Tutorial: How to automate computations using Praat - various sessions* [Workshop Delivered to the Phonetics and Phonology Research Group, Newcastle University].

2.8 Open Source Online Material, e.g., software, Praat scripts, R notebooks, Podcast (22)

1. **Al-Tamimi, J.** (2023). *La voix comme identité* [Online material - Podcast part of Vox - La linguistique sous toutes ses formes]. Retrieved 2023-05-18, from https://podcast.ausha.co/vox-podcast-labex-efl/la-voix-comme-identite
2. **Al-Tamimi, J.** (2022a). *Praat-Spectral-Analyses-fricatives-stops* [Open Source Online Material - Praat Script for Computing Spectral Measurements for Fricatives and

Stops]. Retrieved 2022-07-18, from https://jalalal-tamimi.github.io/Praat-Spectral-Analyses-fricatives-stops/

3. **Al-Tamimi, J.** (2022b). *Praat-Spectral-Tilt* [Open Source Online Material - Praat Script for Computing Spectral Tilt Measurements for Voice Quality]. Retrieved 2022-07-18, from https://jalalal-tamimi.github.io/Praat-Spectral-Tilt/
4. **Al-Tamimi, J.** (2022c). *R-GAMM-LabPhon18* [Open Source Online Material - Supplementary Material as an Introduction to GAMMs (Generalized Additive Mixed-effects Modelling) on UTI (Ultrasound Tongue Imaging) Data. First Release in 2018; Updated in 2022]]. Retrieved from https://jalalal-tamimi.github.io/R-GAMM-LabPhon18/
5. **Al-Tamimi, J.** (2022d). *Praat-f0-Accurate-Estimation* [Open Source Online Material - Praat Scripts to Accurately Estimate F0 Curves, Using the Two-Pass Method and Then Smoothing (Original Version 2018)]. Retrieved from https://jalalal-tamimi.github.io/Praat-f0-Accurate-Estimation/
6. **Al-Tamimi, J.** (2022e). *Praat-Measurement-points* [Open Source Online Material - Praat Script to Allow for an Accurate Measurement Point (Original Version 2018)]. Retrieved from https://jalalal-tamimi.github.io/Praat-Measurement-points/
7. **Al-Tamimi, J.** (2022f). *Praat-Silence-Detection* [Open Source Online Material - Praat Script for Silence/Speech Detection Using the Two-Pass Method for F0 Estimation and Intensity Adapted to Each Speaker]. Retrieved 2022-07-18, from https://jalalal-tamimi.github.io/Praat-Silence-Detection/
8. **Al-Tamimi, J.** (2022g). *Praat-Voicing-detection* [Open Source Online Material - Praat Script for Voicing Detection Using the Two-Pass Method (Original Version 2018)]. Retrieved from https://jalalal-tamimi.github.io/Praat-Voicing-detection/
9. **Al-Tamimi, J.** (2022h). *Praat-VQ-Measurements* [Open Source Online Material - Praat Script for Computing Many Acoustic Measurements for Voice Quality]. Retrieved 2022-07-18, from https://jalalal-tamimi.github.io/Praat-VQ-Measurements/
10. **Al-Tamimi, J.** (2022i). *R training* [Open Source Online Material Aimed as a Self-Introduction to Using R: Introduction, Tidyverse, Visualisation, Inferential Statistics (LM, GLM, CLM, LMM), PCA, Decision Trees, Random Forests]. Retrieved from https://jalalal-tamimi.github.io/R-Training/
11. **Al-Tamimi, J.** (2022j). *R-Estimating Number Of Trees RF* [Open Source Online Material - This R Script Allows to Estimate the Optimal Number of Trees Required to Frow Random Forests via the Conditional Inference Trees Framework (Initial Version in 2018)]. Retrieved from https://jalalal-tamimi.github.io/R-Estimating-Number-Of-Trees-RF/
12. **Al-Tamimi, J.**, Caillol, C., Ferragne, E., & Pélissier, M. (2022). *Many Speech Analyses: Team linckia_nattereri* [Full Analyses of the Many Speech Analyses Project by the Team Members; Contains Processing Files, Praat Scripts to Extract Acoustic Measures, Exploratory and Confirmatory Statistical Analyses]. OSF. Retrieved 2022-07-04, from https://osf.io/8r6x7/
13. Schiel, F., & **Al-Tamimi, J.** (2021). *WebMAUS and WebMINNI Arabic: Online resources for forced-alignment and speech-to-text transcription* [Development of an Arabic Variety-

Independent Romanization System Integrated within an Arabic Variety-Independent WebMAUS Forced Alignment and WebMINNI Services (Integrated with WebMAUS Version 3.8)]. Retrieved from https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface

14. **Al-Tamimi, J.** (2021a). *Introduction to Random Forests* [Open Source Online Material - Introduction to Decision Trees and Random Forests]. Retrieved from https://jalalal-tamimi.github.io/Intro-Random-Forests/

15. **Al-Tamimi, J.** (2021b). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062021.nb.html

16. **Al-Tamimi, J.** (2020). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062020.nb.html

17. **Al-Tamimi, J.**, Heyne, M., & Derrick, D. (2019). *GAMM-Trombone-2019* [Open source online material - Supplementary Material as an introduction to GAMMs (Generalized Additive Mixed-effects Modelling) on UTI (Ultrasound Tongue Imaging) data on trombone vs vowel production]. Retrieved from https://jalalal-tamimi.github.io/GAMM-Trombone-2019/

18. **Al-Tamimi, J.** (2019b). *Introduction to R* [Open Source Online Material as an Introduction R: Introduction, Visualisation]. Retrieved from https://jalalal-tamimi.github.io/R-Introduction-to-R/

19. **Al-Tamimi, J.** (2019c). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062019.nb.html

20. **Al-Tamimi, J.** (2018a). *R-Rating-data* [Open Source Online Material - Supplementary Material as an Introduction to Cumulative Logit Regression]. Retrieved from https://jalalal-tamimi.github.io/R-Rating-data/

21. **Al-Tamimi, J.** (2018b). *R-Voicing-Gemination-VOT* [Open Source Online Material - Supplementary Material for the Statistical Analyses of Mixed-Effects Models and Random Forests]. Retrieved from https://jalalal-tamimi.github.io/R-Voicing-Gemination-VOT/

22. **Al-Tamimi, J.** (2018c). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. Retrieved from https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062018.nb.html

Chapter 3

PhD Diploma

RÉPUBLIQUE FRANÇAISE

Ministère de l'enseignement supérieur et de la recherche

UNIVERSITÉ LYON 2

DOCTORAT

Vu le code de l'éducation, notamment son article L-612-7 ;

Vu le code de la recherche, notamment son article L.412-1 ;

Vu le décret n°2002-481 du 8 avril 2002 relatif aux grades et titres universitaires et aux diplômes nationaux ;

Vu l'arrêté du 3 septembre 1998 relatif à la charte des thèses ;

Vu l'arrêté du 7 août 2006 relatif à la formation doctorale ;

Vu les pièces justificatives produites par M. JALAL-EDDIN AL-TAMIMI, né le 20 août 1976 à KOWEIT (KOWEIT), en vue de son inscription au doctorat ;

Vu le procès-verbal du jury attestant que l'intéressé a soutenu, le 14 septembre 2007, une thèse portant sur le sujet suivant : Indices dynamiques et perception des voyelles : étude translinguistique en arabe dialectal et en français, préparée au sein de l'école doctorale Lettres, langues, linguistique,arts, devant un jury présidé par RENÉ CARRE, Directeur de recherche émérite et composé de THAMI BENKIRANE, Professeur d'université, SALEM GHAZALI, Professeur d'université, JEAN-MARIE HOMBERT, Directeur de recherche, BJÖRN LINDBLOM, Professeur d'université, WILLY SERNICLAES, Chargée de recherche H.D.R.;

Vu la délibération du jury ;

Le **DIPLÔME NATIONAL DE DOCTEUR** SCIENCES DU LANGAGE, *mention très honorable*

est délivré à **M. JALAL-EDDIN AL-TAMIMI**

et confère **le grade de docteur**,

pour en jouir avec les droits et prérogatives qui y sont attachés.

Fait à LYON, le 4 février 2008

UNIVERSITÉ LUMIÈRE LYON 2 — UNIVERSITÉ LUMIÈRE LYON 2 — ACADÉMIE DE LYON

Le titulaire

Le Président

Claude JOURNES

Le Recteur d'Académie,
Chancelier des universités

Académie de Lyon — Rectorat

Roland DEBBASCH

N° LYONII 6593999

/2008200606084

Part II

Research Activities

Chapter 4

Summary of Previous Work

4.1 Introduction

My research falls within the Laboratory Phonology framework, where I look at the role of Fine-Phonetic-Detail (FPD) in various areas of production, perception and learning/acquisition. My main motivation for this type of work is to provide an empirical evidence for FPD and its role in phonological descriptions, but also as part of the mental lexicon, following an Exemplar-based and Usage-based approaches to perception (Foulkes & Docherty, 2006; Hawkins, 2010; Johnson, 2006; McAleer & Belin, 2018; Pierrehumbert, 2001, 2006; Zellou, 2017b). FPD is found in the speech of producers, which allows them to produce specific types of variations to encode their "accent", "traits", and/or "identity"; perceivers use various levels of FPD to identify the units of speech; learners employ specific strategies in their learning journey to reach an ideal native-like and/or target-like production similar to that of native speakers of the target languages.

In many parts of my research, it is clear that speech sounds are organised in such a way to promote the organisation of "speech" as having a major role in the mental lexicon. My research activities summarised below look at various levels. I put forward the role of coarticulation (both local and long-distance), and its impact on how speech sounds are organised in smaller and larger units (from "phonemes" to words). I emphasise on the individuality of producers, perceivers and learners by looking at various types of contrasts and by trying to quantify the FPD. Being an experimentalist with a multidisciplinary approach to my research, I evaluate various phonological units and try to provide an empirical evidence for specific distinctive features. I employ various methods presented in the List 1 below:

1.
   1) Auditory analyses ⇒ Providing a first impression of the data as produced by producers/learners. This is done by emphasising the role of FPD without imposing any *a priori* knowledge on how a system may work
   2) Instrumental acoustic analyses of speech sounds ⇒ From basic to advanced and automated methods in signal processing used in Automatic Speaker/Speech Recognition (ASR)
   3) Instrumental non-invasive articulatory techniques ⇒ Ultrasound Tongue Imaging (UTI) and ElectroGlottoGraphy (EGG), with openings in the future towards employing ElectroMagneticArticulatography (EMA), functional Magnetic Resonance Imaging (fMRI), Electropalatography (EPG), electro PhotoGlottoGraphy (ePGG) and Ultrasound Larynx Imaging
   4) Psycholinguistic approaches:
      a) Using Psychoacoustic approaches, e.g., Auditory representation of speech (e.g., Bark, ERP, Phon/Sone scales); Perceptual Linear Predictive coding (PLP); Correlating acoustics to the Just Noticeable Difference (JND) in speech perception informed by the Signal Detection Theory (SDT)
      b) Perception experiments ⇒ from classical techniques, with identification, discrimination and rating tasks, to openings in the future to using more advanced techniques, such as ElectroEncephaloGraphy (EEG) with Event Related Potentials (ERP), and Eye tracking with Occulometry
   5) Confirmatory Data Analyses ⇒ employing Inferential statistics using a mixed-effects regression modelling framework, which models variations from individual participants, items, and utterances, with various distributions: Gaussian (LMM), Binomial (GLMM), Multinomial (MLR) Cumulative (CLMM); Additive and time series (GAMMs)

6) Exploratory Data Analyses ⇒ employing various machine learning algorithms, from supervised, e.g., Discriminant Analysis, Logistic Regression, Support Vector Machines, Decision Trees, Random Forests, Gradient Boosting Machine, to unsupervised, e.g., clustering techniques, Convolutional Neural Networks (CNN)

One essential component of my research as a whole is the use of multiple approaches to allow for an accurate modelling of the multimodality of speech. Using articulatory to acoustic mapping allows to unravel specific types of constrictions and their acoustic consequences; relating acoustics to perception allows to identify acoustic changes and their impact on perception; comparing results from native speakers and/or adults to L2 learners of a language or infants acquiring their L1 is used to evaluate the role of primary and secondary correlates and how one enhances the other to derive phonological categories (following Keyser & Stevens, 2006; Stevens & Keyser, 1989, 2010). This allows to explain the specific strategies employed by learners/acquirers and how different they are from native speakers/adults. Relying heavily on a psychoacoustic approach allows me to evaluate the perceptual saliency of acoustic results. This allows a better and a more accurate understanding of the impact of the vocal tract on speech production, which allows to empirically evaluate the relationship between articulatory movements, acoustic consequences and perceptual saliency (Following the Quantal Theory of Speech Production, cf. Stevens, 1989).

Finally, my research examines the role of FPD in various types of speakers: Adults, including males and females, Infants acquiring their L1, Native speakers of a language (L1) and learners (L2); Types of material: from controlled lab-speech, read speech, image description to spontaneous speech; On various languages: Dialectal Arabic (Jordanian, Lebanese, Libyan, Moroccan, Algerian, Bahraini, Saudi, Egyptian, etc.), French (standard and varieties); English (British, Australian, New-Zealand); Tongan; Thai, Mandarin Chinese, Malayalam, Finnish, Japanese, Spanish/Catalan, etc... This combination of types of speakers, speech material and languages on which I worked provides a rich environment to assess within and between speaker, dialectal and language variation, which allows to understand the specificity and universality of phonological systems.

In the next sections, a summary of my main research activities will be presented. As will be highlighted below, my research will be divided into seven main activities. We first look at the role of the "Dynamic Specification Approach" of Vowels in Arabic (Section 4.2), by specifically examining the links between production and perception of vowels within-language (Section 4.2.1) and the role of dynamic specification to explaining patterns found in Arabic and French (Section 4.2.2). We continue next by discussing some extensions to the "Dynamic Specification" approaches proposed in the literature as part of the work completed during my PhD supervision, discussing the main findings from various angles (Section 4.2.3).

We move next to the first major part of my research activities, identifying how gemination is phonetically implemented in Lebanese Arabic, starting with how Adults produce it in terms of a contrast combining length and strengthening (Section 4.3). We specifically look at how the temporal domain is better explained by employing a moraic account (Section 4.3.2), trying then to identify whether the contrast is based solely on temporal information and how non-temporal correlates can explain subtle differences in how geminates are realised (Section 4.3.3). We follow this by looking at how infants' acquiring Arabic as their mother tongue produce this contrast (Section 4.3.4) and how they seem to implement a tertiary-type of contrast, with singleton, and geminates at the extremes and a "fortis" or "strong" category, in-between. We take the results from both adults and children as a empirical evidence for the role played by the feature [+Tense] in Arabic phonology.

As an extension to the work done on gemination, we look at implementation of the Voicing contrast in Arabic, starting with the interaction between gemination and voicing by positing the presence of a possible four-way contrast in Arabic. Here, we show how the marked category (i.e., voiced geminates) behaves differently from the other three (i.e., voiceless singleton, voiced singleton, voiceless geminate), where the feature specification emphasises the presence of both [+long] and [+tense]; but with gradation (Section 4.4). We also look at extensions to the voicing profile in Arabic (Section 4.4.3), by examining how voicing is phonologically implemented in Saudi Arabic that shows on the one hand that in Najdi Arabic, the two features [+Voice] and [+Spread Glottis] are active in stops (Section 4.4.3.1) and on the other hand, that in Jazani Arabic, the two features [+Voice] and [+Tense] are active in both fricatives and stops (Section 4.4.3.2).

Then we move to the second major part of my activities, by looking at a special class of sounds in Arabic: "guttural" consonants. Guttural consonants are usually classified as a natural class and they involve a constriction (primary or secondary) throughout the pharyngeal cavity. They are composed of uvulars, and pharyngeals, in addition to the secondary articulated pharyngealised coronals. We first examine the acoustic correlates of pharyngealised coronal plosives in two Arabic dialects, namely Jordanian and Moroccan Arabic (Section 4.5.2), by evaluating the phonetic implementation of the feature [+CET] (for "Constricted Epilaryngeal Tube") following the predictions of the "Laryngeal Articulator Model" (LAM, Esling, 2005; Esling, Moisik, Benner, & Crevier-Buchman, 2019). The results reported in this section claim that the contrast is based not only on spectral changes in the filter, i.e., retraction of the tongue into the pharynx, but also in the source, through voice quality changes associated with a tense voice quality. Following from that, we look at quantifying the (dis-)similarities between "gutturals" using non-invasive techniques, such as Ultrasound Tongue Imaging (Section 4.5.3.2), Electroglottography (Section 4.5.3.3) and combined with acoustic analyses of the source and filter (Section 4.5.3.4). The combination of articulatory and acoustic examination allows for a fine-tuned description of the contrast and emphasises the major role of FPD in feature specification. Then in Section 4.6, we look at two case studies where automated methods in signal processing, either informed by the theory (e.g., prosodic and voice quality profiles) or by employing advanced signal processing algorithms (e.g., openSMILE) can be used to detect vocal changes associated with language disorders. We look first at the case of schizophrenia (Section 4.6.1), before moving to the case of Dementia identified in individuals with a diagnostic of Alzheimer disease (AD) and its precursors in the form of Subjective Cognitive Decline (SCD) or Mild Cognitive Impairment (MCI; Section 4.6.2). We then look at a relatively recent development of a romanisation system in and a forced-alignment system in Arabic developed in collaboration with "The Munich Automatic Segmentation System (MAUS)" team, which led to the development of a free to use, open source system that provides either a time-alignmed segmentation of pretranscribed speech or a text-blind Speech-to-Text transcription (Section 4.7). We end the presentation by looking at various studies identified as "Collaborative Work", which puts forward my expertise in various levels (Section 4.8), before ending with the open source material I developed and made available to the research community (Section 4.9).

4.2 Dynamic Specification of Vowels

Monophthongal vowels are usually described in terms of their first two formant frequencies obtained at the most stable area (around the midpoint). This is identified as the canonical form named the "Simple Target" approach (see, Joos, 1948; Peterson & Barney, 1952, among others). It is a unifying concept that identifies the "Vowel Target" as the most stable configuration in vocal tract, from an articulatory point of view, which is represented by F1 & F2 from single static spectral slice, from acoustic point of view, which implies that F1 & F2 are usually enough for Vowel Target identification (see, Strange, 1989b, for a summary). This seems to be exemplified in more detail by the use of the effective formant, F', rather than the absolute formants in vowel perception, where formant peaks that are within 3.5 Bark are spectrally integrated (Chistovich & Lublinskaya, 1979; Vaissière, 2011).

Due to various issues related to the use of the "Simple Target" approach, researchers were primarily interested in the use of more "elaborated" approaches to resolve the issues related to the normalisation" problem. It is well-known that in perceiving speech sounds, listeners try to remove any intra and inter-speaker variation when trying to identify speech sounds (Johnson, 1997; Johnson & Sjerps, 2021), though it is clear that they still retain some specific details about the speaker producing the items, following an exemplar-based approach to speech perception (Johnson & Sjerps, 2021; Pierrehumbert, 2001, 2006, 2016). This led to the development of the "Elaborated Target" approach, which postulates that vowels are best identified when looking at the distance between formant frequencies via the Bark-Distance approach (Syrdal & Gopal, 1986) or when using formant-ratios rather than absolute formant or formant distances (Miller, 1989). As will be highlighted later on (Section 4.5.2.3, pp. 97), the Bark-Distance approach has been used to correlate specific acoustic patterns to phonological features, such as [±High], or [±Back].

Both approaches have a major drawback. First, the "Simple Target" approach, which relies heavily on the formant values obtained at a single point within the total duration of the vowel, should be regarded as flawed and can be considered as a laboratory artefact (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). It is well-known that vowels are usually co-produced with consonants rather than being produced on their own, leading to a mixture of cues used by listeners to identify vowels categories. Formant frequencies are highly unstable, due to inter & intra-speaker variability among speakers of the same or different physiological sex and age groups (Peterson & Barney, 1952). Finally, when vowels are produced in a sequence, their quality is distributed during the whole voiced portion. Due to these issues, it is clear that when exposed to vowels sliced from real speech (whether isolated items, or from connected speech), listeners will be using different acoustic cues to identify coarticulated vowels. These same issues apply to the "Elaborated Target" approach, which although tries to account for more details in comparison to the "Simple Target" approach, still suffers from the fact that formant frequencies are still obtained at a single point within the total duration of the vowel! Both approaches fall within what is termed in the literature as the "Static" approaches to monophthongal vowel production and perception (Strange, 1989b).

This led researchers to look at vowels from a more dynamic way, which led to the development of a new approach to vowel production and perception, named "Dynamic Specification of Vowels". This dynamic approach puts forward the importance of going beyond the static approaches to monophthongal vowel production and perception, by examining either intrinsic or extrinsic vowel dynamics (Strange, 1989b). The former was termed as the "Vowel Inherent Spectral Changes" (VISC: Hillenbrand, 2013; Morrison & Assmann, 2013) quantifying vowel's own trajectories after ignoring any coarticulatory effects of surrounding consonants; while the latter refers to the role of formant transitions and the portions at the onset/offset of the vowel for a more accurate identification of coarticulated vowels (Strange, 1989a; Strange & Jenkins, 2013; Strange, Jenkins, & Johnson, 1983; Strange, Verbrugge, Shankweiler, & Edman, 1976). A major claim advanced from the proponents of the extrinsic dynamic specification approach is that onsets and offsets of a syllable, without the central part, but in addition to the temporal information provide more detailed cues to the listener to allow for an accurate identification of the vowel even when the central parts were excised (for more details on the silent-centre approach, see Strange & Jenkins, 2013, pp. 93-96).

It is important to note that the majority of research conducted within the intrinsic or extrinsic dynamic specification approaches looks at monophthongal vowels rather than diphthongs, which are dynamic by nature due to them generally containing two stable areas with on-glides/off-glides depending on how they are realised in various languages

In the following sections, a summary of my research activity in this theme will focus on within and between language variation in production and perception, before ending with the current results from dynamic specification of vowels as produced by Arabic learners of English as a foreign language.

4.2.1 From "Static" approaches...

My research activities within this section relied heavily on the "static" approaches to vowel production (J. Al-Tamimi, 2001) and vowel production vs perception (J. Al-Tamimi, 2002; J. Al-Tamimi & Barkat-Defradas, 2002; J. Al-Tamimi, Girard, & Marsico, 2002; Barkat-Defradas, Al-Tamimi, & Benkirane, 2003).

4.2.1.1 Vowel production

First, my dissertation J. Al-Tamimi (2001), under the supervision of Prof. Lolke Van Der Veen, focused on investigating the interferences between vowel production in Modern Standard Arabic (MSA) and Jordanian Arabic, looking specifically at within- and between-speaker variation. As is the case in all Arabic countries, a speaker is generally competent in two registers, a "formal" (MSA) and a "less formal" form, represented by Jordanian Arabic. This diglossic situation leads in some instances to bilingual competency in the two forms, even though MSA is a learnt register mainly from school age, whereas the dialectal form is considered as the mother tongue of the speaker (Ennaji, 1991; Ferguson, 1959; Suleiman, 1985). When looking at the case at hand here, and comparing the productions of vowels in both settings, as produced by 6 male speakers from various regions in Jordan, it was evident that there was a systematic difference between the vowels produced in MSA and in Jordanian Arabic, on the one hand and between speakers in their dialectal form, on the other. Speakers were producing slightly more centralised vowels in Jordanian Arabic in comparison to their own productions in MSA, which seemed to be more peripheral. This seems to correlate well with an extreme type of production in the more "formal" form (MSA), in comparison to a more "reduced" form in the dialect. In this situation, it is possible to correlate the differences between the two registers in the light of the *H&H* Theory (Lindblom, 1990). Between-speaker variation was present, but did not impact on the location of vowels in the acoustic vowel space. Crucially, this study provided an empirical evidence for an 8 monophthongal vowel system in Jordanian Arabic (/iː i eː a aː oː u uː/), in comparison with a 6 monophthongal vowel system in MSA (/iː i a aː u uː/) in addition to two diphthongs (/aj aw/). The results of this first study put emphasis on clear differences between the two forms and most importantly proposed a revision to the descriptions of both the consonantal and vocalic systems of dialectal Arabic. This motivated my subsequent studies due to the scarcity of research in dialectal forms in Arabic, as it puts emphasis on the role of individual variation and differences in speech styles between a more formal (i.e., MSA) to a more colloquial form represented by Jordanian Arabic.

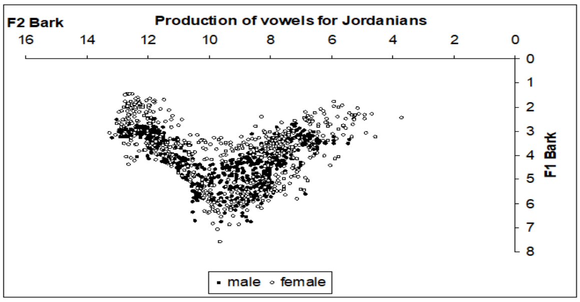
4.2.1.2 Vowel production and perception

The specific differences in vowel realisation identified in (J. Al-Tamimi, 2001) highlighted within- and between-speaker variation in the production of Jordanian Arabic male speakers. As a follow up and expansion of the study, I integrated the research project: "Cognitique (P3) : Variabilité phonétique en production et perception de parole : rôle et limites des stratégies individuelles"; PI René Carré (funded between 2001 and 2003). The project looked at a cross-linguistic comparison between French, Italian and Arabic (represented by Jordanian and Moroccan Arabic).

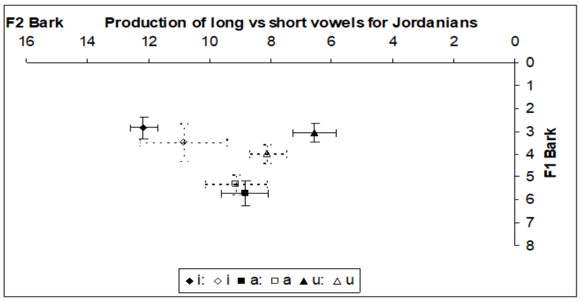
The research conducted during my master, under the supervision of Prof. Jean-Marie Hombert, focused on the vowel production and perception in Jordanian Arabic (in comparison with French), in male and female participants; aiming at quantifying individual strategies, impact of the physiological sex, differences between short and long vowels and links between production and perception of vowels in Jordanian Arabic from 16 participants; 8 males and 8 females, in comparison with French (see J. Al-Tamimi, 2002; which was subsequently published in J. Al-Tamimi, Girard, & Marsico, 2002 on Arabic vs French, 24^{th}JEP; J. Al-Tamimi & Barkat-Defradas, 2002, on Jordanian vs Moroccan Arabic, 5^{th}AIDA Proceedings; Barkat-Defradas, Al-Tamimi, & Benkirane, 2003, on Jordanian vs Moroccan Arabic, 15^{th}ICPhS)

Starting with the production experiment (see Figure 4.1), participants were asked to produce the vowels of their language as produced in words, syllables and in isolation. They were exposed to real words with one of the 8-monophthongal vowels (/iː i eː a aː oː u uː/) in Jordanian Arabic preceded by variable consonantal environments /b d tˤ dˤ k q sˤ w/. After obtaining formant frequencies at the temporal midpoint and applying a normalising technique to reduce physiological sex differences (following Henton, 1989, 1995a, 1995b), there were still clear male/female differences emerging in the data. Female subjects were producing a more peripheral vowels overall than male speakers: Front vowels are more front and more closed; back vowels are more back and more close, and open vowels are produced as more open (Figure 4.1a). These results correlated with classical findings reported by (Henton, 1995a; Lee, Potamianos, & Narayanan, 1999; Peterson & Barney, 1952) that shows clear differences between male and female subjects, with and without normalisation of vowel spaces. Between-subject variation was also evident which led to variable vowel dispersion, which is expected due to individual strategies in producing vowels and impact of surrounding environments. However, speakers were systematic in their production and there were no major shifts in the vowels qualities within and across males/females. Figure 4.1b shows the results differentiating between short and long vowels in Jordanian Arabic. The vowel spaces, represented by the means and SDs, suggest that the short /i a u/ are best described in terms of quantitative and qualitative differences, as these were realised as lax vowels [ɪ ɐ ʊ] by both males and females, in comparison with their tense counterparts /iː eː aː

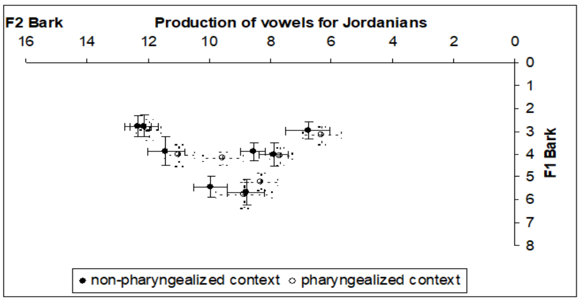
oː uː/. Figure 4.1c shows the impact of the pharyngealised context on the production of vowels; they show an increase in their F1 and a decrease in F2, associated with the features [+open] and [+back], confirming results from previous literature[1]. These results highlighted the systematicity of vowel production by all participants, regardless of their physiological sex in differentiating between the short and long vowels, in addition to the impact of pharyngealisation on the vowels.



(a) Male vs Female



(b) Short vs Long vowels



(c) Non-Pharyngealised vs Pharyngealised

Figure 4.1: Vowel space from the production task in (a) Male vs Female and (b) Short vs Long vowels (reproduced from J. Al-Tamimi, 2002, pp. 32 and 34-35)

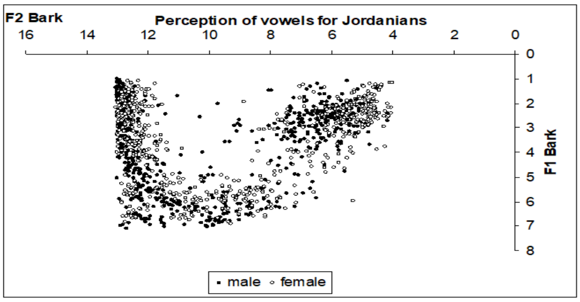
The second part of the study looked at how the same 16 participants identified the vowels of their system in a perception experiment as we wanted to evaluate whether they will be identifying the short vs long vowels and impact of pharyngealisation in a similar fashion to that in the production task. We implemented a perception experiment (designed by François Pellegrino) based on a modified version of the "Method of Adjustment" (Johnson, Flemming, & Wright, 1993). The listener was asked to identify the best prototypical isolated vowel as found in a reference word in either /d/ or /dˤ/ context, e.g., identify the vowel /iː/ as found in the word /diːn/ "religion" in Jordanian Arabic. The listener moved the cursor (controlled by the mouse) searching for the "best" prototypical vowel in an F1*F2 vowel space[2] with two *f*o values at 120Hz and 240Hz [3]. The best prototypical forms chosen by the listeners are shown in Figure 4.2. It is clear there are some differences emerging between males and females in their choices with an increase in vowel space size in females (p<0.05), but most the results pointed to "similar" choices across vowels (Figure 4.2a). In terms of the differences between short and long vowels (Figure 4.2b), short vowels were identified as more centralised than the long vowels; confirming the results from production. Finally, vowels identified in a pharyngealised context were only

[1]See Section 4.5.2 for a more detailed account

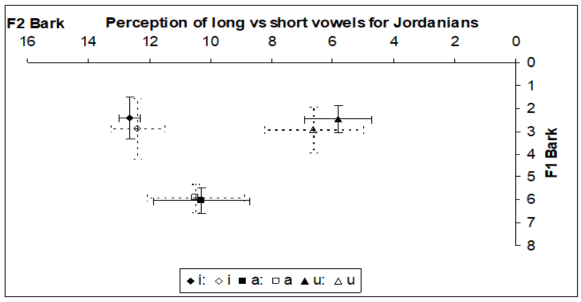
[2]F3 was controlled to be that of the average for each vowel in a /d dˤ/ context.

[3]These were used as canonical forms for male vs female pitch values to evaluate potential impact on preferred vowels. The results showed no clear patterns nor impact on choices

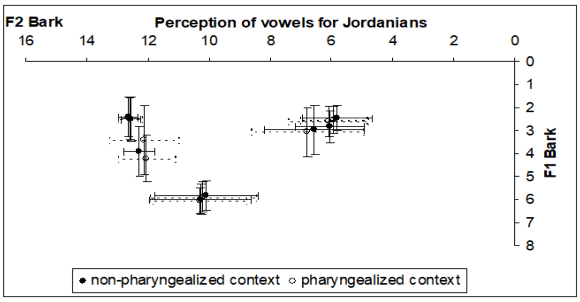
different on the F1 dimension, with a marginal change on F2 (Figure 4.2c); confirming partially the results from the production experiment.



(a) Male vs Female



(b) Short vs Long vowels

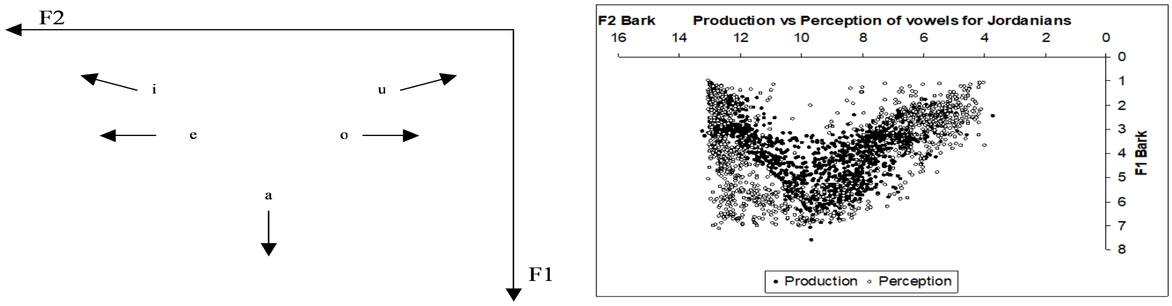


(c) Non-Pharyngealised vs Pharyngealised

Figure 4.2: Vowel space from the perception task in (a) Male vs Female and (b) Short vs Long vowels (reproduced from J. Al-Tamimi, 2002, pp. 32 and 40)

Finally, a final aim of the study was to assess whether vowel spaces in production and perception were linked to each other and whether there is a "hyperspace" effect (Johnson et al., 1993). The Figure 4.3a shows the predictions according to the "hyperspace" effect; vowel spaces in perception will be maximised as listeners tend to identify vowels as extreme forms to those observed in production for the same task. Ours results presented in Figure 4.3b show an expansion of the vowel space in the perception; confirming that listeners prefer extreme vowels by trying to maximise the differences between each category (Barrett, 1998; P. Kuhl, 1991) going for a form similar to that observed in hyper-articulated speech (Lindblom, 1990), even though (Johnson et al., 1993, Figure 5, pp. 520) shows that perceptual vowel spaces are actually more extreme than those obtained from hyper-articulated forms.

One interesting and unexpected finding from the results of this study was the fact that both Jordanian and Moroccan Arabic participants found the task of producing and identifying isolated vowels a difficult task; aspect not found within French nor Italian participants. In production, nearly all isolated vowels were produced in a CV sequence, where C = /ʔ/. In perception, they reported hearing either "something that was missing" or a "beep" with variable qualities, i.e., an /i/-beep, an /o/-beep, etc. We suspected at the time that structural differences between Arabic and French/Italian are possibly the reason behind such a difference; aspect that motivated my PhD research.



(a) Prediction (b) Results

Figure 4.3: Vowel space from production vs perception tasks (reproduced from J. Al-Tamimi, 2002, pp. 50-60)

4.2.2 To "Dynamic" approaches

As highlighted in the introduction to this section and following the results presented in J. Al-Tamimi (2002), I worked on the role of dynamic specification of vowels in Arabic in comparison with those in French. This work formed part of my PhD thesis, under the supervision of Prof. Jean-Marie Hombert and Prof. René Carré (see J. Al-Tamimi, 2007a; which was subsequently published as: J. Al-Tamimi, 2004 on Locus equations and their use to quantify coarticulatory patterns, 25^{th}JEP; J. Al-Tamimi & Ferragne, 2005, on comparing impact of vowel density on the dispersion of the vowel space, 9^{th}Interspeech; J. Al-Tamimi, 2006, on the use of formant slopes to quantify degree of reduction, 26^{th}JEP; J. Al-Tamimi, 2007b, on comparing static vs dynamic cues in production of vowel in Arabic, 16^{th}ICPhS; J. Al-Tamimi, 2008, on impact of vowel density on perceptual vowel spaces, 27^{th}JEP).

My PhD had two main aims. First, I wanted to examine the role of static (i.e., vowel midpoint) vs dynamic approaches (transitional parts in the form of "Formant Slopes") and their impact on discrimination of vowels in production and perception, in a cross-linguistic view point, comparing Arabic to French. Second, the cross-linguistic comparison aimed at quantifying structural and morphophonological differences between the two system, on the one hand, and on quantifying impact of vowel density on the dispersion of vowel spaces, on the other. Three systems were examined: Moroccan Arabic (with a 5-vowel system: /iː aː ə ʊ uː/), Jordanian Arabic (with an 8-vowel system: /i iː eː a aː oː u uː/) and French (with an 11-vowel system: /i e ɛ a ɑ ɔ o u y ø œ/).

The work was motivated by structural differences between the two languages. Arabic is known to be a non-concatenative language, where the consonantal templatic structure impacts heavily on word derivation, while vowels are used as a grammatical tool to guide deriving grammatical functions (Versteegh, 2001). French on the other hand is concatenative language with an inflexional morphology, where the vowel has a major role in word derivation (Haspelmath & Sims, 2010). This leads to a difference in the vowel status and use. In French, a vowel can form a word and is not only used for grammatical derivation; whereas in Arabic, a vowel cannot form a word on its own and is always associated to a sequence of consonants and/or preceded by a glottal stop, leading to the vowel /aː/ to be produced as an [ʔaː]; this fact led to questioning the status of this sequence being either phonetic or phonological. Hence, we formulated the hypothesis that in Arabic, an isolated vowel is an artefact due to the fact that it carries no meaning for a listener, whereas it is part of the grammar in French. This may lead to clear differences between the two languages in terms of the mental representation of the vowel. Our claim then is that if the two languages show differences in terms of their morphological structures, we cannot expect Arabic listeners to be able to easily identify isolated vowels; a more dynamic representation is required to account for that.

For the second aim, we wanted to evaluate the role of vowel density and its impact on the dispersion of vowel spaces. it is well-known in the literature that languages with large inventory sizes will expand their vowel spaces to maximise the distinctivity between vowels, whereas those with reduced vowel spaces will have more freedom to either reduce the vowel space and/or to decrease the distance between vowels (Engstrand & Krull, 1991; Johnson, 2000; Livijn, 2000; Manuel, 1990; Meunier, Frenck-Mestre, Lelekov-Boissard, & Besnerais, 2003; Mok, 2013). The expectation then is that French, with a larger vowel space, will show an increase in vowel space size with a decrease in distance between vowels to retain distinctivity, whereas Arabic, will show the reverse. We would expect Jordanian Arabic to behave differently to Moroccan Arabic due to differences in number of vowels between the two systems. We aimed then to examine if the density size impacts on both vowels produced by speakers of the three systems and in identification of the best prototypes.

During my PhD, I joined two research projects. The first was: "Structuration et Dynamiques des Systèmes Phonético-Phonologiques", led by François Pellegrino (between 2003 and 2005). The project aimed at evaluating the dynamics of vowels and consonants in various languages and my role was to evaluate the dynamic representation of vowels by evaluating the role of coarticulation in production and perception. The second project was "Complexité, Langage et Langues – CL^2", led by François Pellegrino (between 2005 and 2008), which was interested in the complexity of systems, from a phonetic-phonological, morphosyntactic and acquisition. I evaluated the role of structural differences between languages and the interface between phonetic-phonological and morpho-syntactic representation. We look below at the main findings from production followed by perception.

4.2.2.1 Vowel Production

The first part of the work was to look for an empirical evidence for the dynamic specification of vowels and identify if and whether structural differences between languages impact on how the vowels are produced. We reused some of the production data obtained from the research project: "Cognitique (P3) : Variabilité phonétique en production et perception de parole : rôle et limites des stratégies individuelles"; PI René Carré (funded between 2001 and 2003). We also recorded data from missing participants (for the three systems). A total of 30 male speakers (10 for each system) were recruited. We decided not to use female speakers' recordings as we did not have any a priori hypothesis regarding gender-specific variation (beyond physiological sex differences) in the implementation of dynamic specification of vowels.

The data from all participants came in the form of vowels produced in words, in the syllable and in isolation, in three places of articulation (bilabial, alveolar and velar, in additional to alveolar pharyngealised in Arabic). As an example, speakers were asked to produce the vowel /iː/ as produced in the word /diːnak/ ("your religion" in Jordanian Arabic). They were instructed to produce it as follows: /diːnak - diː - iː/. The aim of this structure was to obtain as similar as possible the production of the vowel in the three forms.

Given that our aim was to compare static and dynamic specification of vowels following previous literature, we decided to use the vowel midpoint as a reflex of the static approach and to devise a new metric for the dynamic approach. To quantify the dynamic nature of vowel representation, I borrowed notions from the Locus Equations (LE). LE are widely used in the literature as a potential invariant index to both place of articulation and to coarticulation between consonants and vowels (see Lindblom & Sussman, 2012; Sussman, Hoemeke, & Ahmed, 1993; Sussman, Hoemeke, & McCaffrey, 1992; Sussman, McCaffrey, & Matthews, 1991; Yeou, 1997, among others). LE quantifies the relationship between a vowel's F2 value obtained at the onset and at the midpoint; a relationship quantified via a linear regression (see Equation 4.1). This leads to identifying a slope *a*, which can be rising, falling or flat, and an intercept *b*, representing the point where the regression line crosses the y-axis (Lindblom, 1963). Using these two metrics allows researchers to differentiate between various places of articulation and relating these to coarticulatory changes and resistance (see Lindblom & Sussman, 2012; Yeou, 1997, for an excellent summary).

$$F2_{Onset} = a * F2_{mid} + b \tag{4.1}$$

In our case, and because we were interested in quantifying the dynamic changes on each of the first three formant frequencies as a function of time, we modified the Equation 4.1 by integrating the time component as can be seen in Equation 4.2, where formant frequency 1, 2 or 3, computed each 5 ms from the onset to the midpoint of the vowel is modelled via a linear regression as a function of the time component. We used two versions of the time component, either as normalised to 0 and 0.5 or as raw time, first to assess the shape and direction of the formant slope regardless of time, in the former, and to evaluate strength of change that depends of vowel duration, in the latter.

$$F_{Onset\_Mid} = a * Time + b \tag{4.2}$$

As can be seen from Figure 4.4, a linear regression provides an appropriate fit to the F1 and F3 trajectories; for F2, a linear regression does not fit the data in an accurate manner. Hence, we also applied a 2^{nd} and 2^{rd} order polynomial regressions that extend the original formula (see Equations 4.3 and 4.4, respectively).

$$F_{Onset\_Mid} = a_0 + a_1 * Time + a_2 * Time^2 \tag{4.3}$$

$$F_{Onset\_Mid} = a_0 + a_1 * Time + a_2 * Time^2 + a_3 * Time^3 \tag{4.4}$$

Results obtained from this approach were first quantified via inferential statistics followed by Linear Discriminant Function Analysis (LDA) to show when and which approach is successful at separating vowels within and between the three systems. Looking at within-language results, the LDA showed that dynamic specification via the "Formant Slope" approach allowed for an average increase in classification accuracy by about 10-11% over static approaches when formant frequencies on their own or in association to duration were included (J. Al-Tamimi, 2007a, pp. 186). This dynamic specification puts forward the role of the onset and of formant transitions as a whole in vowel recognition and follows the suggestions in the literature that extrinsic dynamic specification of vowels, in the form of onsets (and offsets) are enough to allow for an accurate identification of vowels (Strange, 1989b).

Between-language comparison showed that vowel density impacted the size and dispersion of vowel spaces (J. Al-Tamimi, 2006; J. Al-Tamimi & Ferragne, 2005). Our results presented in

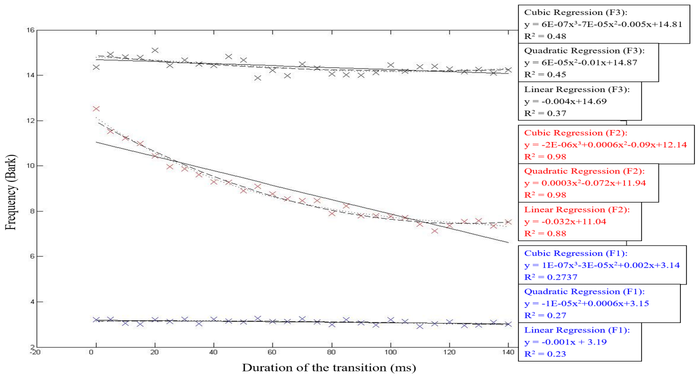


Figure 4.4: Formant Slopes quantified via linear (continuous lines), 2^{nd} Order (discontinuous lines) and 3^{rd} Order (dotted lines) polynomials for F1, F2 and F3 of the vowel /u:/ as produced in the word /du:d/ (worms, in Jordanian Arabic) as produced by speaker 09MAJ, with equations and R^2 (J. Al-Tamimi, 2007a, pp. 162)

Figure 4.5 show the connected vowel spaces represented by the point vowels /i a u/ (long vowels in Arabic) in the three systems. Vowel spaces in Moroccan Arabic (with 5 vowels) are smaller and more centralised than those seen in French (with 11 vowels), which are more peripheral; the vowel space in Jordanian Arabic is in-between. Looking at the vowel space area (computed via the convex hull method), it is clear that Moroccan Arabic shows the smallest size at 7.11 Barks^2, Jordanian Arabic at 8.14 Bark^2 and French at 8.95 Bark^2. These results are conform with the predictions from the literature and follow the predictions of the "Adapted Dispersion Theory" (Liljencrants & Lindblom, 1972; Lindblom, 1986) and of the "Dispersion-Focalisation Theory" (J.-L. Schwartz, Boë, Vallée, & Abry, 1997).

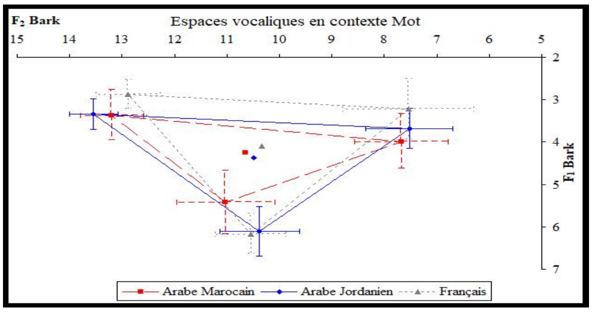
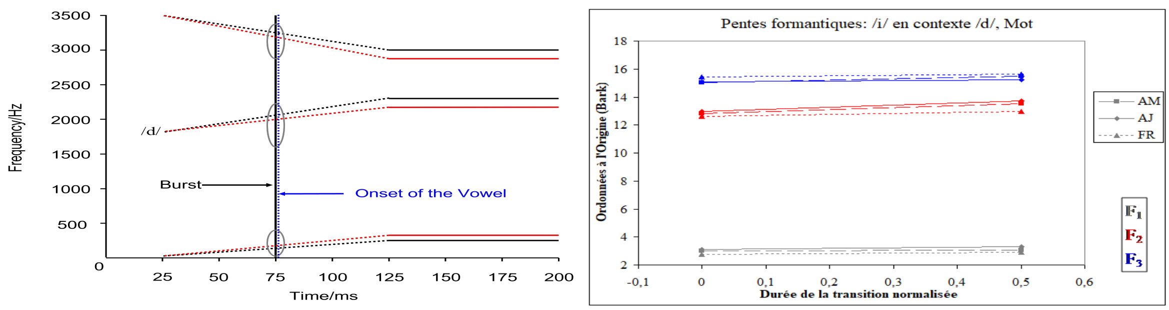


Figure 4.5: Between-Language vowel spaces for the vowels /i a u/ (long vowels in Arabic) in Jordanian Arabic, Moroccan Arabic and in French with the centre of gravity of the vowel space (J. Al-Tamimi, 2007a, pp. 132)

In terms of dynamic specification, and as can be seen from Figure 4.6, the predicted change in terms of vowel reduction can be quantified via changes observed directly at the onset and midpoint values of the three formants. Figure 4.6a shows the predictions, whereby a partially reduced vowel (in red) will have its formant frequencies changing in comparison to a non-reduced vowel (in black), with an increase to F2 and a decrease to both F2 and F3 seen from the onset; the curves in blue show the consequences of an extreme vowel reduction into a /ə/ (see J. Al-Tamimi, 2006, for an additonal account). The results presented in Figure 4.6b confirm this change. Overall, French shows the least reduction on F1 and F3, and variable change on F2, suggesting that the vowel /i/ in French is influenced the most by the alveolar context. Moroccan and Jordanian Arabic seem to pattern in a similar fashion, with the latter showing an increase in vowel reduction. It should be noted that the results do not match the predictions to 100%. This can be explained by a difference either in place of articulation of /d/ (alveolar vs dental) or by how the vowel /i/ is produced across the three systems (pre-palatal in French, possibly mid-palatal in Arabic). This interaction is expected to impact on coarticulatory patterns in the three systems. It is interesting however to note that we already see change across the three systems at the very beginning of the vowel at its onset; this is likely to be a cue to allow for a more accurate discrimination between languages: if at the onset of the vowel (/i/ in our case), listeners are already able to identify language-specific patterns that will allow them to 1) accurately identify vowels as belonging to a specific language and 2) in different contexts[4].



(a) Predictions (b) Results

Figure 4.6: Between-Language Formant Slopes for the vowel /i/ (long vowels in Arabic) in an alveolar context represented via linear regressions; (a) predicted and (b) true results (reproduced from J. Al-Tamimi, 2007a, pp. 164 and 177)

4.2.2.2 Vowel Perception

Given the dynamic patterns identified in the production of vowels across the three systems, we wanted next to assess the contribution of this dynamic specification approach to vowel perception (J. Al-Tamimi, 2007a, Chapter 5). 30 new male listeners who did not participate in the first study were recruited (10 per system). They all shared similar demographics to the 30 speakers in the production task, except that all participants were monolinguals with no knowledge of any other language, to avoid impact on the status of the vowel especially for Arabic listeners. We designed two perception experiments. The first was a reimplementation of the Method of Adjustment prototype identification task that was initially proposed by (Johnson et al., 1993) and was designed by René Carré. The second was a vowel categorisation task similar to that originally used by (Hombert & Puech, 1984) and designed by François Pellegrino.

In both experiments our aim was to compare the performance of listeners in identifying the prototypes and in categorising the vowel spaces using static and dynamic specification. For the prototype identification task, the stimuli were composed of either a steady state vowels with variable length (100ms for short and 200ms for long/plain vowels) representing the "static" approach (see J. Al-Tamimi, 2007a, pp. 307, for the reference values of first 4 formants and Bandwidth). For the dynamically specified vowels, they were composed of the same steady state vowels preceded by a transitional part of variable lengths (between 15-25 ms) and of murmur (variable between 15-25ms) reflecting variable transitions related to the bilabial, alveolar and velar places of articulations (see J. Al-Tamimi, 2007a, pp. 309, for the reference values for the virtual

[4]See appendices in J. Al-Tamimi (2007a), which show the patterns across vowels and contexts

locus values for the first 4 formants)[5].

For the categorisation task, a set of 56 stimuli were resynthesised using the software: SyntFormVoy developed by René Carré. These had an f_0 pattern of raising-falling (values 120-132-96 Hz), F1 frequencies varied by 100 Hz step between 250 and 750 Hz and F2 by a 200Hz step, between 650 and 2350 Hz and two F3 values: 2300 and 2700 Hz reflecting rounded and unrounded vowels (see J. Al-Tamimi, 2007a, pp. 311, for the reference values for the virtual locus values for the first 4 formants for each stimuli). The dynamic stimuli were composed of the same static vowels preceded by a 25ms with a transition reflecting an alveolar place of articulation.

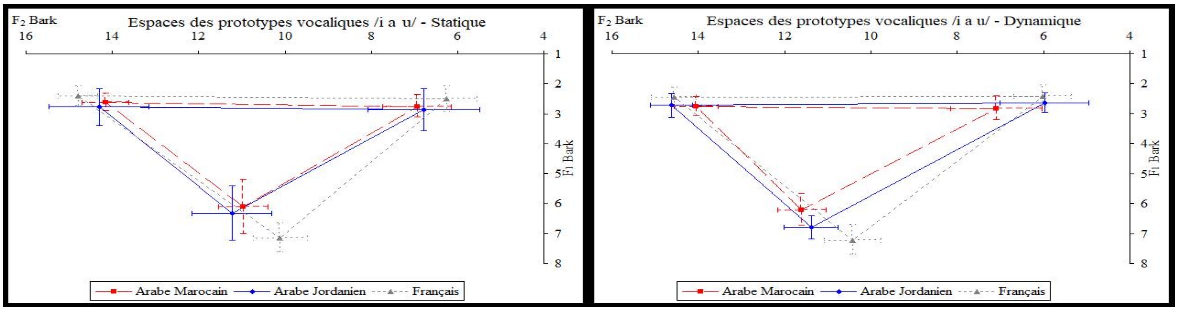
In both cases, we termed the static stimuli as V-Target and the dynamic stimuli as CV-Targets to reflect how the stimuli were generated and what the expectations are in terms of potential differences in listeners behaviour. We used real words for each system and asked participants to identify the best prototypical vowel as would be identified within that word, e.g., identify the best prototypical vowel for /iː/ as in the word "/diːn/" ("religion") in Jordanian Arabic. For each system, we chose various words that all started with one of the three contexts /b d k/ and was followed by one of the vowels, in a stressed syllable position. For Moroccan Arabic, we decided to use six vowels instead of the five reported in the literature due to the fact that the vowel */i u/ from classical Arabic merged into the vowel /ə/ in the modern variety and we wanted to assess whether listeners will still be able to distinguish the two vowels from a perception point of view[6]. For Jordanian Arabic and French, we retained the two systems used in the production task, i.e., /i iː eː a aː oː u uː/) for the former and /i e ɛ a ɑ ɔ o u y ø œ/ for the latter. Finally, because we used two types of experiments (Prototype and Categorisation) with two types of stimuli (V-Target and CV-Target), it was essential to have as much control as possible to the order of presentation. Hence we used a split-plot design, with a complete counterbalancing: half of the listeners in each language started with the V-Target and the other half with the CV-Target, **and** the one half started with the Prototype experiment and the other with the Categorisation experiment. These were important to remove any impact of training and carryover effects that can lead listeners exposed first to the CV-Target to find it "easier" to identify the target vowels using the V-Target stimuli, and vice versa.

Starting with the Prototype identification task and looking first at the between-language comparison, the results presented in Figure 4.7 shows the average and SD of the three prototypical point vowels /i a u (long vowels in Arabic) as identified by the listeners of the three systems, in the V-target (Figure 4.7a) and the CV-Target (Figure 4.7b). It is clear from the results that vowel spaces in the three systems were different in dispersion and in their sizes. In the V-Target (Figure 4.7a), vowel spaces in Moroccan Arabic were the smallest (at 12.49 $Bark^2$) than in Jordanian Arabic (at 13.79 $Bark^2$) or in French (at 20.12 $Bark^2$). This change in the vowel space size depends on the vowel density and is correlated with the findings from production that vowel spaces are impacted by the vowel density size (see J. Al-Tamimi, 2008). When looking at the CV-Target (Figure 4.7b), it is that vowel spaces and positions are different between-languages and V-Targets and CV-Targets. This possibly highlights the fact that listeners were able to identify both prototypes as being different from each other. In the CV-Target again, vowel spaces seem to have been impacted by the vowel density: Moroccan Arabic showing the smallest vowel space (at

[5] We used the two software: PercepCV et PlagePercepCV developed by René Carré to obtain the F2 and F3 formant values for the virtual locus, following the Klatt and Klatt (1990)'s approach

[6] In the production results presented in J. Al-Tamimi (2007a, Figures 69-71, pp. 215-217), Moroccan Arabic speakers were variable in their production of the vowel /ə/ varying between an [ɪ ɐ] in [d k] contexts

11.92 $Bark^2$) than in Jordanian Arabic (at 17.83 $Bark^2$) or in French (at 20.50 $Bark^2$). It is interesting to note that in both Jordanian Arabic and French, vowel spaces increase in size and dispersion in the CV-Targets highlighting a possible compensation for coarticulated vowels. The vowel spaces using the V-Targets were already larger than those in production by 5 to 8 $Bark^2$ (in words or in isolation, see J. Al-Tamimi, 2007a, Figure 37, pp. 143); in this case, it is possible to hypothesis that when using CV-Targets, a true expansion and an extreme position of the prototypes can be achieved. The results of the LDA showed that across all places of articulation and vowels, the average rate of the V-Targets was around 60%, whereas it was around 67% for the CV-Target; with an improvement of about 7%.

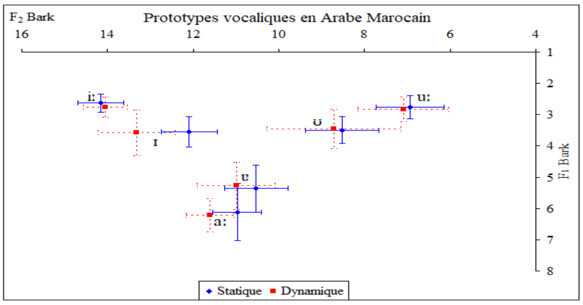


(a) Static (b) Dynamic

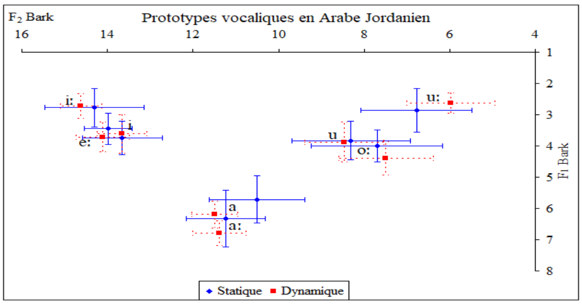
Figure 4.7: Between-language Prototype identification for the point vowels /i a u/ (long vowels in Arabic) in Moroccan Arabic, Jordanian Arabic and French; (a) V-Target; (b) CV-Target (reproduced from J. Al-Tamimi, 2007a, pp. 373 and 380)

Moving on to the within-language results, Figure 4.8 provides a comparison between the vowel spaces for all prototypical vowels identified using the V-Target and the CV-Target in Moroccan Arabic (Figure 4.8a), Jordanian Arabic (Figure 4.8b) and in French (Figure 4.8c). In both Moroccan and Jordanian Arabic, it is clear that the prototypical vowels in the CV-Targets were identified as more peripheral than in French, which seems to show an inverse relationship. Moroccan Arabic listeners were able to identify three separate prototypes for their "short" vowels /ɪ ɐ ʊ/, which were different both qualitatively and quantitatively from the long counterparts (Figure 4.8a). The same can be said for Jordanian Arabic listeners, were the short vowels received the same qualities /ɪ ɐ ʊ/, although they were more peripheral than those in Moroccan Arabic (Figure 4.8b). In both dialects, listeners were identifying vowels differently from each other, whereby the distinctivity was increased using the CV-Target. In French, listeners identified the 11 prototypes as different from each other, although the two pairs /ø œ/ and /a ɑ/ were close to each other possibly signalling merger in perception, although our listeners all spoke a standard version of French as spoken in the city of Lyon.

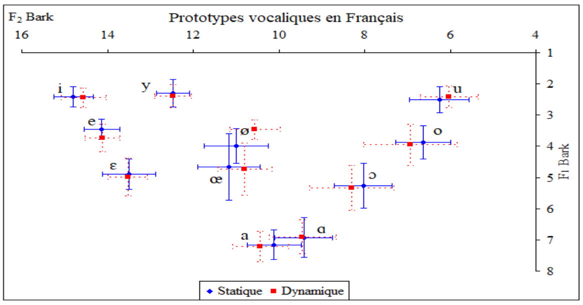
Looking at the results of the LDA, the use of the CV-Target in Moroccan Arabic yielded an average rate of 78% in comparison with that of the V-Target at 86%; a statistically significant decrease in accuracy that is possibly related the having three distinct forms for the vowel /ə/ as /ɪ ɐ ʊ/, as the rates for the point vowels /iː aː uː/ were similar across both types of stimuli varying between 78-98%; most of the confusions were within the short vowels. In Jordanian Arabic, the rates for the V-Target were on average close to 61%, whereas they were around 76% for the CV-Target; with an increase of around 15%. Finally, in French the rates for the V-Target were of 82% against 84% for the CV-Target. Classification rates for individual vowels varied across the three languages from 60% to 95%. Interestingly, listeners of the three languages explained a the



(a) Moroccan Arabic



(b) Jordanian Arabic



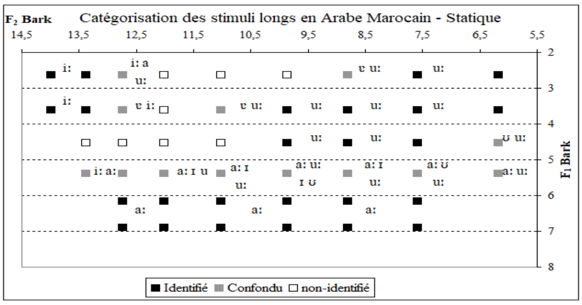
(c) French

Figure 4.8: Within-language Prototype identification in the V-Target vs CV-Target approaches; (a) Moroccan Arabic; (b) Jordanian Arabic; (c) French (reproduced from J. Al-Tamimi, 2007a, pp. 325, 344 and 364)

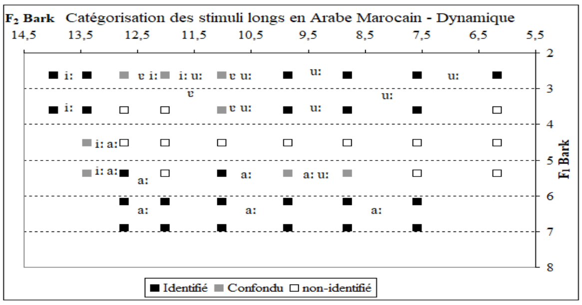
debriefing after the experiments that when the CV-Targets were used, they seemed to correspond closely to vowel quality in comparison to that using the V-Target; the French listeners perceived a slightly "better" vowel quality when using the CV-Target.

The results of the second experiment are presented in Figure 4.9, for the V-Target (Figures 4.9a, 4.9c and 4.9e) and the CV-Target (Figures 4.9b, 4.9d and 4.9f). Overall, the results point to similarities in the patterns of responses reported in the Prototype Identification task and here, across the three systems. When comparing the results of the V-Target to those of the CV-Target, it is clear that all listeners performed slightly differently in the latter, by extending the range of vowels categorised (especially for the back vowels) into more fronted positions; this indicates that coarticulated vowels were categorised differently to non-coarticulated ones. There were of course more confusions in the responses provided by Moroccan and Jordanian Arabic listeners in comparison to the French ones, possibly due to more freedom in responses provided by the former group. This correlates well with the predictions of the "Adapted Dispersion Theory" (Liljencrants & Lindblom, 1972; Lindblom, 1986) and of the "Dispersion-Focalisation Theory" (J.-L. Schwartz et al., 1997) in terms of the internal organisation of vowel spaces and of the impact of density of vowel systems (Manuel, 1990).

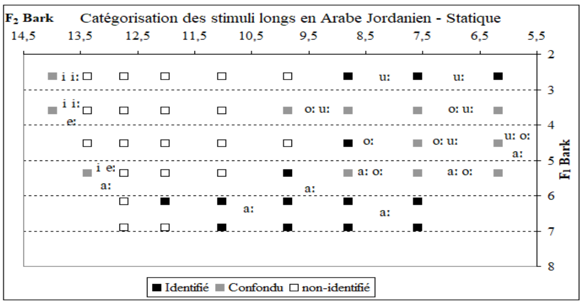
The results of the two perception experiments provided a clear evidence for the role played by the Dynamic specification in terms of the CV-Targets. They both highlighted the fact that contextual information present in the form of transitions added to steady-state vowels (V-Targets) provided a more fine-tuned identification of the Prototypical Vowels and those in categorisation. This provides evidence in favour of a different mental representation of isolated and of coarticulated vowels for the three systems. Our claim here is that CV-Targets attracted the



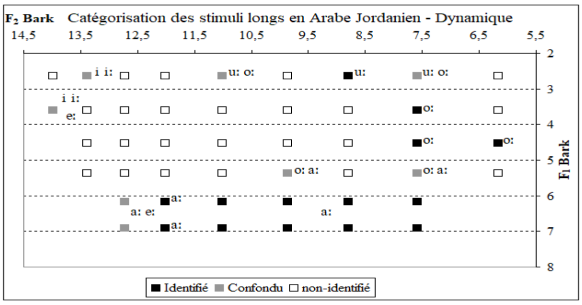
(a) Moroccan Arabic - Static



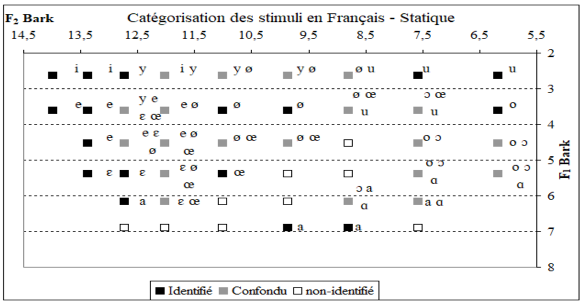
(b) Moroccan Arabic - Dynamic



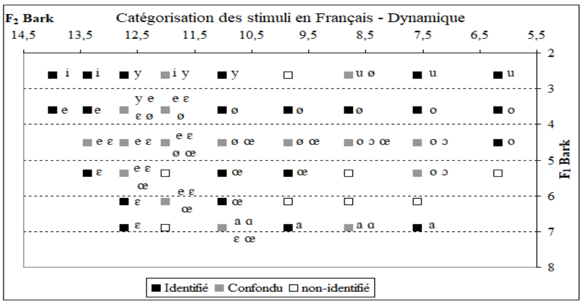
(c) Jordanian Arabic - Static



(d) Jordanian Arabic - Dynamic



(e) French - Static



(f) French - Dynamic

Figure 4.9: Within-language Categorisation of vowels V-Target (a, c, e) vs CV-Target (b, d, f) approaches; (a, b) Moroccan Arabic; (c, d) Jordanian Arabic; (e, f) French. Black: ≥7 stimuli identified as belonging to the same category; Grey: ≥7 stimuli identified as belonging to several categories; White: ≤6 stimuli identified as belonging to several categories; (reproduced from J. Al-Tamimi, 2007a, pp. 398, 403, 408, 412, 418 and 420)

best prototypical form for a target vowel to a different location on the vowel space in both experiments; hence the CV-Target seems to operate as a magnet that exists in perceiving allophones, suggesting that listeners have a rich mental representation of vowels that is not only composed of steady-state V-Target, but also contains details of transitional parts available in the CV-Target (J. Al-Tamimi, 2007a, 2008; Barrett, 1998). This work allowed to unravel new evidence for the rich mental representation that listeners have. It correlates well with the predictions from the Exemplar-based approaches to speech perception whereby listeners store a lot of detail of the sounds they produce and perceive and reuse these when needed (Foulkes & Docherty, 2006; Johnson, 1997; Nguyen, 2001, 2005; Nguyen, Wauquier, & Tuller, 2009; Pierrehumbert, 2001). Our initial hypothesis was that the dynamic specification used here in the form of CV-Targets will benefit heavily Arabic listeners to identify the prototypical vowels of their system. Given that vowels in Arabic are never produced in isolation and do not have a lexical meaning, we suppose that listeners have stored this Fine-Phonetic-Detail in their mental representation, and this facilitated their access to the mental lexicon because they were in our CV-Targets. However, our results showed that even French listeners benefitted from this rich representation and our claim then is that this dynamic specification is part of the grammar as listeners will use it when required by the communicational situation.

4.2.3 Extensions of "Dynamic specification" approaches

My work within the "Dynamic Specification" of vowels looked specifically at the extrinsic approaches in the form of transitional parts identified both as Formant Slopes, in production, and CV-Targets in perception. These were concerned with vowel specification and impact of the onset and the transition from the preceding consonant on vowel discrimination in production and perception.

In what follows, I present some of the work done by PhD students who worked with me on their topics, starting with deriving the locus by integrating the frequencies of F2 (at onset and midpoint) and F3 via a newly developed metric "$F2_R$" (Section 4.2.3.1); accounting for the VISC in L1 Arabic and in L2 Arabic learners of English (Section 4.2.3.2) and finally, dynamic specification of diphthongs and impact of tone in Standard Mandarin (Section 4.2.3.3)

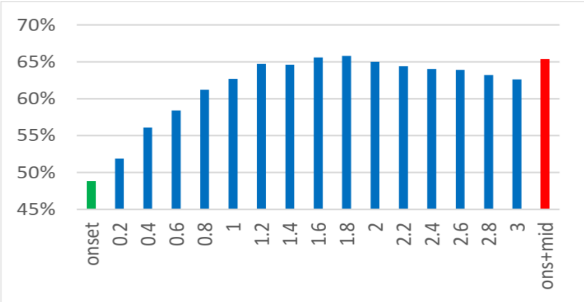
4.2.3.1 Deriving the "Locus"

To provide an insight as to how the Formant Slope approach and specifically Locus Equations can inform classifications of consonants, the work led by Dan McCarthy during his PhD under my supervision (D. McCarthy, 2019) looked at how both onset and midpoint values of F2 and F3 can provide insights into discriminating place of articulation of plosives in British English (in addition to other spectral metrics specific to the burst). Rather than relying on the absolute frequencies of F2 at the Onset and Midpoint in addition to F3 at the onset, to quantify an "invariant" metric that can differentiate consonant's place of articulation (similar to the approach advanced by Lindblom, 1996), D. McCarthy and Al-Tamimi (2019, pp. 3778; following D. McCarthy, 2019, pp. 159-161 and 162-208) devised a new metric named "$F2_R$" (for "F2 Reconstructed" value) that allows to reduce the relationship between F2 (at onset and midpoint) and F3 (at onset) into a simple value using Equations 4.5 and 4.6:

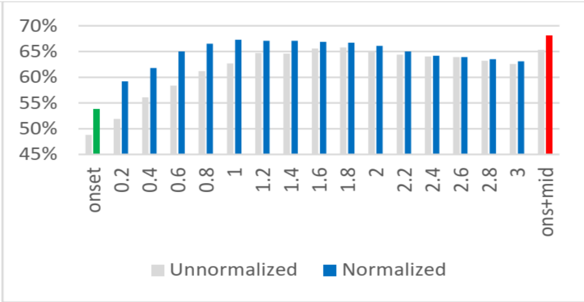
$$F2_R = F2_{Onset} - \{(F2_{Mid} - F2_{Onset}) * c\} \tag{4.5}$$

$$F2_R Norm = F2_R - \mu F3_{Individual} \tag{4.6}$$

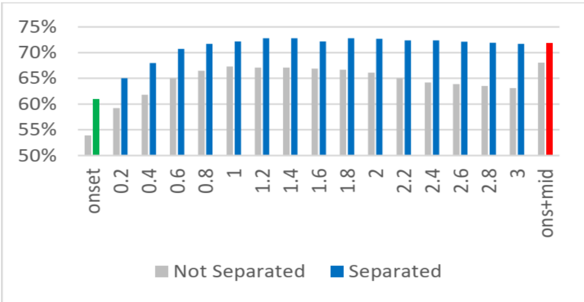
Equation 4.5 integrates a time constant that allows to provide an estimate as to where the virtual locus can be located, while equation 4.6 proposes a normalisation of the $F2_R$ using the average F3 value for each individual subject. Figure 4.10 shows the classification accuracy for various combinations. It is worth noting that when using the $F2_{onset}$ on its own, the rates are too low (less than 50% in majority of cases); using $F2_{onset+mid}$ yields an improvement to classification accuracies; finally, using $F2_R$ with a c constant = 1.8 yields an almost identical classification rate as that of $F2_{onset+mid}$ (65.8% versus 65.3%, respectively, see Figure 4.10a). Thus, it is clear that collapsing these two metrics into one yields a marginal improvement in classification accuracy and prevents loss in recognition. When using the normalised version (Figure 4.10b), it is clear that the classification rates of $F2_R$ at c = 1 are very close to those of combining $F2_{onset+mid}$ (67.3% vs 68.1%). Finally, when separating the vowels into front and back (Figure 4.10c), the rates for $F2_R$ peak at 73% (with c = 1.2, 1.4, 1.8, and 2) over $F2_{onset+mid}$ (at 72%).



(a) $F2_R$ by Time



(b) $F2_R$ Normalised



(c) $F2_R$ by Front vs Back velars

Figure 4.10: Classification accuracy of the $F2_R$; (a) integrating time with a $c = 1$ = no time constant used; <1 = between the release and the vowel onset; >1 = before the release; (b) following the normalised version and (c) separating front and back vowels of the normalised version (reproduced from D. McCarthy & Al-Tamimi, 2019, pp. 3779-3780)

Overall, the results of this study showed that in lieu of using $F2_{onset+mid}$ together or $F2_{onset}$ on its own, a new measure termed $F2_R$ retains all essential information related to the CV transition. This goes hand in hand with the predictions advanced in my own research on the role of Formant Slopes in discriminating between vowels (see Section 4.2.2.1); here we find a similar conclusion on the role of formant transitions at discriminating place of articulation in voiced plosives (see D. McCarthy, 2019, for additional estimations).

4.2.3.2 VISC in L1 and L2

The introduction presented at the beginning of Section 4.2 described the "Dynamic Specification" approaches that were looking at within vowel variation. This was termed as the intrinsic dynamic approach looking specifically at "Vowel Inherent Spectral Changes" (VISC Hillenbrand, 2013; Morrison & Assmann, 2013). This mainly concerned monophthongal vowels, and especially in English. The work led by Wael Almurashi during his MA (IPhD in Phonetics and Phonology summer project) and PhD under my supervision, extended this work to look at the VISC specification in Arabic. This first part of the work conducted during Wael's MA thesis looked first at quantifying the intrinsic dynamic patterns in the productions of 12 male Hijazi Arabic (HA) speakers by examining their role in vowel specification and whether they would allow to differentiate between the long and short vowels. This follows from the results of J. Al-Tamimi (2002, 2007a) suggesting that short vowels in Arabic are spectrally centralised and temporally reduced, in comparison to the long vowels. The work presented in Almurashi, Al-Tamimi, and Khattab (2019, 2020b); Almurashi, Khattab, and Al-Tamimi (2018) compared between the static and dynamic VISC results of HA in an [hVd] environment. Using various

VISC approaches, the results showed that all eight vowels in HA Arabic are best described in terms of VISC in comparison with the static approach. Figure 4.11 show the results of the direction model (computed for each formant as: Direction = Offset80% - Onset20%). Upon examining the results, it is clear that the short vowels /i u/ have a falling F1, while /i: u:/ have a rising F1; /a/ has a steeper falling F1 in comparison to that of /a:/. On F2, all vowels have a rising change, except from /i i: e:/. The results on F1, especially for the [+high] short vowels confirm that they should be described more as similar to the lax vowels in English (following Slifka, 2003, pp. 922). The classification accuracy of the Linear Discriminant Analysis function (LDA) showed that the direction model in addition to duration allowed the discrimination of all eight vowels with a rate of 73% that increase to 90% when separating the tense long vowels from the lax short vowels (see Table I and II from Almurashi et al., 2020b, pp. 2920-2921). When the duration was not included, a reduction in accuracy of about 10-15% was observed; when the three point-model (i.e., onset20%, mid and offset80%) was used, the rates without duration were of 95-98% that increased to 99% when duration was added. These results highlight that short vowel in Arabic are both temporally and spectrally reduced in comparison to their long counterparts.

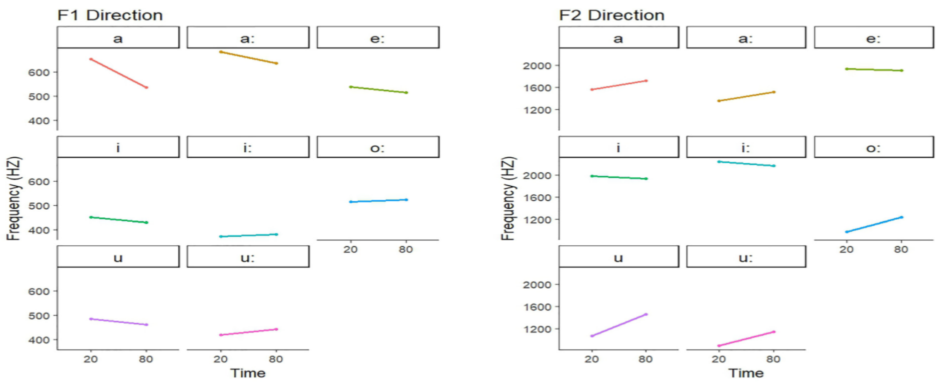
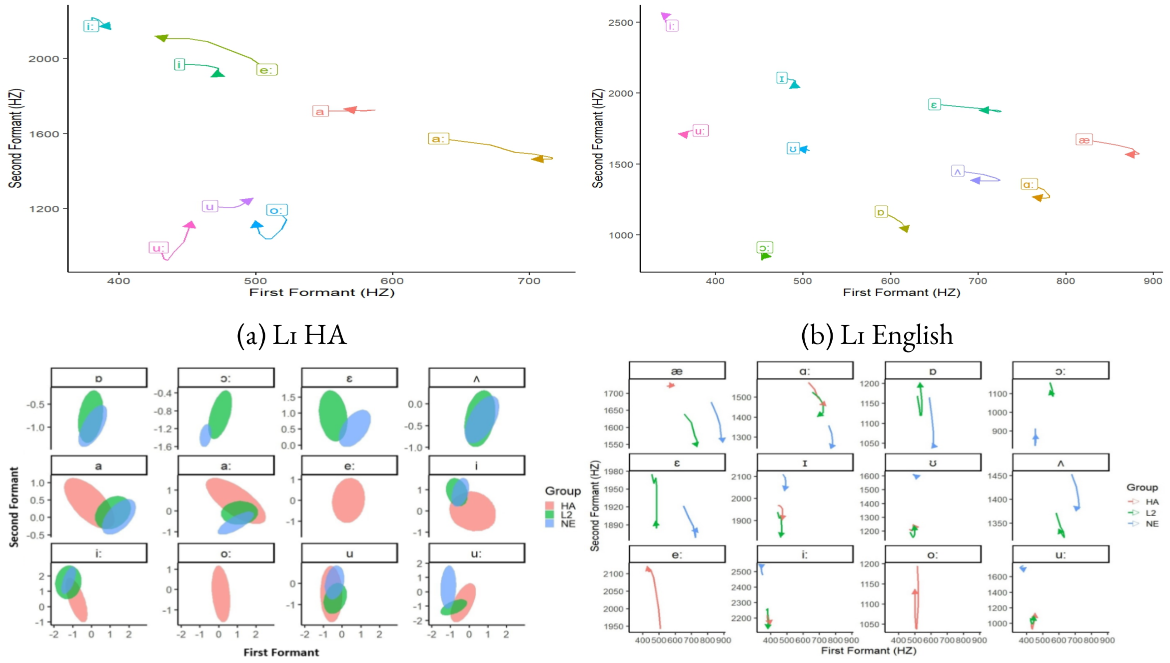


Figure 4.11: Results of the direction model of HA vowels on F1 and F2 (reproduced from Almurashi et al., 2020b, pp. 2923)

As a continuation and expansion to the initial work, Almurashi (2022, see also: Almurashi, Al-Tamimi, & Khattab, Under Revisions on L1 HA; Almurashi, Al-Tamimi, & Khattab, Under Review on L1 English; Almurashi, Al-Tamimi, & Khattab, 2020a, 2021, 2023; Almurashi, Khattab, & Al-Tamimi, In preparation on L2 English) looked at the VISC "Dynamic specification" account of 20 L1 HA speakers' productions in various contexts (10 males and 10 females), of 20 L1 English SSBE speakers' productions (10 males and 10 females) and finally of 20 HA L2 learners of English (10 males and 10 females; intermediate levels of English). The L1 HA and L2 learners were the same is we wanted to track the change between their own L1 HA and L2, in comparison with native L1 English speakers.

Using the static (midpoint) and various dynamic approaches to the VISC (offset, slope, direction, 3-points, 5-points and 7-points), the results presented in Figure 4.12 show the dynamic specification of the vowels in L1 HA, L1 English and L2 English using the 7-points model (Figures 4.12a, 4.12b and 4.12d) in comparison to the static midpoint approach (Figure 4.12c). The results of the dynamic approaches in L1 HA and L1 English point to similarities and differences in the patterns observed for various vowels (see Figures 4.12a, 4.12b). Vowels specified for the features [+High, +Tense] show opposite directions on both F1 and F2; similarly, those specified for [+High, -Tense] show opposite and sometimes variable change. For the remaining vowels specified for [-High, -Low]/[-High, +Low], it is evident that similar patterns of spectral change is observed between the two languages.



(a) L1 HA (b) L1 English

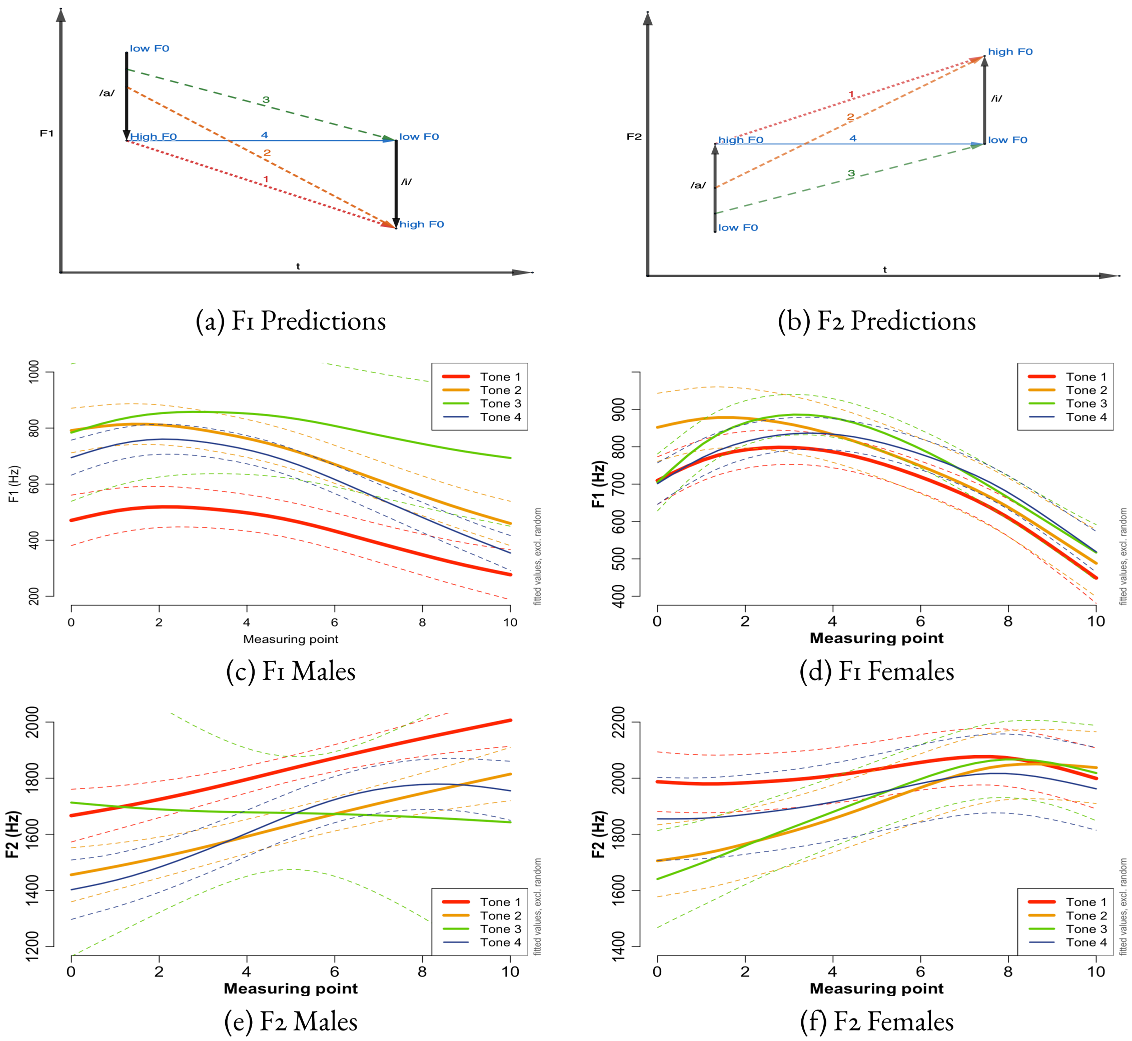
(c) L2 English - Static (d) L2 English - Dynamic

Figure 4.12: Vowel formant trajectories using the 7-point model (a, b, and d) and ellipses around vowel categories (c) for each of (a) L1 HA, (b) L1 English, (c) L2 English static and (d) L2 English dynamic accounts (reproduced from Almurashi, 2022, pp. 88, 115, 149 and 164; see also: Almurashi et al., Under Revisions on L1 HA; Almurashi et al., Under Review on L1 English; Almurashi et al., 2020a, 2021, 2023, In preparation on L2 English)

The results point to an increase in spectral change observed in L1 HA in comparison to that of L1 English highlighting the impact of vowel density on dispersion of vowels seen previously (J. Al-Tamimi, 2007a; Manuel, 1990) and is correlated well with the predictions of the “Adapted Dispersion Theory” (Liljencrants & Lindblom, 1972; Lindblom, 1986) and of the “Dispersion-Focalisation Theory” (J.-L. Schwartz et al., 1997). Interestingly, to our knowledge, these results have not been reported before using the VISC approach. Given the differences between the two languages in terms of the spectral change, we wanted next to assess whether L2 learners of English will produce vowels with spectral change similar to that of their L1 or of the learnt L2. The results presented in Figure 4.12c clearly show variation in terms of vowel dispersion across the L1 HA, L1 English and L2 English. There is a clear overlap between all phonologically “similar” and “different” vowels in Arabic and English (Flege, 1987, 1995). Interstingly, we observe that the L2 learners produce vowels that are closer to the L1 HA and in some instances, they produce a vowel that is between the L1 HA and L1 English. The picture becomes more complex when looking at the patterns using the VISC approach (Figure 4.12d), where the learners are producing vowels that are dissimilar to the two ambient languages; in most instances, they are closer to their L1 HA than to the L1 English, but in other instances, they are similar to the L1 English. Clearly, the VISC approach provides a clearer account of the learners’ variety, highlighting a potential for them to produce vowels that are either closer to the L2 language or that is between the two ambient languages (see Almurashi et al., 2023, In preparation, for more details).

4.2.3.3 Dynamic specification of Diphthongs

In his Master thesis under my supervision, C. Li (2022) used GAMMs (Generalised Additive Mixed-effects Models) to quantify the dynamic specification of the diphthong /ai/ as a function of the four lexical tones in Standard Mandarin: tone 1 (high tone - 55), tone 2 (rising tone - 35), tone 3 (low tone - 21 in natural speech) and tone 4 (falling tone- 51). Inspired by this thesis, C. Li, Al-Tamimi, and Wu (2023) looked specifically at the impact of tones on the /ai/ sequence across males and females to identify how and whether they can describe the degree of diphthongisation-monophthongisation of this diphthong. Figure 4.13 shows the predictions (Figures 4.13a and 4.13b) and the results of the study split by males and females, looking specifically at the F1 (Figures 4.13c and 4.13d) and F2 (Figures 4.13e and 4.13f) curves for the diphthong /ai/.



(a) F1 Predictions (b) F2 Predictions

(c) F1 Males (d) F1 Females

(e) F2 Males (f) F2 Females

Figure 4.13: Formant frequency curves for each of F1 (a and b) and F2 (c and d) for Males (a and c) and Females (b and d) in Standard Mandarin as a function of measurement point (11-points) and according to the four tones (reproduced from C. Li et al., 2023)

The predictions for each of F1 and F2 presented in Figures 4.13a and 4.13b predict a direct relationship between vowel realisations and *f*o, following the predictions of the intrinsic *f*o reported in the literature. Hence, we predicted that a higher *f*o will lead to the vowel to be produced as more close and more front, whereas a lower *f* o leads to a more open and a more back production. Hence, using tones, the predictions are that F1 curves will be higher for a low falling tone (tone 3) and lower with a high tone (tone 1), flatter with a falling tone (tone4) and steeper with a rising tone (tone 2). The F2 curve will have an opposite pattern to that of F1.

The results obtained in C. Li et al. (2023) showed a direct correlation with the predictions, with the tones 1 and 3 showing a partial monophthongisation effect, whereas tone 2 shows a minimal monophthongisation and tone 4 show a flatter F1 curve (see Figures 4.13c and 4.13d), with similar opposite patterns for F2 curves (see Figures 4.13e and 4.13f).

Using a Articulatory Phonology framework, C. Li (2022) started exploring this question, which is currently developed further in C. Li and Al-Tamimi (Under Review) to show how the c-centre effect can also be quantified in the case of diphthongs and is influenced by the lexical tones in Standard Mandarin. Further exploration of dialectal differences between the varieties of Mandarin will be explored in C. Li (In preparation) to evaluate whether there is a relationship between tones and the degree of monophthongisation in the various varieties explored.

4.3 How important is the feature [+Tense] in Arabic?

4.3.1 Introduction

In this direction of research, we were interested in how lexical gemination, i.e., phonological consonantal length distinction is implemented in Lebanese Arabic in adult speech and how children acquire this contrast during their first two years of life. This was part of a project led by Prof. Ghada Khattab at Newcastle University, UK, when I joined the team as a post-doctoral researcher in December 2007. The aim of the project was to identify how infants acquire the gemination contrast in Lebanese Arabic. We started off our work by looking at how adults implemented the gemination contrast in their productions, while assessing the various strategies employed by infants.

Consonantal gemination in Arabic have an important morpho-phonological status as they lead to complex word formation (Versteegh, 2001). When associated with the phonological vowel length distinction, they allow for variation in terms of disyllabic structures that are composed of singleton vs geminate consonants, on the one hard, and of short vs long vowels, on the other. List 2 presented below shows the major disyllabic structures (with examples from Lebanese Arabic), known to be the first five templates in morphophonological word formation that can also covary with stress assignment between trochaic and iambic structure (J. McCarthy, 1979; Versteegh, 2001):

2. 1) ˈCVCVC (/ˈʕadad/ *"number"*)
   2) ˈCVC:VC (/ˈʕad:ad/ *"he enumerated"*)
   3) ˈCV:CVC (/ˈʕa:ded/ *"counting"*)
   4) ˈCV:C:VC (/ˈʕa:d:e(h)/ *"having counted"*)
   5) CVˈC:V:C (/ʕaˈd:e:d/ *"counter"*)

The majority of studies looking at geminate consonants highlighted the fact that the most prominent acoustic correlate is consonantal length, which allows geminate consonants to be assigned the phonological feature [+Long] (Ham, 2001; Ridouane, 2003, 2007, 2010, among others). Reduction in the length of the preceding vowel is sometimes reported in the literature as a secondary consequence of the geminate structure (Esposito & DiBenedetto, 1999; Ham, 2001), while other studies reported changes in the vowel following the geminate (Idemaru & Guion, 2008; Local & Simpson, 1999). It is clear from previous literature that geminate consonants impact heavily on surrounding vowels leading to temporal compensation, which can be quantified as a case of prosodic change on the full word containing a geminate (see Khattab & Al-Tamimi, 2014b). This prosodic change allows possibly for geminates to operate beyond the actual geminated consonant and we used this in our subsequent analyses to evaluate the impact of gemination on syllable affiliation (Khattab & Al-Tamimi, 2014b).

The changes observed on the temporal domain have also been reported to impact on the spectral detail of geminate consonants and their surroundings. For instance, previous literature on other languages reported spectral changes in the consonants themselves or the surrounding vowels, including clearer consonant resonance, especially for sonorants, formant and voice quality changes in the surrounding vowels (Abramson, 1999; Arvaniti & Tserdanelis, 2000; Esposito & DiBenedetto, 1999; Local & Simpson, 1999; Ridouane, 2007).

The interaction between changes observed in the temporal and the non-temporal domains is reported in the majority of studies to simply operate at the phonetic level: due to the longer production of the geminate, any changes observed within the consonant are simple mechanical *phonetic* changes due to the lengthening effect. When a sound is produced for a longer time, the constriction location is reached maximally, which then favours a hyper-articulated form (Lindblom, 1963, 1990; Maniwa, Jongman, & Wade, 2009).

However, various researchers offered an alternative explanation to this contrast. Due to the prevalence of non-temporal consequences of gemination, DiCanio (2012); Kohler (1984); McKay (1980); Nellis and Hollenbach (1980) reported that the gemination vs singleton contrast can be better accounted for as a fortis/lenis or tense/lax contrast. The non-temporal correlates can be seen in various languages as the primary consequence of the contrast that leads to a longer production. For instance, articulatory strength and longer durations are related to each other: Due to the lengthening of geminates, they require longer articulatory strength to maintain the constriction for longer, which leads to higher articulatory effort (Catford, 1977). Additional changes are required when maintaining a longer constriction, which leads to the non-temporal correlates to become more salient in specific languages. Finally, language-specific prosodic constraints impacting on timing needs to be accounted for: in Arabic, both phonetic and phonological timing, syllable structure, the prevalence of phonological consonantal and vocalic length can impact the internal organisation of gemination as a major feature in Arabic phonology (Ham, 2001). This may lead to non-temporal correlates to become more enhanced in Arabic than in other languages; aspects we aimed to quantify in our subsequent studies (see J. Al-Tamimi & Khattab, 2015, 2018b).

In the following sections, we start by looking at gemination (Section 4.3.2), first, evaluating temporal compensation and how a moraic account can explain the patterns in the data. We then look at the acoustic consequences of gemination throughout the VCV sequence in fricative consonants (Section 4.3.3). We next look at the impact of the feature [+Tense] on the acquisition of gemination in Arabic (Section 4.3.4).

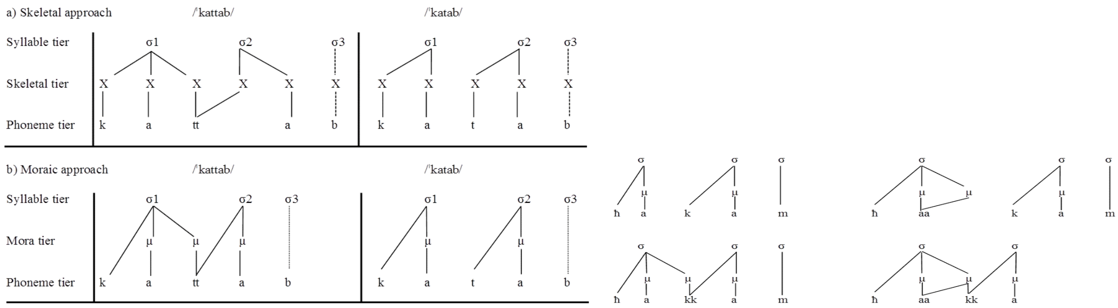
4.3.2 Gemination in Lebanese Arabic - A moraic account

Our first study wanted to assess the phonetic implementation of this phonemic consonantal length distinction and its interaction with the phonemic vocalic length distinction, in this specific dialect. The project as a whole had recordings of about 9000 words, from 20 speakers (10 males and 10 females) who produced the five syllable structures listed above (see List 2 on pp. 68), with the medial consonant being one of the following 24 consonants in their singleton and geminate version /b, t, tˤ, d, dˤ, k, ʔ, f, s, sˤ, z, ʃ, ʒ, x, ɣ, ħ, h, m, n, l, ɾ, w, j, ʕ/ and the vowels surrounding were one of the two open vowels /a aː/, although the long vowel /aː/ is realised as an [eː] in this dialect due to Imala (Nasr, 1966). The following sections present in chronological order the various activities that led to identifying the contrast of singleton vs geminates in Lebanese Arabic to be based on both a temporal and non-temporal contrast, with two active features [+Long] and [+Tense].

Due to the primacy of the temporal domain reported for other languages, we explored durational measures of the singleton/geminate consonant and the vowels surrounding it, in a VCV sequence. In Khattab and Al-Tamimi (2014b, see as well Khattab & Al-Tamimi, 2008), we looked at the absolute and proportional durations of the four syllable structures in the trochaic context from the list above (see items 1) to 4) from the List 2 on pp. 68) for all manners and places of articulation. Our aim was to provide a comprehensive account of the temporal domain and tried to look at how the phonetic implementation can explain the phonological patterning in the data.

Our study wanted to evaluate which of the two non-linear approaches to phonological representations can account for the temporal organisation of the four disyllabic structures including the consonantal and vocalic length contrasts (Figure 4.14). The comparison of the predictions of the autosegmental vs moraic approaches is presented in Figure 4.14a. A major difference between the two approaches is that the former does not directly predict links between phonetic and phonological weight, contra to the latter. Following the Autosegmental approach (see Figure 4.14a; skeletal), in comparison with a singleton consonant (which is associated with a single timing slot), a geminate is associated with two timing slots that can be shared between two syllables (Davis, 2011). The two timing slots allow to predict phonological length but not phonological weight! On the reverse, the moraic theory explains that "*...geminates are attached to a moraic node and a syllabic node and there is a connection between weight and length on the one hand, and weight and syllable position on the other.*" (Khattab & Al-Tamimi, 2014b, pp. 234).

Based on the predictions of the moraic approach (see Figure 4.14a; moraic), a geminate consonant is associated with one mora and when in an intervocalic position, it is linked to the coda of the first syllable and the onset of the second. This has clear consequences for phonological length and weight. In fact, there are various views as to the links between abstract and surface timing. Hayes (1989) predicts that moras only serve as the link between prosodic and segmental information, whereas Cohn (2003) argues for a closer relationship between phonological structure and phonetic timing. This was further explored by Broselow, Chen, and Huffman (1997) who showed that durational patterns between consonants and vowels will be language-specific and depend on the syllable structure. Onset consonants are weightless, while codas acquire a mora due to the Weight by Position rule (Hayes, 1989, pp. 258). In Lebanese Arabic, CV syllables are light, CV: and CVC are heavy, while CV:C and CVC: are superheavy. These differences can predict how mora sharing can operate; in CVC:, the final C shares a mora with the preceding long vowel (following the Adjunction-to-Mora principle, see Broselow et al.,



(a) Predictions (b) Analysis

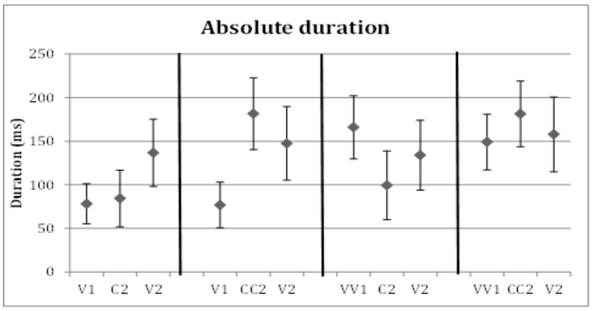
Figure 4.14: Phonological account of gemination in a VCV sequence in Lebanese Arabic, (a) predictions of the skeletal vs moraic approaches to timing, (b) phonological analysis using the moraic approach of the four syllable structures CV(:)C(:)C including mora sharing between consonants and vowels (reproduced from Khattab & Al-Tamimi, 2014b, pp. 234 and 242)

1997, pp. 14–15). For instance, working on Levantine Arabic, Broselow et al. (1997, pp. 59–60) showed that mora sharing impacts on the length of vowels and consonants in different syllable structures: vowels are longer in CV: than CV:C syllables, with no differences between CV and CVC; coda consonants are longer in CVC than in CV:C.

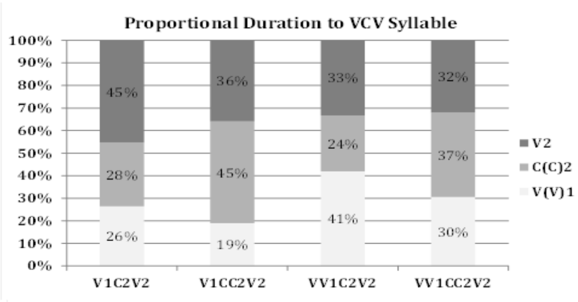
We explore the moraic account on the four syllable structures in Lebanese Arabic (see Figure 4.14b). Looking at the first three cases: ˈCVCVC, /ˈħakam/ ("referee MASC-SG"), ˈCV:CVC, /ˈħa:kam/ ("tried MASC-SG") and ˈCVC:VC, /ˈħak:am/ ("treated MASC-SG"), no mora sharing is operating here. However, in the more complex cases: ˈCV:C:V(C), /ˈħa:k:a(h)/ ("scratched FEM-SG"), mora sharing is evident between the long vowel in the first syllable and the geminate that is shared between the first and second syllables.

Due to mora sharing observed in the phonological analysis presented in Figure 4.14b, we examined the temporal results for the medial V(:)C(:)V sequences in the four syllable structures. Figure 4.15 shows the absolute (Figure 4.15a) and proportional durations as a function of the VCV syllable (Figure 4.15b) or the whole word (Figure 4.15c). Looking at the absolute duration (Figure 4.15a), the geminate consonant (C:) is longer than the singleton (C) as expected from previous literature, and the preceding long vowel (V:) shows temporal reduction when it is followed by a geminate (but not by a singleton!). When looking at the proportional durations (Figures 4.15b and 4.15c), it is evident that the geminate in the syllable structure ˈCV:C:VC is shorter than that in ˈCVC:VC and the vowel preceding the two geminate consonants (ˈCVC:VC and ˈCV:C:VC) are shorter than those preceding a singleton (ˈCVCVC and ˈCV:CVC). Temporal compensation in the vowels is evident in our case and this is direct consequence of mora sharing predicted from the moraic approach.

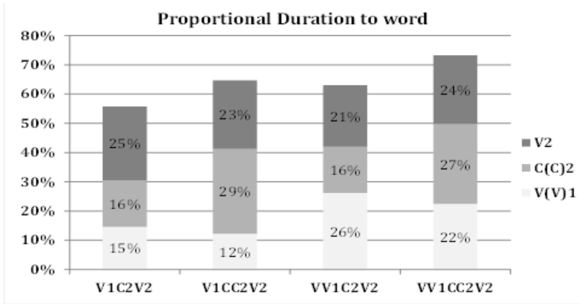
The results reported in the first study explained how gemination operates in Lebanese Arabic by providing a phonological analysis employing a moraic account and a phonetic analysis. The latter confirmed the predictions of the former showing a clear evidence for a connection between phonological and phonetic weight. The conclusion drawn from Khattab and Al-Tamimi (2014b) provides an empirical evidence explaining temporal compensation as observed in the data. In fact, this provides evidence for gemination to operate beyond the segmental level and is better accounted for as a prosodic change. We claim that the overall change observed in the geminate structure causes gemination to possibly operate differently in Arabic; a topic we explored further in subsequent research.



(a) Absolute



(b) Proportional - Syllable



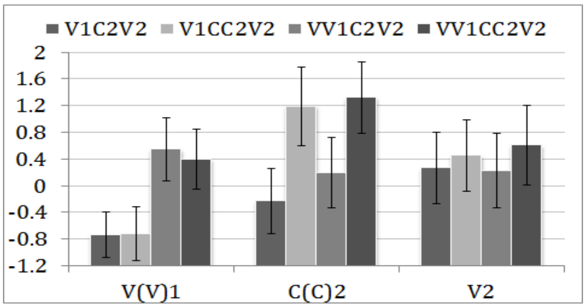
(c) Proportional - Word

Figure 4.15: Durational results for the four syllables within the frame ˈCV(:)C(:)V(C): (a) Absolute duration, (b) Proportional duration as a function of the syllable and (c) as a function of the whole word, for each of the V(:)1, C(:)2 and V2 (reproduced from Khattab & Al-Tamimi, 2014b, pp. 253)

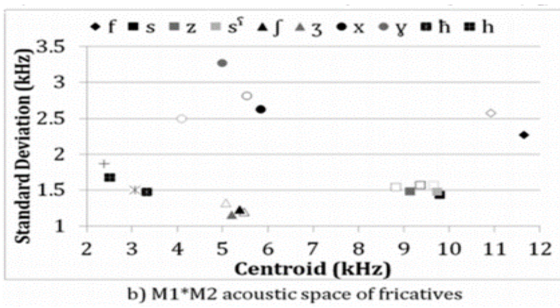
4.3.3 Acoustic consequences of [+Tense] in geminate fricatives

Given the descriptions of how gemination impacts both the temporal and non-temporal domains reported in the introduction (Section 4.3.1) and the possible different behaviour of how phonological weight is implemented, we wanted to evaluate further if and whether gemination is associated with non-temporal correlates reported in the previous literature. Our studies were the first looking specifically at this variety of Arabic where phonological weight can predict phonetic weight and a moraic approach accounts for the temporal compensation observed in the data. We explored the combination of the two types of correlates to assess its strength and provided a way to assess which of the two is to be considered as primary and which one is secondary. Relying on the relational invariance and enhancement of features advocated by Keyser and Stevens (2006); Stevens and Keyser (1989, 2010), we systematically compared the acoustic correlates in the full prosodic V(:)C(:)V sequences (see J. Al-Tamimi & Khattab, 2010 on stops and fricatives, BAAP; J. Al-Tamimi & Khattab, 2011 on stops and fricatives, 17th ICPhS; Khattab & Al-Tamimi, 2015b, on nasals and lateral, GemCon2015 - satellite workshop to 18th ICPhS; J. Al-Tamimi & Khattab, 2015, on fricatives, JASA; see Section 4.4.1 for the interaction between voicing and gemination).

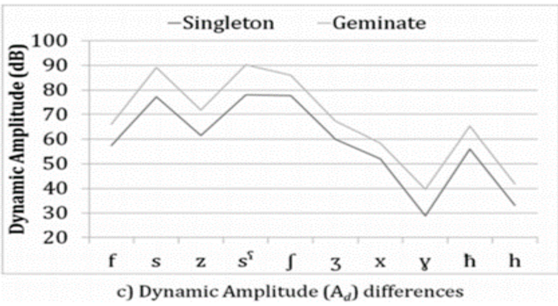
J. Al-Tamimi and Khattab (2015) evaluated a variety of temporal and non-temporal correlates within the medial singleton or geminate fricative and the vowels preceding and following it. A total of 45 measures spanning a combination of temporal and non-temporal correlates, such as: intensity; fundamental frequency; spectral peak and shape, dynamic amplitude, and voicing patterns of medial fricatives; vowel quality and voice quality of surrounding vowels. This is by large the most comprehensive study that accounted for all possible changes that were in majority reported in previous literature, with explorations of new measures not used specifically in the singleton vs geminate contrast (spectral shape and dynamic amplitude). Figure 4.16 reports the main findings for the temporal (Figure 4.16a), the M1*M2 acoustic space for fricatives (Figure 4.16b), the Dynamic Amplitude A_d (Figure 4.16c) and the voicing patterns in the medial voiced or voiceless fricatives (Figure 4.16d). We used z-scored values for all acoustic results, and reported on raw results when needed; the former allowed to put all acoustic predictors on an equal weight, while the latter allowed to evaluate specific predictors' contributions.



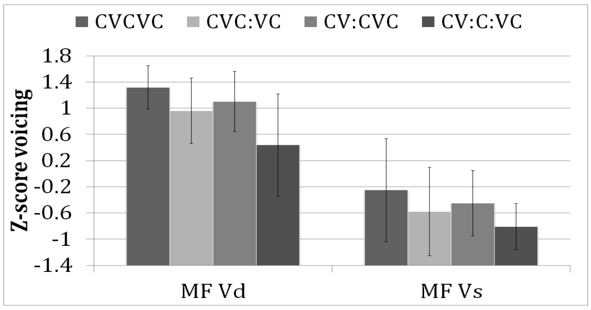
(a) Duration



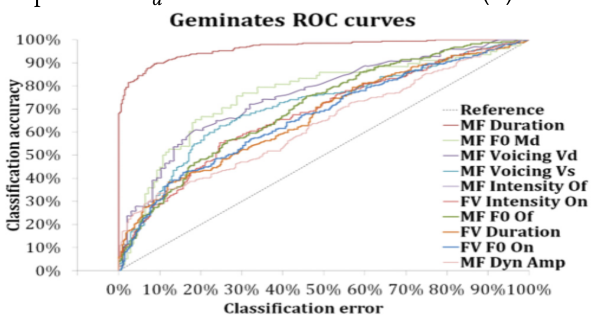
(b) M1*M2 acoustic space



(c) Dynamic Amplitude - A_d



(d) Voicing patterns



(e) ROC curves

Figure 4.16: Results for the singleton vs geminate contrast in fricative consonants: (a) z-scored duration for V(:)1, C(:)2 and V2; (b) M1*M2 acoustic space for fricatives in kHz; (c) Dynamic Amplitude A_d in dB; (d) Voicing patterns in medial consonants; (e) ROC-Curves for classifications, PV = preceding vowel; MF = medial fricative, FV = following vowel, Vd = voiced, Vs = voiceless, On = onset, Md = midpoint, Of = offset, Dyn Amp = dynamic amplitude A_d, (reproduced from J. Al-Tamimi & Khattab, 2015, pp. 350, 354, 355 and 358)

First, Figure 4.16a shows the temporal results for fricatives and surrounding vowels and they do match those reported previously (see Figure 4.15, pp. 72). Temporal compensation of the preceding vowel was only evident when the vowel was long (i.e., ˈCV:CVC vs ˈCV:C:VC); in the latter, the vowel was shorter by around 9 ms. The medial geminate was almost two-folds longer that the singleton; there was shortening of the geminate in the more complex syllable structure ˈCV:C:VC when compared with the simpler structure ˈCVC:VC. Finally, an interesting pattern not reported often in the literature showed the following vowel to be impacted: A longer V2 was observed in the two geminate structures, which was even longer in the more complex structure ˈCV:C:VC.

Moving on to the spectral shape that was estimated using time-averaged spectra, we proposed an acoustic fricative space by plotting the Centroid (M1) on the x-axis against the Standard Deviation (M2), using the raw values in kHz. The results presented in Figure 4.16b show how each of the singleton (white filled) and geminate (dark filled) are scattered. It is clear that a geminate fricative showed an increase in M1 and a decrease in M2, which are partially correlated with an increase in effort reported in the previous literature on clear vs casual speech (Maniwa et al., 2009) or on vocal effort (Jesus & Shadle, 2002).

Next, we reported for the first time on the results of the Dynamic Amplitude A_d measures that was previously used to assess vocal effort in fricatives produced in a clear speech format (following Jesus & Shadle, 2002). This measure looks at the difference in the amplitudes of the high frequency (0.5–18 kHz) and the low frequency regions (0-2 kHz). Figure 4.16c shows that geminates have a systematically higher A_d than singletons and this was evident across all places of articulation. This again correlated with an increase in effort and air pressure associated with geminates (Catford, 1977; Nellis & Hollenbach, 1980).

Lastly, we looked at the voicing patterns between phonologically voiced and voiceless fricatives. Figure 4.16d shows that when the fricative was in a geminate structure, it showed systematically less voicing frames, regardless of its phonological status. In the phonologically voiced fricatives, there was more devoicing, whereas in the phonologically voiceless fricatives, there was less voicing lead (or voicing shadow). These results are comparable with the patterns normally found in tense (fortis) categories (Jaeger, 1983).

Overall, the results reported in J. Al-Tamimi and Khattab (2015) showed variable patterns depending on how each predictor contributed to the singleton vs geminate contrast, with variable effects sizes. To assess the strength of each predictor, we used Linear Discriminant Analysis (LDA) on each predictor and then reported the AUC (Area Under the Curve) of the ROC curve (Receiver Operating Curve, following Swets, Dawes, & Monahan, 2000). Figure 4.16e shows the AUC of the top ten classification models that best distinguished the singletons from the geminates in a binary classification task. It is clear that the duration of the medial consonant is the best predictor to the contrast (with 89% classification accuracy; 96% AUC), followed by variable predictors, including f0 within the phonetically voiced medial fricatives, %Voicing, intensity, duration and f0 of the following vowel and the Dynamic Amplitude A_d, with rates between 51% and 71% and AUC percentages ranging between 51% and 76% (see J. Al-Tamimi & Khattab, 2015, Table 3, pp. 357). One important note here: the results reported in Figures 4.16b and 4.16c showed clear patterns that were evident across places of articulation; when using the binary classification task, place of articulation was ignored, which is likely to be the reason behind not finding these predictors in the top ones from the LDA nor from the AUC.

Overall, there results reported in J. Al-Tamimi and Khattab (2015) were conclusive as to the primacy of the medial fricative duration as the primary correlate that can be represented with the feature [+Long] following an SPE-based approach to feature specification (Chomsky & Halle, 1968). This confirms previous research on gemination in general and in Lebanese Arabic.

However, the systematic changes observed for the geminate structures on the non-temporal predictors is a clear indication that these secondary correlates have an important role as various features used to enhance the primacy of the primary feature quantified via consonantal length (following Keyser & Stevens, 2006; Stevens & Keyser, 1989, 2010). In our case, these secondary features are similar to those reported previously as potential correlates for the feature [±Tense] (Jessen, 1998; Jessen & Ringen, 2002; Kohler, 1984). Our claim then is that both [±Long] *and* [±Tense] are active features in Lebanese Arabic, with the geminate being specified as both [+Long] and [+Tense] in comparison with the singleton being specified for the features [−Long] and [−Tense].

4.3.4 The importance of [+Tense] in child acquisition

Given the results reported above (see Sections 4.3.2 and 4.3.3), it is clear that geminates in Arabic impact on the overall structure of a syllable. Words with medial geminates are produced differently from singletons and have a prosodic impact on the overall structure of the syllable. This is true for any of the syllable structures listed in the List 2 (pp. 68). We also showed that the more complex syllable structure ˈCVːCːVC allows for mora sharing to occur, which impacts directly on the temporal compensation observed in our data. A secondary consequence to gemination is tenseness, assigned the feature [+Tense] that we claimed is active in Lebanese Arabic due to the multitude of acoustic consequences reported in J. Al-Tamimi and Khattab (2015) on fricative consonants.

While working on identifying the phonetic correlates and feature specification of gemination in adult speech, we also looked at how children acquired gemination in Lebanese Arabic, first by looking at the phonetic correlates in geminate acquisition and then by identifying the cross-linguistic patterns differentiating child acquisition across English, Arabic, and French (Project funded by the ESRC, between 2009-2012, PIs: Marilyn Vihman, Tamar Keren-Portnoy, Rory DePaolis, Ghada Khattab and Sophie Wauquier). These double projects and the work conducted within them led to various publications and presentations (see Khattab & Al-Tamimi, 2011a, Child Language Seminar; Khattab & Al-Tamimi, 2011b, 12th IASCL; Khattab & Al-Tamimi, 2011c, International Child Phonology Conference; Khattab & Al-Tamimi, 2012a, International Child Phonology Conference; Khattab & Al-Tamimi, 2012b, BAAP; Khattab & Al-Tamimi, 2013a, Book chapter "the Emergence of Phonology"; Khattab & Al-Tamimi, 2013b, International Child Phonology; Khattab & Al-Tamimi, 2014a, BAAP; Khattab, Vihman, Al-Tamimi, Nakai, & Kunnari, 2014, LabPhon Conference; Khattab & Al-Tamimi, 2015a, 18th ICPhS; Khattab & Al-Tamimi, 2017, 14thIASCL).

Taking as a basis for comparison the results obtained for adults producing gemination, we looked at the developmental trajectories via a longitudinal study spanning the period of acquisition of the first 50 words (≈between 12 and 18 months). Longitudinal audio-visual recordings of spontaneous mother-child interactions were obtained at the home of the participants, who were in a multilingual setting, with Arabic as the main primary language, but due to the sociolinguistic context in Lebanon, French and English are also ambient languages spoken sometimes at home by their parents. Each session was recorded once per month, for around 30-45 minutes per session, when the children were aged from 9 months to 3 years old, with a total 20 children recorded.

The data from only 10 children (5 in Beirut and 5 in London) were processed using Praat

(Boersma & Weenink, 2013) and Phon (Rose et al., 2006). The former allowed for segmentation of the acoustic signal, while the latter was used to identify visual cues to identifying words as produced by the children, which was also informed by the acoustic signal. Over 100 sessions (30-45 minutes each) from the 10 children aged between 10 and 20 months were processed (in addition to those from their mothers).

We first transcribed all productions from the children, and used the criteria set in Vihman and McCune (1994) to decide whether a word is a real word and not a simple "jargon" or productions that are out of context. Here the auditory and visual contexts were important as they allowed us to unpack and decide whether specific productions were simple repetitions after the mother, or true spontaneous productions from the child. In total, I transcribed impressionistically 11875 words that were produced across the 10 children, and from over 100 sessions. This impressionistic phonetic transcription was informed by both auditory and acoustic cues that were essential to identify minute articulations specific to the children without imposing any phonological type of productions. Due to the prevalence of the disyllabic structures in the productions of the mothers, we decided to only look at disyllabic productions either "targetted" or "realised" by the children. By "targetted", we mean the words that were identified as belonging to one of the words that have been identified due to the context (visual, repetition, etc.), while the "realised" were those that were produced by the children. This yielded 5697 disyllabic words targetted/realised by the ten Lebanese Children. Based on our detailed phonetic and phonological analysis of the data, we identified a templatic pattern composed of the syllabic shape ˈCV(ː)C(ː)VC. This templatic pattern was the most targetted by children with their realisations ranging between monosyllables, to disyllables, to multisyllables.

We profiled the productions of the five children living in Beirut, a total of 2863 words were targetted/realised. Our results showed that overall 65% of the words targetted were in Arabic, 18% in English, 8% in French and 9% multilingual (i.e., words with multiple origines, see Khattab & Al-Tamimi, 2013a, 385). Figure 4.17 presents the distribution of the words targetted or realised by the children, in the templatic frame ˈCV(ː)C(ː)VC. In the targetted words, 51% contained a singleton consonant, 28% as geminate and 21% as "complex" (containing secondary articulations or affricates). In the words realised by the children, 19% contained a singleton, 54% contained a geminate and 30% contained a complex "realisation" identified as either a "tense" or a "fortis" production. Examples of productions from a Child, aged 1;7 for the word /ˈmaːma/ "mummy": [ˈh̃ɐ̝mæˑ], [ˈmˑɐ̃], [ˈmæ̃ˑʷmˑɐˑh] and [ˈmæ̠̃ˑmmæ̃ˑ]. The use of the half long diacritic [ˑ] is important here as it signals within consonants a "fortis"-like production with short production. This can be quantified as half a geminate or a [+Tense, −Long] consonant. This clearly shows how variable children were in their productions with different types of short, strong and long consonants.

One of the most important findings from our first study was that children produced the targetted median consonants not as based on a binary split, but rather tertiary: there were simple vs geminates and "fortis"/"tense" productions in-between. These results show a direct link with those obtained in the adults data, except from the fact that our results from adults showed a binary distinction, with emphasis on secondary correlates of the [+Tense] feature being more important in Arabic than in other languages. It seems then that children acquiring gemination in Arabic were influenced by the input received and were latching on specific types of productions that signalled the importance of the feature [+Tense]. As highlighted previously (see end of

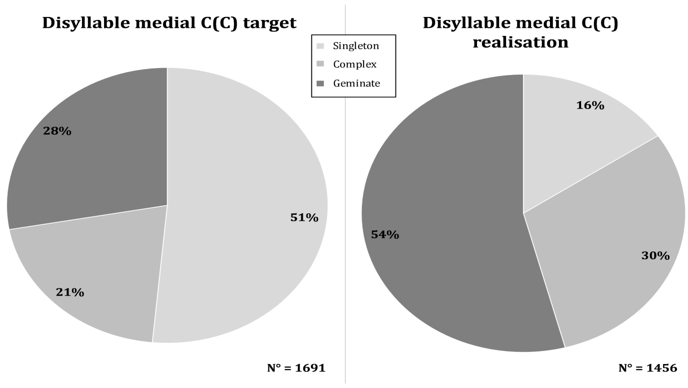
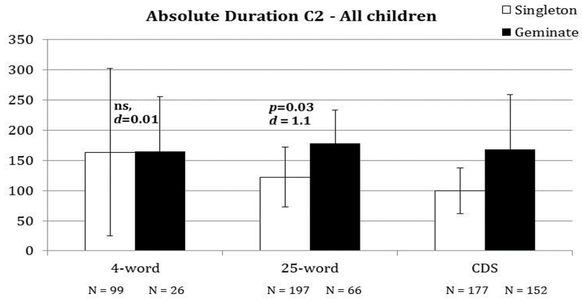


Figure 4.17: Pie charts representing disyllabic productions by Lebanese Arabic children, targetted (left) and realised (right) with percentages for the three categories: simple, complex or geminate (Khattab & Al-Tamimi, 2013a, pp. 387)

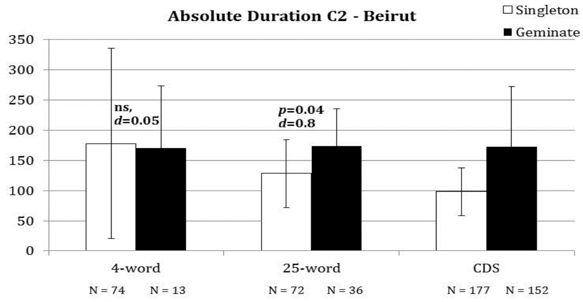
Section 4.3.3), gemination in Arabic can be described as having the two features [+Long] as primary and [+Tense] as secondary. The results presented so far for children acquiring gemination in Arabic suggest that around 84% of the words contained the feature [+Tense], while only 54% the feature [+Long]. Our interpretation of this is that the gemination contrast in acquisition can be described in terms of the two features [+Tense] as primary and [+Long] as secondary; an opposite to that of the adult, however, this account puts more emphasis on the feature [+Tense] is being important in the phonology of Arabic. Of course our initial results were based on the productions of children living in Beirut; we extended the analysis to those living in London and the same patterns emerged.

The first patterns reported here were concerned with all productions regardless of age. We continued our evaluation of how the singleton/geminate contrast is acquired relying on acoustic measurements of durations of the consonants. Our results suggest that singleton consonants were generally more stable and closer to the productions of adults in CDS (Child Directed Speech), whereas the productions of geminates were more variable (Khattab & Al-Tamimi, 2013b, 2014a, 2015a). We looked specifically at the two critical periods identified by Vihman and McCune (1994), when the child is around 1 years old and has a small productive vocabulary, with approximately 4 words spontaneously produced in a 30-minutes setting, and ending around 6-8 months later, when the child's vocabulary progressed to around 50 words spontaneously produced in a 30-minutes setting. These two critical periods are referred to as the 4-words and the 25-words periods.

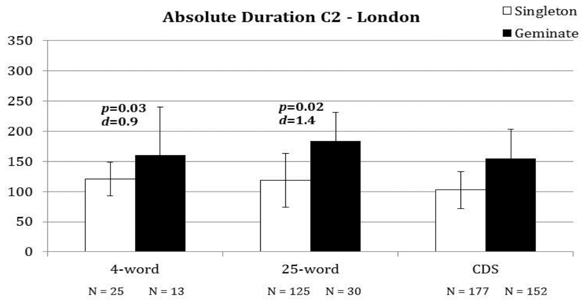
Figure 4.18 shows the overall results for these patterns. Specifically, starting with Figure 4.18a, it is clear that in comparison with CDS (obtained from the mothers' productions in similar contexts), all 10 children showed no differences between the singleton and geminate consonants at the 4-words period, but the temporal differentiation started to emerge at the 25-words period. There were clear differences between the settings where the children living in Beirut did not show differences between the two categories at their 4-words period, but those living in London did show a difference (see Figure 4.18b for the former and Figure 4.18c for the latter). It is interesting to note that at the 25-words period, the productions from children in both settings show similarities in their ranges that, although follow a similar pattern to that of the adult in CDS, are still produced as longer than those in CDS. In both settings, it seems that children were variable in their productions at both critical periods, which were clearly influenced by the input received. It seems that children living in Beirut received variable type of input, where those receiving only



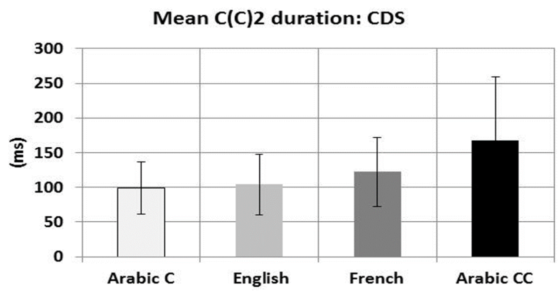
(a) All Children



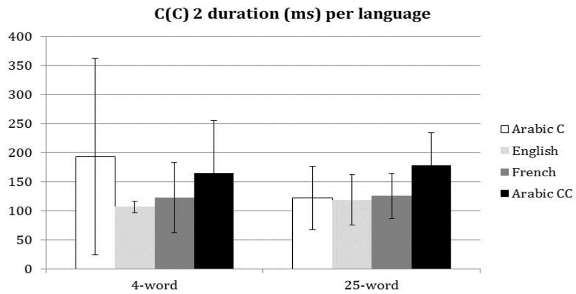
(b) Beirut setting



(c) London setting



(d) Per Language - CDS



(e) Per Language - Child

Figure 4.18: Results for the singleton vs geminate contrast in acquisition: (a) Durations at 4-words, 25-words and CDS for all children; (b) those living in Beirut; (c) those living in London; (d) Multilanguage patterns in CDS; (e) in Child acquisition at the 4-words, 25-words, (reproduced from Khattab & Al-Tamimi, 2013b, 2014a, 2015a)

Arabic as an input showing more stability in their realisations, in comparison to those receiving variable multilingual input. Those living in London seem to have received a marked difference between singletons and geminates, and seem to be more stable in their productions (Khattab & Al-Tamimi, 2013b, 2014a, 2015a).

Looking at the multilingual productions, Figure 4.18d presents the temporal information for the medial consonants produced in CDS in Arabic singleton (/C/) and geminates (/C:/), with those in French and English. As can be seen, Arabic singleton consonants are almost of a similar length to those in English and French; geminates are 70 ms longer (on average) from Arabic singletons. When looking at child productions presented in Figure 4.18e, it is clear that at either 4-words or 25-words periods, children are producing similar productions in English and French, slightly longer productions in Arabic geminates, with opposite patterns for Arabic singletons. At 4-words, Arabic singletons are extremely variable and are actually produced as longer than Arabic geminates; at 25-words, Arabic singletons are shortened and are of comparable length to that of English or French. All in all, the patterns of productions seen in child speech are similar to those of adults CDS, but are much longer due to the fact that they do not yet have a fully matured motor control. However, it is interesting to see that children are following similar patterns as those of the adults' CDS productions.

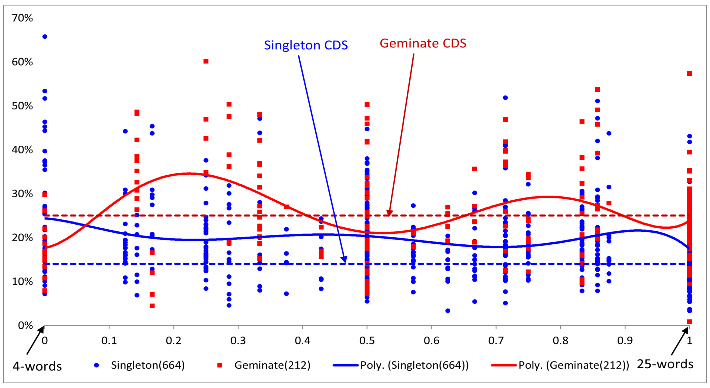


Figure 4.19: Gemination contrast development overtime, with individual datapoints for Singleton (blue) and Geminates (red) across all children and 6^{th} order polynomial regression curves (reproduced from Khattab & Al-Tamimi, 2012a, 2012b)

Given that we followed the children during various sessions (minimum 5, maximum 8), it was important to track the developmental trajectories in their acquisition of the singleton vs geminate contrast. Figure 4.19 presents an account of the change overtime by showing the range of CDS for singleton and geminates (linear horizontal lines, in blue and red respectively), individual datapoints across all children for both categories, and a 6^{th} order polynomial regression curves for both categories, and the x-axis is the normalised time from the 4-words to the 25-words periods (between 0 and 1, respectively). It is clear that children were following variable trajectories when looking at their individual results, but overall patterns are interesting in showing that the geminate category was the most variable, with the singleton being less variable. However, the actual durational values of the singleton consonants were more similar to those of the geminate productions in adults CDS. The geminates were showing a clear case of hyper-articulation or overshooting as suggested by (Scobbie, Gibbon, Hardcastle, & Fletcher, 1996), whereby children's productions reach extreme position to potentially signal the contrast to adults. When looking at the 4-words period, it is clear that the patterns are reversed, where singleton consonants were much longer than geminates; a pattern that was reversed after the first session. It is likely that the immature vocal tract in these children led to hyper-articulated forms for both categories; a pattern that was stabilised throughout the remaining sessions. Similarities in the durational measures were observed towards the middle (around 2-4 months after the first session; see time = 0.5); double peaks before and after. At the end and towards the 25-words period, a divergence between the two contrasts can be seen whereby singletons in child speech became much closer to those in CDS; the same can be said for geminates. Our results allowed to unravel new evidence for the role of input in shaping acquired phonological structures. Parents were already variable in their productions, in Arabic but also in French and English, and they favoured the production of disyllabic structures in Arabic and French, but monosyllables in English; the same patterns were observed in the child productions (Khattab & Al-Tamimi, 2015a). Our analyses in child acquisition were informed by the clear patterns observed in the adult speech;

without quantifying the major role of the feature [+Tense] in adult speech, we would likely have not identified these patterns in child speech and likely to have suggested that these results were simple reflection of the immature vocal tract in children leading to longer productions overall. The prevalence of the disyllabic templatic shape ˈCV(:)C(:)VC in adult speech was also found in child speech. This disyllabic templatic shape with a medial geminate is a productive shape that children were latching on and reproducing to allow to form their productions (Khattab & Al-Tamimi, 2013a, 2015a). Given that the input was rich in these shapes, it is likely that children acquired these forms and developed rich mental representations that contained this templatic shape. A usage-based and exemplar-based approaches explain best these results. Children were exposed to variable productions in the three languages they were exposed to (Arabic, French and English) that led to multiple acoustic correlates (both temporal and crucially non-temporal) representing the singleton and geminate contrasts to be stored in their mental representations (Johnson, 1997; Nguyen, 2005; Nguyen et al., 2009; Pierrehumbert, 2001, 2006). As highlighted previously, research has shown that speech perception is gradient and dynamic, which allows for multiplicity of correlates to be used to influence the lexical items stored in the mental representation (McMurray, 2022). It is likely that information related to the language structure in terms of the prevalence of the disyllabic structure with gemination and the fine-phonetic-detail for the features [+Long] and [+Tense] are stored. Productions by the mothers in French and English contained clear lengthened productions of the medial consonants (Khattab & Al-Tamimi, 2015a), leading to the same items produced by children to contain lengthening and strengthening features similar to those produced by the adults. Finally, it is clear that there is a match between the patterns reported in both adult and child speech; gemination in Arabic is represented by the two features [+Long] and [+Tense], with a clear preference for children before the 25-word period for the feature [+Tense] to be primary, which changes throughout their acquisition to likely become secondary.

As a follow up to this line of research, we looked at how gemination impacts the voicing profile in Arabic (Section 4.4), by examining the interaction between gemination and voicing to evaluate the presence of a possible 4-way contrast (Section 4.4.2). As an extension to the work done on the voicing profile in Arabic identified, we look at two studies done by my PhD students who examined how voicing is implemented in two varieties of Saudi Arabic that show opposite feature specification (Section 4.4.3). We look at the results from Najdi Arabic that shows a special case of the two features [+Voice] and [+Spread Glottis] being active in stops (Section 4.4.3.1) and then at Jazani Arabic where the two features [+Voice] and [+Tense] are active in both fricatives and stops (Section 4.4.3.2)

4.4 The Voicing contrast in Arabic

4.4.1 The Voicing contrast across languages

In this work, we further explored the phonological features that can best describe the voicing contrast in stop consonants, first in interaction with gemination in Lebanese Arabic. J. Al-Tamimi and Khattab (2018b) examined both primary and secondary correlates that were reported in the world's languages to mark the voicing contrast in stop consonants. Arabic, is well-known to be classified within the Voicing languages, where the Voiced set is predicted to show pre-voicing, whereas the Voiceless set is not associated with any aspiration portion (or positive VOT), regardless of the position. This pattern is the reverse when one looks at Aspirating languages (e.g., English or German), whereby the Voiced set is traditionally associated with devoiced productions in initial stressed syllables, but with passive voicing intervocalically, whereas the Voiceless set in initial stressed syllables is produced with increased duration for the aspiration portion (or positive VOT). Traditionally, this dichotomy was used by Lisker and Abramson (1964) and revised by Abramson and Whalen (2017) as a way to classify languages as belonging to one or the other category. This dichotomy led researchers to identify the VOT as the main correlate to differentiate between Voicing and Aspirating languages. For instance, negative VOT is claimed to be the main correlate for the former, whereas positive VOT is the main correlate for the latter (see Abramson & Whalen, 2017; Lisker & Abramson, 1964, among others).

Table 4.1: Phonological account for the voicing contrast following SPE, traditional and non-traditional approaches, partially following J. Beckman et al. (2013, Table 1, pp. 260); HSP = Heightened Subglottal Pressure; SG = Spread Glottis.

Language	Orthography	Phonetic	SPE	Traditional	Non-Traditional
Voicing Language	*p, t, k*	[p t k]	[−Voice, −HSP]	[∅]	[∅]
	b, d, g	[b d g]	[+Voice, −HSP]	[Voice]	[Voice]
Aspirating Language	*p, t, k*	[pʰ tʰ kʰ]	[−Voice, +HSP]	[∅]	[SG]
	b, d, g	[p t k]	[+Voice, −HSP]	[Voice]	[∅]

From feature specification point of view, there is a debate surrounding the features that best account for the voicing contrast (see Table 4.1 for more details). Traditionally, and following the SPE approach, Chomsky and Halle (1968) described the two features [±Voice] and [±HSP] (for Heightened Subglottal Pressure)[7] as the main features that can account for cross-linguistic differences in the voicing contrast. As highlighted in Table 4.1 (column SPE, following Chomsky & Halle, 1968, Table 8, pp. 328), Voiced consonants in stressed syllables in English (and Aspirating languages) written orthographically as *b, d, g*, are phonetically realised as [p t k][8] rather than *[b d g]; they are produced as [b d g] intervocalically due to passive voicing. They are often referred to in the literature as "lenis" consonants (as opposing "fortis" for the Voiceless aspirated set). This specific category is assigned the features [+Voice, −HSP]. The Voiceless aspirated set, written orthographically as *p, t, k*, are phonetically realised as [pʰ tʰ kʰ] specifically is stressed syllables; in unstressed syllables, they are realised as [p t k]. From a feature specification point of view, the former group is assigned the features [−Voice, +HSP], while the latter only differs in the sign of HSP (i.e., [−Voice, −HSP]). Looking at the case of Voicing languages, it is clear that the feature specifications are similar, albeit for HSP and the fact that "aspiration" is not specified for these languages. The assumption then is that the feature [±Voice] has the same status across the two types of languages. This seems to have been followed by proponents of the "traditional" approach described in J. Beckman et al. (2013): [Voice] serves as the unique feature to describe the Voiced sets in both Voicing and Aspirating languages (see Table 4.1, column Traditional, following Keating, 1984; Kingston & Diehl, 1994, among others). While this approach is appealing as it highlights the need to have one feature across Voicing and Aspirating languages with multiple acoustic correlates serving to describe the contrast in an elegant way (see Kingston

[7]The feature [±HSP] (for Heightened Subglottal Pressure) was subsequently replaced by the feature [±SG] (for Spread Glottis) described in Halle and Stevens (1971)

[8]Or /b̥ d̥ g̊/. It should be noted that phonetically, the two forms can be considered as identical, though [p t k] indicates a truly voiceless production, whereas [b̥ d̥ g̊] can be indicative of partial devoicing.

& Diehl, 1994, Table 1, pp. 427), it is important to note that there are clear differences in how the feature [Voice] is implemented in both types of languages (see below).

On the contrary, and following the Laryngeal Realism approach (Honeybone, 2005; G. K. Iverson & Salmons, 1995), Non-Traditional approaches assume two privative features that are used to distinguish the two types of languages (see Table 4.1, column Non-Traditional J. Beckman, Jessen, & Ringen, 2009; Harris, 1994; Honeybone, 2005; G. K. Iverson & Salmons, 1995; Jessen & Ringen, 2002). The results of these studies are in favour of putting emphasis on cross-language differences in how the Voicing contrast is implemented. For instance, it is clear from Table 4.1 (column Non-Traditional) that the feature [Voice] is only active in the Voiced set in Voicing languages, whereas the feature [SG] (for Spread Glottis), is only active for the Voiceless set in Aspirating languages; the other two receive [∅] as they are not active in either case. This account can easily be criticised in terms of not accounting for the case of unaspirated Voiceless stops in English, however, the fact is that this specific category will receive the same feature specification as the Voiced set in Aspirating languages (i.e., [∅]), given that phonetically speaking, the two are realised as [p t k] (J. Beckman et al., 2013, although see, Kingston & Diehl, 1994 for an alternative view).

Based on the views advanced by Jessen (1998, 2001), it is clear that two auditory features can account for the differences across Voicing and Aspirating languages: the features [Voice] and [Tense]. Jessen (2001, Figure 1, pp. 244) described the basic correlates for each feature and the shared non-basic features. [Voice] has "closure voicing" as the basic correlate; "aspiration duration" for [Tense]. The latter is in fact similar to that observed for [SG], if one only relies on the dichotomy of the VOT highlighted above (in addition to other views highlighted in Ridouane, 2006). While this VOT dichotomy and the basic correlates identified in Jessen (2001) provide a clear cut separation between the two types of languages, it is clear this is far from ideal. The shared non-basic correlates highlighted by Jessen (2001) are important to look at further; they seem to follow what Kingston and Diehl (1994, see Hawkins, 2010 for an elegant reinterpretation) have also identified. However, in both cases, the shared non-basic features are assigned to different phonological features [Voice] or [Tense] for Jessen, which was meant to be used to describe any Voicing contrast, and [+Voice] or [−Voice] for Kingston and Diehl, which was primarily identified to account for the voicing contrast in English. Following Jessen (2001); Kingston and Diehl (1994), it is interesting to note that the [Voice] (for Jessen) or [+Voice] (for Kingston and Diehl) set has a lower F1 and *f* o, presence of voicing in the closure, reduced closure duration, longer preceding vowel and changes in voice quality correlates. The [Tense] (for Jessen) or [−Voice] (for Kingston and Diehl) set shows overall reverse patterns, with a higher F1 and *f* o, absence of voicing in the closure, increased closure duration, shorter preceding vowel and changes in voice quality correlates.

Clearly these primary and secondary correlates work together to enhance the features [Voice] and [Tense]/[SG] in both Voicing and Aspirating languages and their combinations allow for the features to become active in one or the other language. In the following sections, we provide three different studies that looked at the primary and secondary correlates in the Voicing contrast within Arabic. The first looks at the interaction between Voicing and Gemination in Lebanese Arabic, where we provide empirical evidence for a possible 4-way Voicing contrast in Arabic (Section 4.4.2). We then look at two subsequent studies by PhD studies who worked with me on their topics (Section 4.4.3), by examining how the Voicing contrast is implemented in two Saudi

Arabic dialects that differ in their feature specifications, especially for the Voiceless set. These three studies put forward the hypothesis that Arabic, traditionally classified as a Voicing Language is overspecified and using a binary [±Voice] feature is not enough to account for the minute differences in both Voiced and Voiceless sets; additional features, e.g., [SG] and [Tense] account better for the contrast. In the three studies, we adopt a numeric Privative feature account following J. Beckman et al. (2013), who follow Chomsky and Halle (1968) by assuming that *"...features are privative and stops in a true voice language are either specified as [voice] or not specified for a laryngeal feature, and in an aspirating language, stops are specified as [sg] or not specified for a laryngeal feature. This is in the phonology. Let us assume, further, as has often been suggested, that at some level prior to the phonetics privative features are transformed into numerically specified features (see Chomsky & Halle 1968), and that every segment has to have a positive numerical specification for the feature that is active in that language – but not for any feature that is not active."* (J. Beckman et al., 2013, pp. 279).

It is important to note that J. Beckman et al. (2013) consider numerically specified phonological features to operate before application of any phonological rules and this is not a simple surface level notation. This means that these numeric specifications of active features are part of the grammar and the mental representation of the speaker. We take this forward in our investigation here.

4.4.2 A 4-way contrast in Arabic?

As highlighted previously (Sections 4.3), the feature [+Tense] associated with geminate consonants has an important role in Arabic Phonology. In Section 4.3.3, we presented the results of the voicing patterns in geminate fricative consonants (following J. Al-Tamimi & Khattab, 2015). The results presented in Figure 4.16d (pp. 73) highlighted that geminate fricatives were produced with less voicing frames than singletons, which were identified to be similar to the patterns found in tense (fortis) categories (Jaeger, 1983).

The debate surrounding whether gemination is considered to be primarily temporal with secondary consequences due to the longer duration of the consonant or whether it is the reverse, due to increase in air pressure and changes associated with stronger production that geminates are produced as long has already been presented in the Introduction (see Section 4.3.1, pp. 68). We explored this topic further as shown in (J. Al-Tamimi & Khattab, 2018a, BAAP[9]; Khattab & Al-Tamimi, 2018, 16th LabPhon; J. Al-Tamimi & Khattab, 2018b, Journal of Phonetics[10]).

While the VOT was seen as a primary correlate for the Voicing contrast across languages, we have seen how additional correlates can contribute to defining this contrast. When looking at languages with 3 or 4-way contrasts, it is clear that the VOT fails to account for these and secondary correlates become promoted to primary places, e.g., the case of Korean with a 3-way contrast that does not rely solely on the VOT (Kim, Beddor, & Horrocks, 2002; Lisker & Abramson, 1964; Shimizu, 1989).

We saw previously (Section4.3.3; see J. Al-Tamimi & Khattab, 2015) that gemination in many instances is described in terms of articulatory strength, whereby long consonants are better

[9]This work received the Peter Ladefoged prize for the presentation most in the spirit of the work of the late Peter Ladefoged, British Association for Academic Phoneticians, 2018.

[10]This work was selected as an invited manuscript for the special issue of Journal of Phonetics,"Marking 50 Years of Research on Voice Onset Time and the Voicing Contrast in the World's Languages", eds., T. Cho, G. Docherty & D. Whalen

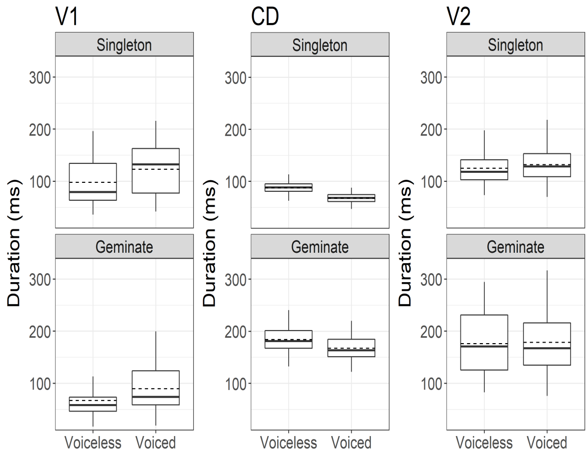
described in terms of tenseness rather than simply by temporal differences as articulatory strength leads to longer productions (Catford, 1977; Jaeger, 1983; Nellis & Hollenbach, 1980, but see, Abramson, 1986; Ridouane, 2010 for an alternative view).

Looking at how gemination and voicing interact with each other, it is clear that a special account needs to be identified here. As highlighted above, there are similarities between the two approaches identified in Jessen (2001) and in Kingston and Diehl (1994) in the fact that secondary (or shared non-basic) correlates play a major role in distinguishing the Voicing contrast across the world's languages (see descriptions and Table 1, pp. 270 in Jessen, 2001). Clearly, the case of geminates is treated differently; it is assigned the feature [Tense] following Jessen (2001). When dealing with the Voicing contrast in geminates, Jessen (2001) assumes that the Voiceless geminate is assigned the features [+Tense, −Voice], whereas the Voiced geminate is assigned [+Tense, +Voice]. From a phonetic implementation point of view, this will create a conundrum: If the geminates are [+Tense], the expectations are that in the Voiceless geminate, we should see similar patterns to those observed for "Fortis" consonants in Aspirating languages. For the Voiced geminate, this becomes more complex, given that they are typologically rare. The two features [+Tense] and [+Voice] will be acting in opposite direction given that the phonetic correlates for the former are different from those for the latter (following any of the views advanced in Kingston & Diehl, 1994, or Jessen, 2001). We then ask the following questions: which acoustic features will allow the Voiced set to be distinguished from any other category? Can we assume that the Voiced geminate will be more similar to the "lenis" category (given the [+Voice] feature)? We know from the literature that Voiced stops are typically described as having lax or lenis properties, and often undergo shortening in true Voicing languages in order to maintain voicing (Jansen, 2004; Jessen, 2001; Kohler, 1984). If this were the case, can they be produced as both long and short at the same time? If they are not produced as both short and long, can they be produced with both voicing and devoicing at the same time? How would the two features [+Tense] and [+Voice] interact with each other?

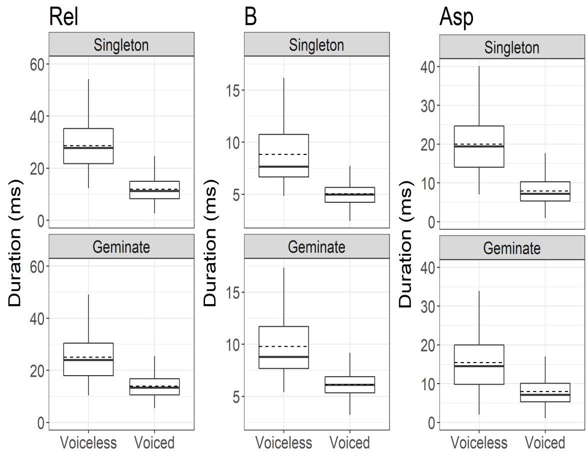
Given that the Voicing contrast can be implemented variably across the world's languages and the situation can become more complex when dealing with gemination, we explored the interaction between voicing and gemination to identify if this can lead to a potential four-way contrast. J. Al-Tamimi and Khattab (2018b) explored the interaction between the two by providing a systematic account of changes happening within the four categories: Voiceless Singleton, Voiced Singleton, Voiceless Geminate and Voiced Geminate, using the five syllable structures presented previously (see items 1) to 5) from the List 2 on pp. 68) specifically in the stop category. We used 19 different acoustic correlates that were identified in the previous literature as potential correlates for the Voicing and/or for the Gemination contrasts. These correlates spanned both temporal (e.g., closure duration, %Voicing, VOT, durations of preceding/following vowels or the release duration[11]) and non-temporal (e.g., F1, *f*0 and *H*1*-*H*2* at offset of V1 or onset of V2 and intensity dB in various positions) correlates.

Figure 4.20 presents the results within the temporal domain, with the durations within the VCV sequence (Figure 4.20a) and within the release phase (Figure4.20b), in addition to the voicing patterns (Figure 4.20c) using both the traditional VOT and the %Voicing. Overall, as

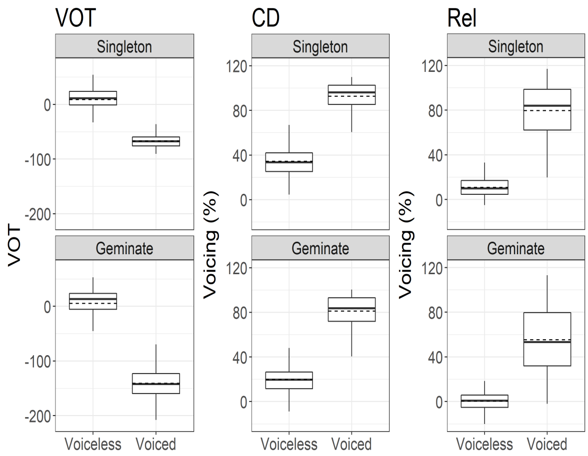
[11]We report on the VOT using the 50% threshold recommended by Abramson and Whalen (2017) to identify the Voicing patterns as positive or negative voicing. In additon, the release phase was segmented into full release, Burst and Aspiration phases to account for any specific impacts of gemination on these portions.



(a) Durations within V1C2V2



(b) Durations within the release



(c) Voicing patterns

Figure 4.20: Results of the 4-way Voicing contrast in Lebanese Arabic: (a) Durations of V1, C2 or V2; (b) Durations of the Release, Burst and Aspiration phases; (c) Voicing patterns with the classical VOT, %Voicing in the closure and the Release; CD = Closure Duration, V1 = Preceding Vowel, V2 = Following Vowel, Rel = Release phase, B = Burst phase, Asp = Aspiration phase, VOT = Voice Onset Time using the 50% threshold, Dotted lines within each boxplot represents the mean (reproduced from J. Al-Tamimi & Khattab, 2018b, see Notebook 1 available at `https://jalalal-tamimi.github.io/R-Voicing-Gemination-VOT/`)

shown in Figure 4.20a in a Voiced stop environment, the preceding vowel (V1) is lengthened, the Closure Duration (CD) is shortened, with no clear differences within the vowel following the medial consonant (V2), and this is regardless of the singleton or geminate status of the C2. Geminates trigger a shorter V1, longer C2 and V2, regardless of the Voicing status, confirming our previous findings in fricatives (J. Al-Tamimi & Khattab, 2015). Moving on to the release phase (taken as a correlate for positive VOT), Figure 4.20b shows that the duration of Rel, Burst or Aspiration phases is shorter in a Voiced context, regardless of the singleton/geminate status of C2. Interestingly, the Rel phase for the Voiceless category shows that the Singleton C2 has an average duration of 30 ms. Geminates trigger an overall shorter Rel and Asp portions, but a longer Burst, due to presence of multiple burst contacts identified in this specific category. Finally, when considering the Voicing patterns in the data, Figure 4.20c shows the traditional

VOT measure (positive or negative VOT), in addition to that of the %Voicing either in CD or the Rel phases. Interestingly, The traditional VOT shows that the Voiceless set has a marginal positive VOT, whereas the Voiced set has variable negative VOT that is conditioned by the singleton/geminate status of the C2. Within the Voiceless set, the traditional VOT failed to distinguish between the Singleton and Geminate as the difference was of a marginal 3ms (p=0.064); within the Voiced set, geminates showed an increase in the negative VOT by an average of -74ms (p<0.0001) that is influenced by its longer CD. Moving on to the %Voicing, it is interesting to note a clear interaction between the Voicing and Gemination status of the C2. The Voiced set shows an overall increase in %Voicing that is more important in the Singleton (92.64%) than in the Geminate (81.4%). However, differences emerged within the Voiceless set, as we show an overall increase in Voicing lead (or passive voicing) in the Singleton (34.35%) in comparison to the geminate (19.90%). These patterns in the Voiced and Voiceless sets are not captured by the traditional VOT as using the 50% threshold of voicing in the closure leads to ignoring any voicing lead in the Voiceless set. Finally, the Rel phase shows an increase in %Voicing in the Voiced sets that is more important in the Singleton (79.56%) than in the Geminate (55.34%). Within the Voiceless set, we observe some Voicing lead in the Singleton (10.84%) in comparison to the Geminate (0.42%).

As shown in J. Al-Tamimi and Khattab (2018b, pp.316-318), the other spectral correlates show differences related to the Voicing contrast. Given that the Geminate consonant has a longer CD, it leads to a reduction of V1 duration, with less Voicing frames (via %Voicing), which leads to a reduction in F1 frequency, f0 patterns and the intensity. These correlates are similar to those identified previously as being the primary and secondary correlates for the "Fortis" set in Aspirating languages (Jessen, 2001; Kingston & Diehl, 1994), leading to Arabic to be identified in a mixed category on the continuum between Voicing and Aspirating languages.

Table 4.2: Confusion matrices in percentages of classification results from the three Random Forests (Models A, B and C), with model accuracy next to each model, and with the prediction in rows and original data in columns. Vd = Voiced; Vl = Voiceless, S = Singleton, G = Geminate (reproduced from J. Al-Tamimi & Khattab, 2018b, pp. 319)

	Model A (92.5%)				Model B (82.3%)				Model C (67.2%)			
	Vd-S	Vl-S	Vd-G	Vl-G	Vd-S	Vl-S	Vd-G	Vl-G	Vd-S	Vl-S	Vd-G	Vl-G
Vd-S	**92.6**	8.6	0	0	**91.9**	8.6	1.4	0	**67.6**	6.7	23.1	0
Vl-S	7.4	**89.5**	0	1.5	6.8	**57.1**	3.3	15	4.7	**51.4**	4.7	16.5
Vd-G	0	1.9	**93.9**	6	0.7	1.9	**90.1**	6	26.4	9.5	**67.9**	5.3
Vl-G	0	0	6.1	**92.5**	0.7	32.4	5.2	**78.9**	1.4	32.5	4.2	**78.2**
Total	100	100	100	100	100	100	100	100	100	100	100	100

To make sense of the complex picture, we used Random Forests grown via the Conditional Inference Trees framework (Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008; Strobl, Boulesteix, Zeileis, & Hothorn, 2007; Strobl, Hothorn, & Zeileis, 2009; Strobl, Malley, & Tutz, 2009), with the four Voicing by Gemination categories: Voiced Singleton, Voiceless Singleton, Voiced Geminate and Voiceless Geminate. We used this version of Random Forests as it allows to compute Variable Importance scores conditioned by the correlation matrices between heavily correlated predictors. Table 4.2 shows the % correct classification accuracy next to each model

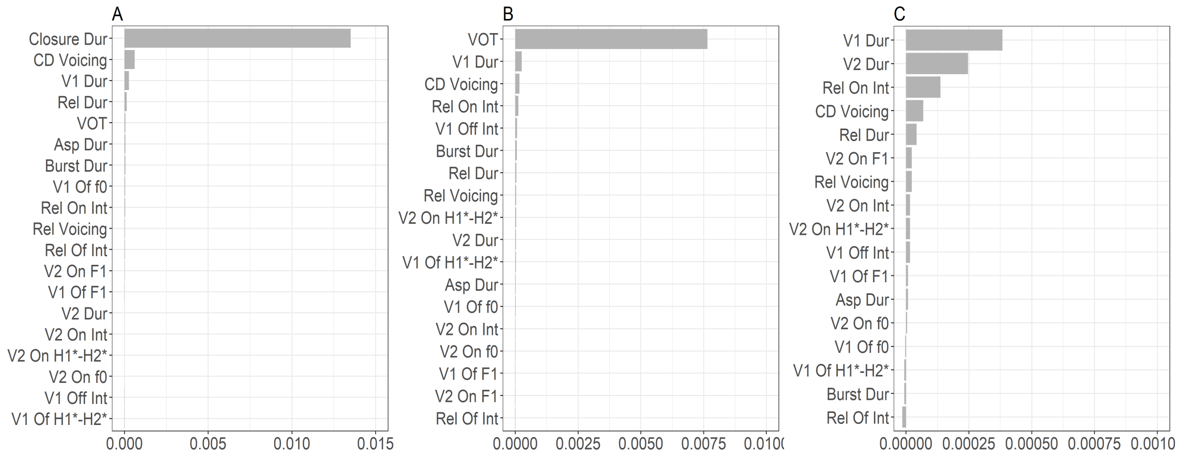


Figure 4.21: Variable Importance scores of the Random Forests grown via the Conditional Inference Trees framework, with Conditional Permutation tests (reproduced from J. Al-Tamimi & Khattab, 2018b, pp. 319; see Notebook 2 available at https://jalalal-tamimi.github.io/R-Voicing-Gemination-VOT/)

and the confusion matrix for each Random Forest. Figure 4.21 shows the Variable Importance scores as a mean decrease in accuracy for three models: Model A contains all 19 predictors used in this experiment; B uses 18 predictors after removing the CD; C uses 17 predictors after removing both CD and VOT.

First, it is clear that Model A performs exceptionally well as it achieves a rate of 92.5% (with an R^2=0.964). The confusions are mostly within the Singleton or the Geminate categories, regardless of the Voicing status (Table 4.2; Model A). Variable Importance scores point to the CD as being the most influential predictor (Figure 4.21; A), followed by %Voicing in CD, V1 and Rel Durations, then the VOT. Due to the conditional nature of this version of Random Forests, the most influential predictor is highlighted (see Notebook 2 available at https://jalalal-tamimi.github.io/R-Voicing-Gemination-VOT/, for the unconditional Variable Importance scores, which shows CD as the most influential followed by the same predictors, but in different orders). Clearly, the traditional VOT is confusable with the CD, hence it is identified as less important than %Voicing in the Closure. Model B reports on all predictors without the CD. It achieves a rate of 82.3% (with an R^2=0.802) at separating the four categories. Interestingly, the two Voiced categories (Table 4.2; Model B) show similar types of confusions as in Model A, but the Voiced Geminate show variable confusions across all remaining categories. The Voiceless sets are clearly interesting. First, the Voiceless Singleton (classified correctly as such in only 57.1% of the cases) is more confusable with the Voiceless Geminate (at 32.4%); the Voiceless Geminate (classified correctly as such in 78.9% of the cases) is confused with the Voiceless Singleton (at 15%). Clearly, these results show that when not using the CD, there are similarities between the two Voiceless categories, which led to more confusability. The Variable Importance scores (Figure 4.21; B) now show the traditional VOT to emerge as the most influential predictor, followed by the V1 duration and %Voicing. Lastly, Model C reports the results of all remaining predictors after removing the CD and the VOT. It is clear that this model is not optimal as it achieves a 67.2% (with an R^2=0.561), indicating that all additional correlates are secondary in nature. The confusion matrix (Table 4.2; Model C) shows most of the confusions are within the Voiced or the Voiceless sets rather than within the

Singleton or the Geminates. Variable Importance scores (Figure 4.21; C) point to the V1 and V2 duration, Intensity of the Release, %Voicing in CD and Release Duration as the most influential predictors. It should be noted that these are widely reported as correlating with the Voicing contrast, as highlighted in both Jessen (2001) and Kingston and Diehl (1994).

The results of our Random Forests clearly show how each of the primary and secondary correlates play a role in differentiating this complex 4-way contrast. Using individual predictors' results and those combined in our Random Forests, we proposed a phonological analysis informed by J. Beckman et al. (2013)'s privative feature with numerical notation. Table 4.3 shows two canonical examples from Russian and German representing respectively a Voicing and an Aspirating language. Following this privative account, active features are assigned a numeric notation ranging between 1 and 9 (1 = feature lacking specification the phonology; 3 = limit to allow a feature to become active; 9 = feature maximally active), whereas inactive features are assigned a [∅]. Based on the results from previous research, J. Beckman et al. (2013) present the privative feature specification for a "true" Voicing or a "true" Aspirating language (see Table 4.3, top row). It should be noted that this feature specification is based primarily on the VOT and voicing patterns in these languages (and others cited in the paper on Thai, Icelandic, English, Hindi, etc.).

Table 4.3: Privative feature specification in Voicing (Russian) and Aspirating (German) languages (Top), and in Lebanese Arabic (bottom). Code for numeric notation: 1 = feature lacking specification the phonology; 3 = limit to allow a feature to become active; 9 = feature maximally active; ∅ = non-active feature; SG = Spread Glottis (reproduced from J. Al-Tamimi & Khattab, 2018b, pp. 320, in addition to J. Al-Tamimi & Khattab, 2018a; Khattab & Al-Tamimi, 2018)

Voicing languages		**Aspirating languages**	
	Russian		**German**
Fortis:	[1Voice] [∅SG]	Fortis:	[9SG] [∅Voice]
Lenis:	[9Voice] [∅SG]	Lenis:	[1SG] [∅Voice]
	Arabic		
	Singleton stops		**Geminates stops**
Fortis:	[3Voice] [∅SG] [2Tense]	Fortis:	[1Voice] [∅SG] [4Tense]
Lenis:	[8Voice] [∅SG] [0Tense]	Lenis:	[6Voice] [∅SG] [3Tense]

Employing a similar approach in our case here was less than straightforward given that we were not only interested in the implementation of the Voicing contrast in general, but in interaction with Gemination. Based on the results presented in our studies, we have shown that both Voiced and Voiceless sets in Singleton or Geminate stops are variably produced. They seem to share acoustic correlates identified for the [Voice]/[+Voice] and for [Tense]/[−Voice] (following Jessen, 2001; Kingston & Diehl, 1994).

Our results highlighted above, especially for the Rel phase (see Figure 4.20b), the range of positive VOT in Lebanese Arabic, ranging between 12-14 ms for the Voiced Singleton-Geminate, and 29-25 for the Voiceless Singleton-Geminate, respectively. These figures are not high enough to justify the feature [SG] to be active in this dialect; hence it is assigned [∅SG]. Next, the interaction between the features [Tense] and [Voice] impact variably the 4-way contrast, we present in Table 4.3 (bottom row) the case for Lebanese Arabic. First, the feature [Tense] is active in this dialect, especially in the case of Geminates and signalled phonetically by variable secondary, rather than primary, correlates. Hence, we assign it a numeric notation of 3 or 4. For the Singletons, we claim that [Tense] is marginally active in the Voiceless Singleton (with [2Tense]) and that it is either lacking specification in the phonology or with no major impact in the Voiced Singleton (with [0Tense]). The feature [Voice] on the other hand, is actively present in all cases, with a near maximal presence in the Voiced Singleton (with [8Voice]) and with an intermediate presence in the Voiced Geminate (with [6Voice]). Contra to its specification for Russian, the Voiceless sets in Lebanese Arabic do allow some passive voicing, possibly due to the medial VCV sequence; hence receiving either [3Voice] for the Singleton or [1Voice] for the Geminate.

These numeric notation show gradiency in the phonological feature specification of this language that are important to consider in the wider context. In our case, we were able to quantify these minute differences after identify primary and secondary acoustic correlates of this complex case. The feature [Tense] is important in Arabic and our results on both fricatives (see Section 4.3.3) and stops (here) confirm this. It has a major role in the phonology of the language as it is actively used by adults but also by children acquiring gemination (see Section 4.3.4). Importantly, and as highlighted in J. Beckman et al. (2013), these numeric notations are not to be seen as simple surface application of rules; they are integral to the phonology following the spirit of Chomsky and Halle (1968).

In the next section, we look at two extensions to the Voicing contrast in Saudi Arabic following the Laryngeal Realism and the privative feature account.

4.4.3 The voicing profile across Arabic dialects

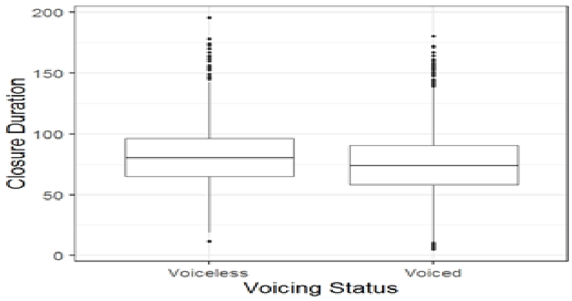
As highlighted above, the privative feature account within Laryngeal Realism puts forward the hypothesis that the numeric notation are integral part of the phonology of a language, and this can be used to identify which features are active or not. In what follows, I present some of the work done by PhD students who worked with me on their topics, first on the privative feature specification in Najdi Arabic where the features [Voice] and [SG] are both active (Section 4.4.3.1) and a second on Jazani Arabic where the features [Voice] and [Tense] are both active (Section 4.4.3.2).

4.4.3.1 [+Voice] and [+Spread Glottis] in Najdi Arabic

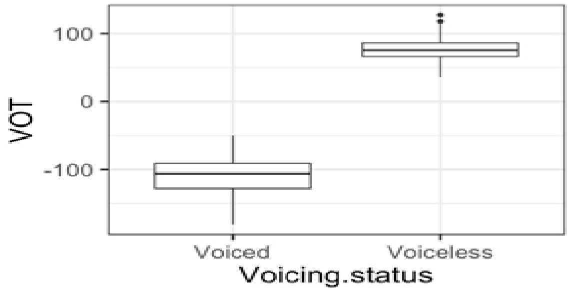
Al-Gamdi, Al-Tamimi, and Khattab (2019, 19^{th}ICPhS, see also Al-Gamdi, Khattab, & Al-Tamimi, 2018, BAAP)[12] quantified the differences between the Voiced and Voiceless stops in Najdi Arabic produced in word-initial position in a carrier sentence, by 12 speakers (6 males and 6 females). Using multiple acoustic correlates allowed to differentiate the two categories. Figure 4.22 shows the results of the Closure Duration (Figure 4.22a), the classical VOT (Figure 4.22b) and *f*0 patterns at the onset of the vowel following the stop (Figure 4.22c). The Voiced stops showed a systematic decrease in Closure duration, in VOT and in *f*0. It is clear that the Voiced set shows a clear negative VOT (i.e., prevoicing), whereas the Voiceless set showed a clear positive VOT. The *f*0 patterns correlated well with the differences reported in both Jessen (2001); Kingston and Diehl (1994) and correlate well with the two features [Voice] and [SG]. The

[12] This work was part of Nief Al-Gamdi's IPhD in Phonetics and Phonology summer project, equivalent to an master thesis

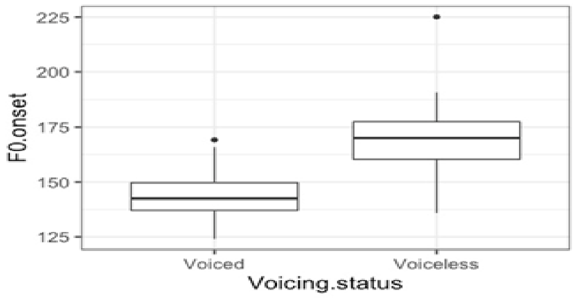
conclusions advanced in Al-Gamdi et al. (2019) is that Najdi Arabic seems to be overspecified with two features, following the privative features account: The Voiced set is clearly assigned the features [Voice, ∅SG], while the Voiceless set has [SG, ∅Voice]. The two features are active in this dialect; if only [±Voice] is active, then [−Voice] assigned to the Voiceless set cannot explain the increase in aspiration (positive VOT) in this dialect.



(a) Closure Duration



(b) VOT duration



(c) Fo patterns

Figure 4.22: Voicing patterns in Najdi Arabic initial stops: (a) Closure duration; (b) VOT; (c) Fo patterns (reproduced from Al-Gamdi et al., 2019)

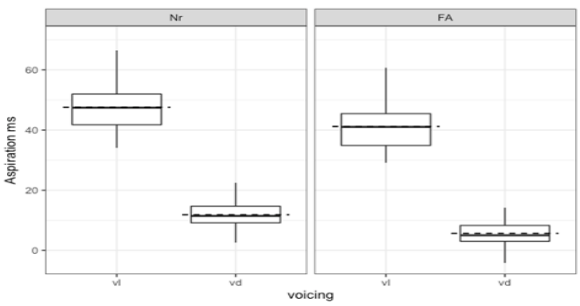
While Al-Gamdi et al. (2019) showed a clear pattern in this dialect, it only looked at stops in word-initial but utterance medial position (due to using a carrier sentence). As a follow up to this study and in his PhD, Al-Gamdi (2022, see also, Al-Gamdi, Al-Tamimi, & Khattab, 2022a BAAP; Al-Gamdi, Al-Tamimi, & Khattab, 2022b 18th LabPhon; Al-Gamdi, Al-Tamimi, & Khattab, Under Review Laboratory Phonology) examined the Voicing profile in Najdi Arabic, by looking at Voiced and Voiceless stops in word-initial utterance-initial, word-medial and word-final positions, as produced by 32 monolingual native speakers of the Najdi dialect (15 males, 17 females)[13]. The aim was to assess how Voicing is implemented in this dialect, by looking at various positions, places of articulation, vowel contexts and in both fast and normal speech rates[14]. Results presented in Figure 4.23 show a snapshot of the multitude of acoustic measures employed in this study[15]. Figure 4.23a shows how the aspiration duration across the Voiced and Voiceless stops is a robust acoustic cue as it is stable across speech rate; Voiceless stops show on average 48 msec in Normal speech rate, that drops to 40 in Fast speech rate. The Voiced set shows a minor aspiration duration of around 12 msec (7.5 msec for the fast speech rate). Looking at the

[13]A total of 40 participants were recorded, but due to data loss and other technical issues, data from 8 participants were discarded; see Al-Gamdi (2022) for more details

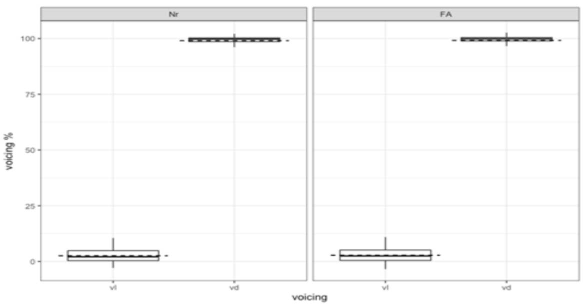
[14]Al-Gamdi (2022) also looked at Cluster-assimilation to validate the Voicing profile in the dialect

[15]There were close to 20 acoustic measures that varied in terms of temporal and non-temporal and across positions, see Al-Gamdi (2022, Table 4.2, pp.79)

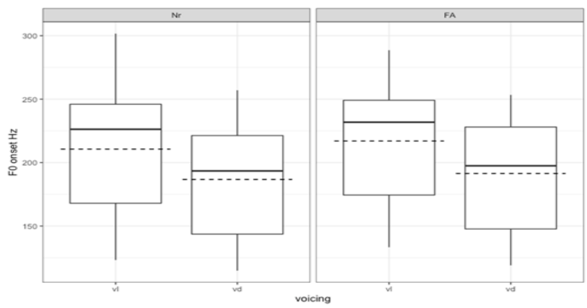
%Voicing, Figure 4.23b shows a dichotomous difference as the Voiced set is clearly prevoiced with %Voicing clearly showing 100% voicing, whereas the Voiceless set shows a robust voicelessness; both are consistent across the two speech rates. Finally, Figure 4.23c shows an overall consistent increase in $f0$ profile obtained at the onset of the vowel following the stop in the Voiceless set across the two speech rates, where there is a marginal increase in the fast speech rate in the Voiceless set in comparison to the normal speech rate. The results of this study are again in favour of describing the Voicing profile in this dialect as overspecified due to the two privative features being active.



(a) Aspiration Durations



(b) %Voicing



(c) F0 patterns

Figure 4.23: Voicing patterns in Najdi Arabic Utterance Initial stops: (a) Aspiration duration; (b) %Voicing; (c) F0 patterns Nr = Normal rate, FA = Fast speech rate(reproduced from Al-Gamdi et al., Under Review, see Al-Gamdi et al., 2022a, 2022b)

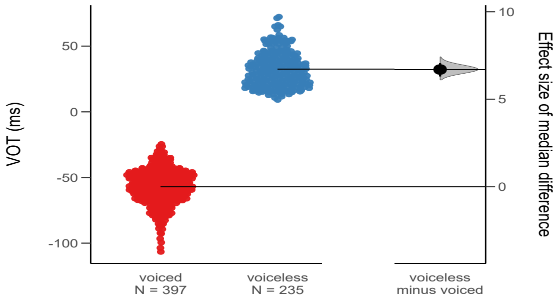
Najdi Arabic seems to be overspecified with two features, following the privative features account: The Voiced set is clearly assigned the features [9Voice, ∅SG], while the Voiceless set has [8SG, ∅Voice]. The numeric notations follow from J. Beckman et al. (2013)'s account and are explained in terms of a robust active Voicing in the Voiced set, especially at the utterance-initial but also medial positions, whereas the Voiceless set is clearly showing a relatively long aspiration duration that is comparable in length to other Najdi, Ghamidi, North Saudi and Qatari Arabic (see Al-Gamdi, 2022, Table 9.2, pp. 226), but not to Lebanese Arabic (J. Al-Tamimi & Khattab, 2018b).

4.4.3.2 [+Voice] and [+Tense] in Jazani Arabic

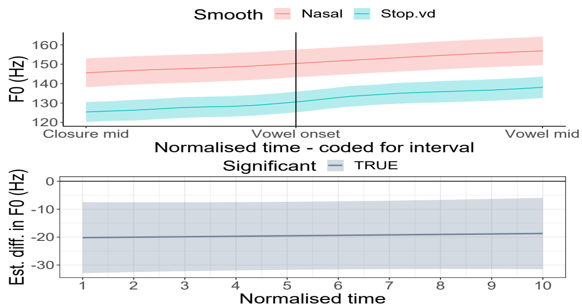
As a second subsequent study, Dallak (2023, see also, Dallak, Khattab, & Al-Tamimi, 2023, 20th ICPhS) in his PhD study under my supervision[16] continued to explore the Voicing profile in another Saudi Arabic dialect: Jazani Arabic, situated in the south-east of Saudi Arabia. This

[16] Part of a continued supervision agreement with Newcastle University, UK.

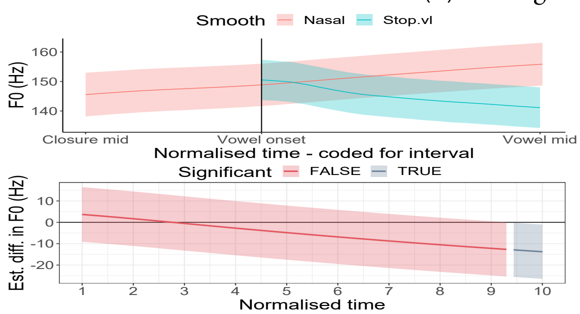
dialect differs from Najdi Arabic in that it is closer to an urban dialect. Dallak explored the Voicing profile as produced by 20 male speakers of the dialect producing the Voiced and Voiceless sets in various places of articulation and vowel contexts (restricted to the long vowels /iː aː uː/). While J. Al-Tamimi and Khattab (2018b) on Lebanese Arabic and Al-Gamdi (2022) on Najdi Arabic identified clear cross-dialectal differences in how the Voicing contrast is implemented in the light of the privative feature account embedded within Laryngeal Realism, they both lacked a direct comparison with a nasal baseline that was previously reported in Ladd and Schmid (2018).



(a) Voicing differences



(b) Voicing - Nasal vs Voiced



(c) Voicing - Nasal vs Voiceless

Figure 4.24: Voicing patterns in Jazani Arabic: (a) Estimation plot voicing differences; (b) Estimated F0 splines for nasal baseline and Voiced stops with difference smooth; (c) Estimated F0 splines for nasal baseline and Voiceless stops with difference smooth (reproduced from Dallak et al., 2023)

Dallak et al. (2023, following Dallak, 2023) explored this further by examining how the VOT and *f*0 perturbations in comparison with a nasal baseline differ in both Voiced and Voiceless sets. Figure 4.24 presents the Voicing profile results in Jazani Arabic. Starting with the VOT, the results presented in Figure 4.24a show the estimated plot of the Voiced vs Voiceless stops where the former shows a clear negative VOT pattern (-75 ms), whereas the latter showed an average positive VOT (32.5 ms), where the estimated difference is of 90.5 ms, exhibiting a robust effect size of 6.7 (see secondary y-axis).

Moving on to the results of the *f*0 perturbation, it is clear that the Voiced set induced a significantly reduced *f*0 contour of -19.45 Hz averaged across all the intervals from the closure midpoint to the vowel midpoint, with a robust lowered pattern throughout the measures intervals (see Figure 4.24b). Figure 4.24c shows the comparison between the Voiceless set and the nasal baseline. It is clear that the Voiceless set induced a statistically non-significant pattern of -8.29 Hz averaged across all the intervals measured relative to the nasal baseline. There was a non-statistically significant rising pattern at the vowel onset, in comparison to the nasal baseline, but a clear falling transition from the onset to the midpoint of the vowel until reaching a significant difference at the vowel midpoint.

Dallak et al. (2023) reported original findings related to the Voicing profile in Jazani Arabic, looking at the interaction between the VOT and the *f*0 perturbations in comparison to the nasal baseline. From a feature representation point of view, Dallak et al. (2023) advocate that the feature [Tense] rather than [SG] is active on the Voiceless set in this dialect, alongside the feature [Voice] for the Voiced set. In his PhD, Dallak (2023) explored this further by comparing the Voicing profile differences across stops and fricatives in order to compare whether the Voiceless fricatives are assigned the feature [SG] as default feature or whether an alternative feature can account for differences related to manner of articulation. The study compared various acoustic metrics in both stops and fricatives and showed a difference in terms of the acoustic correlates of the Voicing profiles used in the dialect. Specifically and following Dallak (2023), there are two active privative features to account for the changes observed in the dialect: the features [Voice] and [Tense]. The privative features following J. Beckman et al. (2013)'s numerical notations are that the stops in Jazani Arabic are specified with [5tense] [1voice] for the Voiceless stops, but [2tense] [8voice] for the Voiced stops. For fricatives, the specification are as follows: [6tense] [2voice] for the Voiceless fricatives, but [1tense] [6voice] for the Voiced fricatives. Clearly, stops and fricatives in this dialect behave differently to those of Najdi Arabic. This is likely due to the sociolinguistic background of participants; an area of further exploration within Arabic dialects.

4.5 Epilaryngeal constriction in Arabic - The feature [+cet]

4.5.1 Introduction

This research theme emerged during my MA (Section 4.2.1.2) and PhD research (Section 4.2.2) and expanded on since 2014, where I was interested in the impact of pharyngealisation on surrounding vowels in both Jordanian and Moroccan Arabic (JA and MA, respectively). As we saw previously, the two dialects are different in terms of vowel distribution, and dispersion and in the type of dynamic cues used to quantify their vowels. Here, we wanted to assess any cross-dialectal differences in how pharyngealisation and the whole guttural class impact on vowels, to establish possible differences related to constriction location and type of correlates used.

Guttural consonants in Arabic have been the subject of debate in phonetics and phonology. Gutturals are assumed to form a natural class due to specific morpho-phonological patterning, and/or to the use of a *common* oro-sensory zone in the pharynx (J. McCarthy, 1994; Sylak-Glassman, 2014a). Traditionally, J. McCarthy (1994) claimed that the members of such a class are restricted to pharyngeals /ħ ʕ/ (or epilaryngeal /ʜ ʢ ʡ/) and uvulars (/χ ʁ q/) due to these consonants having primary constrictions within the pharynx. However, researchers on Arabic and other languages highlighted that Pharyngealised consonants (/tˤ dˤ ðˤ sˤ/) (Sylak-Glassman, 2014a), and/or Laryngeals (/ʔ h/) (Zawaydeh, 1999) are also considered to be part of this class, due to the fact that pharyngealised consonants share a similar place of articulation to that of pharyngeals, albeit with a different degree of constriction (J. Al-Tamimi, 2017b; Laufer & Baer, 1988) or due to the increase in the frequency of the first formant in the vowels surrounding gutturals when compared with plain coronals (Zawaydeh, 1999).

From a phonetic point of view, and as highlighted below (Sections 4.5.2 and 4.5.3), it is clear that researchers failed to identify a single common acoustic and/or articulatory correlate that unites all the members of gutturals as a natural class (Bin-Muqbil, 2006). As highlighted below, the majority of research on Arabic looked at supra-laryngeal changes related to guttural consonants, either using acoustic measures (e.g., F1, F2 and sometimes F3; F2-F1) or articulatory techniques, mostly invasive (e.g., X-Rays, Nasoendoscopy, Videofluoroscopic) with some using non-invasive techniques (e.g., Ultrasound Tongue Imaging, UTI; ElectroMagneticArticulatography, EMA; ElectroGlottoGraph, EGG). In most cases, combined techniques were used with both articulatory and acoustic measures restricted in the majority of cases to formant frequencies.

My research has focused on an alternative view to guttural consonants using a combined phonetic and phonological accounts. First by relying on the predictions of the "Laryngeal Articulator Model" (LAM) (Esling, 2005; Esling et al., 2019), which identifies the epilarynx as the main articulator used in the productions located within the Lower Vocal Tract. According to LAM, when the epilarynx is constricted, there are multiple articulatory (an acoustic) changes, including raising of the larynx with a concomitant retraction of the tongue root and dorsum in a *back and down gesture*. This combined gesture has been claimed to suit well the whole post-velar/guttural natural class (Sylak-Glassman, 2014a, 133). This led me to provide a complete acoustic account of laryngeal and supra-laryngeal changes due to pharyngealisation in two Arabic dialects. I was first informed by research done on languages with [±ATR] harmony (for "Advanced Tongue Root"; see Section 4.5.2 for more details), where the acoustic consequences of such constrictions are well-established: [−ATR] trigger an increase in F1, reduction in F2, reduction in F2-F1 and changes in spectral tilt inducing a more tense voice quality, in comparison to [+ATR] vowels. Then using results identified in singing (e.g., opera signing), allowed to evaluate how an epilaryngeal constriction leads to an enhanced and clearer voice quality due to an increased energy in the higher frequencies. Combining these approaches informed the results of how pharyngealisation operates in these two Arabic dialects.

Next, informed by the results obtained on pharyngealisation, I expanded my research to all guttural consonants to try and identify a clear pattern in the articulatory and acoustic domains, which could allow them to be categorised as belonging to the same natural class. This subsequent research used a combined articulatory and acoustic account using non-invasive UTI and EGG for the former and advanced acoustic analyses evaluating the acoustic correlates of laryngeal and supra-laryngeal changes, for the latter.

My research within this theme have tried to provide an empirical evidence for how supra-laryngeal and laryngeal activities within the pharyngealised set interact with each other to define this category and then for looking at quantifying similarities between guttural consonants providing empirical evidence supporting the legitimacy of gutturals as a natural class.

4.5.2 Pharyngealised coronal stops

The main body of research within this section falls within these publications: J. Al-Tamimi and Barkat-Defradas (2002, 5^{th} AIDA on cross-dialectal differences using static approaches); Barkat-Defradas et al. (2003, 15^{th} ICPhS on cross-dialectal differences using static approaches); J. Al-Tamimi (2004, 25^{th} JEP on using Locus Equations to distinguish place of articulation); J. Al-Tamimi (2007b, 16^{th} ICPhS on using static and dynamic cues in pharyngealised contexts); J. Al-Tamimi (2009, International Workshop on Pharyngeals and Pharyngealisation on using static and dynamic cues in pharyngealised contexts); J. Al-Tamimi (2014, BAAP on quantifying the role of spectral tilt in pharyngealisation); J. Al-Tamimi (2015, 18th ICPhS on quantifying the role of spectral tilt in pharyngealisation); J. Al-Tamimi (2016, 2nd Arabic Linguistic Forum, ALiF on quantifying the role of spectral tilt in pharyngealisation and epilaryngeal constriction); J. Al-Tamimi (2017a, 2nd Phonetics and Phonology in Europe, PaPE on quantifying the role of spectral tilt in pharyngealisation and epilaryngeal constrictions); J. Al-Tamimi (2017b, Article in Laboratory Phonology on quantifying the role of spectral tilt in pharyngealisation and epilaryngeal constrictions).

As will be highlighted below, the research from 2014 onwards presented an account of the role of laryngeal and supra-laryngeal activity in describing pharyngealisation.

4.5.2.1 Traditional account

Traditionally, pharyngealisation (or emphasis) in Arabic is a secondary articulation where a secondary pharyngeal constriction is superimposed on a primary constriction involving coronal consonants, which allows the distinction between non-pharyngealised /t d ð s/ and pharyngealised consonants /tˤ dˤ ðˤ sˤ/, although other consonants are considered as having a secondary pharyngealisation impact on surrounding vowels, e.g., /m b r l/ (Watson, 2007). It generally involves retraction of the tongue dorsum towards the upper pharyngeal area, which leads to a lowering of the second formant in the surrounding vowels (e.g., Bin-Muqbil, 2006; Ghazeli, 1977; Watson, 2007; Zawaydeh & de Jong, 2011). Although these characteristics are mostly agreed upon, pharyngealisation is also associated with a retracted epiglottis, a raised larynx, a pressed/tense voice quality, and/or a protruded lip posture (see e.g., F. Al-Tamimi & Heselwood, 2011; Cantineau, 1960; Hess, 1998; Laufer & Baer, 1988; Lehn, 1963; Zeroual & Clements, 2015; Zeroual, Esling, & Hoole, 2011, among others). It is clear that there are differences between "true" pharyngeals and pharyngealised consonants, where both share the same place but vary in the degree of constriction (Laufer & Baer, 1988). This led researchers to identify variable places of articulation for pharyngealised consonants in Arabic which varies from a (post-)velar to a (mid-)low pharyngeal (Khattab, Al-Tamimi, & Heselwood, 2006).

From an acoustic point of view, nearly all studies looked at formant frequencies as correlating with supra-laryngeal changes (J. Al-Tamimi, 2017b). A plethora of studies reported the primary *acoustic* correlate differentiating between the two categories to be lowering of the second formant (F2) in the vowel adjacent to the consonant, which correlated very well with *tongue dorsum retraction* reported in the majority of studies (Al-Ani, 1970; F. Al-Tamimi & Heselwood, 2011; J. Al-Tamimi & Barkat-Defradas, 2002; Barkat-Defradas et al., 2003; Ghazeli, 1977; Jongman, Herd, Al-Masri, Sereno, & Combest, 2011; Laufer & Baer, 1988; Zawaydeh & de Jong, 2011). This led researchers to identify the main phonological feature describing these patterns to be [+Back] and/or [+RTR] (for "Retracted Tongue Root", see Watson, 2007, among others).

Two apparent issues have been highlighted in J. Al-Tamimi (2017b). For the first (see Section 4.5.2.4 for the second), if the constriction location is somewhere in the pharyngeal area, then there are expectations for changes to occur beyond the F2. Indeed, a few researchers explored the contribution of the "pharyngeal" component of pharyngealisation, by reporting an inconsistently higher first formant (F1) in the pharyngealised context (F. Al-Tamimi & Heselwood, 2011; J. Al-Tamimi & Barkat-Defradas, 2002; Barkat-Defradas et al., 2003; Ghazeli, 1977; Jongman et al., 2011; Khattab et al., 2006; Laufer & Baer, 1988). Fewer researchers reported on changes associated with the third formant (F3), which was shown to be variable and to depend on the vowel, with a higher F3 in back vowels, e.g., /uː/, potentially reflecting an upper-pharyngeal constriction, and a lower F3 with front vowels, e.g., /iː/, potentially reflecting a mid-pharyngeal constriction, with no differences in /aː/ (F. Al-Tamimi & Heselwood, 2011; J. Al-Tamimi, 2007a, 2007b; Jongman et al., 2011; Norlin, 1987; Zeroual et al., 2011).

The findings reported in the literature go hand in hand with the predictions from the "Acoustic Theory of Speech Production" with regards to the acoustic-to-articulatory mapping. For instance, a pharyngeal constriction (towards the middle/low part of the pharynx) is known to induce the following acoustic changes: ↑F1, ↓F2, ↓F3; when the constriction is located higher-up (i.e., upper pharyngeal/uvular), the following patterns are expected: marginal ↑F1, ↓F2, ↑F3 (Carré & Mrayati, 1992; Chiba & Kajiyama, 1941; Mrayati, Carré, & Guérin, 1988; Stevens, 1989).

4.5.2.2 Epilaryngeal Constriction

The "Laryngeal Articulator Model" (LAM) was initially developed by Esling (2005), extensively modelled by Moisik (2013a), expanded on in Esling et al. (2019), and theoretically and typologically evaluated in Sylak-Glassman (2014a). The Lower Vocal Tract is "is bounded inferiorly by the glottis and superiorly by the oropharyngeal isthmus and velo-pharyngeal port." (Moisik, 2013a, p. 84). It is located within the Lower Vocal Tract above the larynx and has the ventricular folds as its lower part and the rim formed by the epiglottis and aryepiglottic folds as its upper part (Moisik, Czaykowska-Higgins, & Esling, 2012). Following LAM, we postulated the hypothesis that pharyngeals in general and potentially pharyngealised consonants in particular are produced by sphincterally constricting the epilarynx through constricting the intrinsic and/or the extrinsic laryngeal muscles; tongue retraction seen in both cases is caused by this constriction and is seen as a facilitator and an enhancer of the pharyngeal articulation (Esling, 2005; Moisik, 2013a; Sylak-Glassman, 2014a, 2014b). In fact, LAM predicts that constriction of the hyoglossus muscle draws the tongue as a whole backward and downward, which leads to retraction of the tongue root and dorsum (Moisik, 2013a, pp. 372-373; Sylak-Glassman, 2014b, pp. 5). This leads to vowels influenced by an "epilaryngeal" constriction to be produced as "retracted" with a *back and down* gesture.

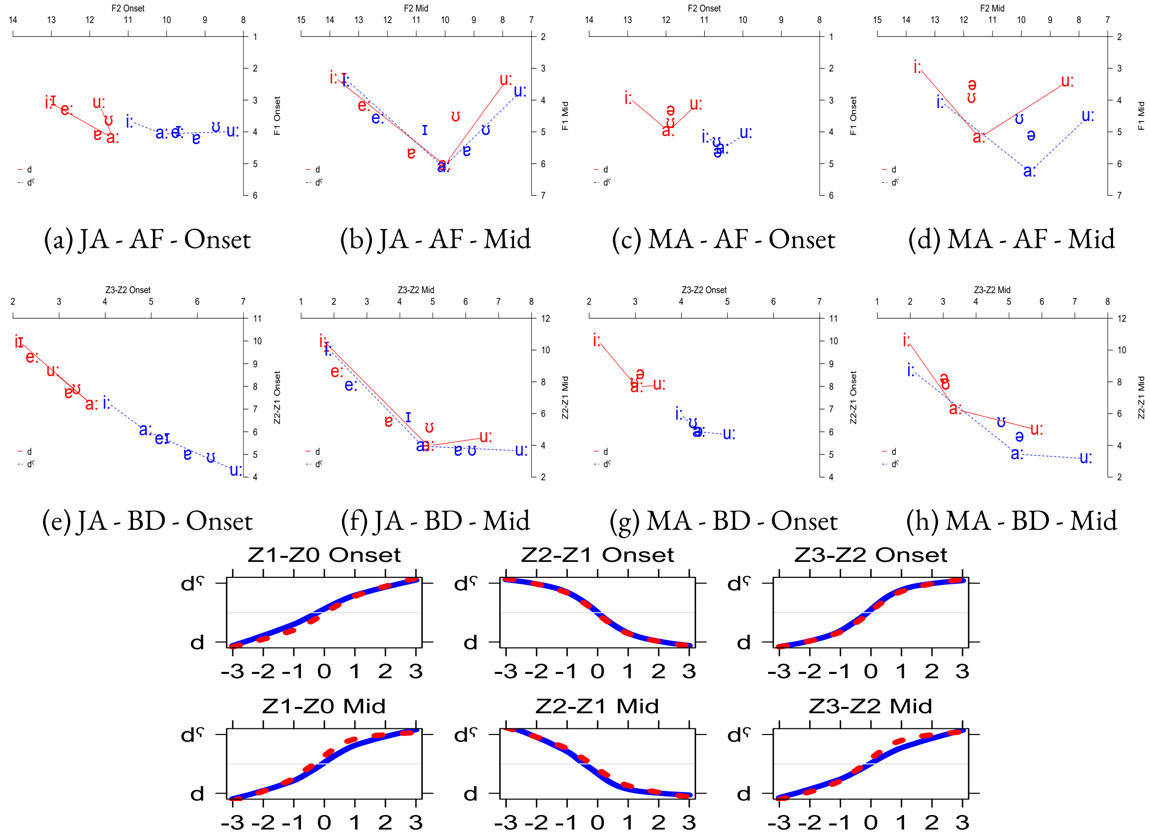
In addition, the epilaryngeal constrictor mechanism leads to voice quality changes, due primarily to a raised larynx posture that accompanies an extreme epilaryngeal constriction and secondarily due to two actions: 1) ventricular folds are coupled with the vocal folds and are brought down to the glottis to allow for creaky phonation to occur (Moisik, 2013a; Sylak-Glassman, 2014a, 2014b) and 2) when ventricular folds are constricted, they lead to laryngealisation, tense and/or harsh voice qualities (Edmondson & Esling, 2006; Stevens, 1977). Aryepiglottic fold constriction is associated with either trilling observed in "true" pharyngeals, which lead to them being actually named as epilaryngeal consonants rather than pharyngeals within LAM (Esling, 1999, 2005; Moisik, Esling, & Crevier-Buchman, 2010), or with an enhanced and clearer voice quality especially in singing due to an increased energy in the higher frequencies as well as 'ringing' seen in specific types of singing (Story, 2019; Titze & Story, 1997). J. Al-Tamimi (2017b) hypothesised that there will be a clear difference between true "pharyngeals" and pharyngealised consonants, following LAM: the former are produced by constricting the aryepiglottic folds as a primary feature, whereas the latter by constricting the ventricular folds as a secondary feature. There are of course additional minute changes that will be discussed later on.

4.5.2.3 Supra-laryngeal changes

The main claim advanced in J. Al-Tamimi (2017b) is that previous literature failed to identify the specific patterns related to impact of pharyngealisation due to two aspects: 1) looking mostly at the second (F2) and in some instances the first formant (F1), without systematically considering the combined impact of the first three formants and 2) not systematically quantifying the proximity of formants via the Bark-Difference metrics as an index to constriction location. In fact, we saw previously that the proximity between formants, following the "Elaborated Target" approach presented above (see Section 4.2, pp. 49), especially that between F1 and F2 or F2 and F3 helps in describing vowels, reducing overlap and is used as a normalisation technique. Based on the psycho-acoustic distances between formants, the effective formant, F', rather than the absolute formant is used in perception as the peaks become spectrally integrated when the distance between formant peaks is within 3.5 Bark (Chistovich & Lublinskaya, 1979; Vaissière, 2011). Syrdal and Gopal (1986) were the first to systematically quantify vowel differences using the Bark-Distance approach, which allowed researchers to correlate the distance to specific distinctive features. For instance, Z1-Z0 (i.e., F1-f0 in Bark) correlates well with the openness dimension (i.e., [±High]) (Fahey, Diehl, & Traunmüller, 1996; Hoemeke & Diehl, 1994; Syrdal & Gopal, 1986; Traunmüller, 1981): closed ([+High]) vowels have a Z difference lower than 3 Bark; around 5 Bark for open ([−High, +Low]) vowels (Syrdal & Gopal, 1986, pp. 1090). Z2-Z1 correlates well with compactness of the spectrum (Sylak, 2011; Syrdal & Gopal, 1986) and is highest for front ([−Back, +High]) vowels, and lowest for (mid-)open back ([+Back, −High, −Low/+Low]) vowels (Syrdal & Gopal, 1986, pp. 1090). And finally Z3-Z2 correlates well with the backeness dimension: highest difference for (mid-)open back ([+Back, −High, −Low/+Low]) and back ([+Back]) vowels; smallest for front vowels ([−Back]) (Syrdal & Gopal, 1986, pp. 1090). In addition, Z3-Z2 correlates well with "spectral flatness" or "divergence" of spectral peaks for F2 and F3 as seen in pre-palatal /i/ to distinguish it from /y/ (Wood, 1986).

In the context of pharyngealisation, the hypothesis advanced by J. Al-Tamimi (2017b) is that a direct consequences of the constriction in the pharyngeal area will lead to an increased difference between F1-f0 (↑Z1-Z0) due to increased openness, a smaller difference between F2-F1 (↓Z2-Z1) due to increased compactness, and a greater difference between the second and third formants (↑Z3-Z2) due to spectral divergence (Carré & Mrayati, 1992; Mrayati et al., 1988; Stevens, 1989). Hence we expect the vowels in the vicinity of pharyngealised consonants to show the following pattern: ↑Z1-Z0, a ↓Z2-Z1, and a ↑Z3-Z2. These differences should be more obvious at the onset rather than the midpoint of the vowel due to coarticulatory changes associated with the impact of pharyngealisation on surrounding vowels.

Figure 4.25 shows the patterns reported in J. Al-Tamimi (2017b) for supra-laryngeal changes quantified via formant frequencies. Looking at the results for both JA and MA, it is clear that when using Absolute Formants (F1*F2 vowel space) at the Onset, vowel spaces to be more back for the former and more back and low for the latter (Figures 4.25a and 4.25c); these correlate with the [+Back] and [+Low] features, especially in MA. At the midpoint, the impact of pharyngealisation on vowels in JA (Figure 4.25b) is minimal, with marginal differences in back vowels becoming more back and with short vowel centralisation; in MA, the full vowel space is still produced as more back and more low (see Figure 4.25d). These results show a clear difference between the two dialects in how pharyngealisation impacts on the vowels and is likely to be correlated with the differences between dialects with a more "guttural" quality reported in Bellem (2007). Looking at the Bark-Difference results, it is clear that at the Onset, JA (Figure 4.25e) and MA (Figure 4.25g) vowels show



(a) JA - AF - Onset (b) JA - AF - Mid (c) MA - AF - Onset (d) MA - AF - Mid

(e) JA - BD - Onset (f) JA - BD - Mid (g) MA - BD - Onset (h) MA - BD - Mid

(i) Predictions of /dˤ/ vs /d/ responses Bark Difference

Figure 4.25: Formant Frequencies for both JA (left: a, b, e and f) and MA (right: c, d, g and h) across /dˤ/ or /d/ using Absolute Formants (AF, top) or Bark-Difference (BD, bottom) and onset (a, c, e and g) or midpoint (b, d, f and h): (a) JA F1∗F2 vowel space at onset; (b) JA F1∗F2 vowel space at midpoint; (c) MA F1∗F2 vowel space at onset; (d) MA F1∗F2 vowel space at midpoint; (e) JA Z3-Z2∗Z2-Z12 vowel space at onset; (f) JA Z3-Z2∗Z2-Z1 vowel space at midpoint; (g) MA Z3-Z2∗Z2-Z1 vowel space at onset; (h) MA Z3-Z2∗Z2-Z1 vowel space at midpoint; (i) Predicted Probabilities for /dˤ/ vs /d/ in JA (blue solid) or MA (red dashed) curves across z-scores (-3 to 3), at onset (top) and midpoint (bottom) for each of Z1-Z0 (left), Z2-Z1 (middle) and Z3-Z2 (left) (reproduced from J. Al-Tamimi, 2017b, pp. 20 and 21)

more compact spectrum (↓Z2-Z1), with spectral divergence/flatness (↑Z3-Z2). At the midpoint, spectral compactness and divergence/flatness are more evident in MA (Figure 4.25h) than in JA (Figure 4.25f).

Finally, using a predictive modelling approach (Baguley, 2012; Hastie, Tibshirani, & Friedman, 2017; Kuhn & Johnson, 2013), we quantified the differences between /dˤ/ and /d/ using a Bayesian Generalised Linear Mixed-effect Model with each acoustic correlate used as a predictor (z-scored) and the class (/dˤ/ and /d/) as an outcome. Figure 4.25i shows the predicted probabilities for the two classes using each of the three Bark-Difference measures (i.e., Z1-Z0, Z2-Z1 and Z3-Z2) between z-scores -3 to 3 (inspired by the approach used in Baumann & Winter, 2018, on perceptual identification of prominence in German). Z2-Z1, especially at the Onset (Figure 4.25i middle column), shows a perfect falling sigmoid curve from /dˤ/ (at z-score -3) to /d/

(at z-score 3), indicating a decrease in Z2-Z1 values in /dˤ/ in comparison with that in /d/. Z3-Z2 (Figure 4.25i left column) shows a perfect rising sigmoid curve indicating a reverse pattern with increased values in /dˤ/, whereas Z1-Z0 (Figure 4.25i right column) shows a near flat-rising sigmoid curve. The patterns seen here match perfectly the predictions of pharyngealisation to be associated with a ↑Z1-Z0, a ↓Z2-Z1, and a ↑Z3-Z2.

In terms of the constriction location, previous results showed that a low pharyngeal constriction is predicted to have a Z2-Z1 difference around and below 3 Bark (obtained within pharyngeal consonants, see Heselwood & Al-Tamimi, 2011, Figure 10, pp. 123), whereas at the vowel onset, a Z2-Z1 was estimated to be around 4.46 Bark for a pharyngealised contexts (averaged across males in all vowel contexts; with 5.1 Bark in /iː/, 3.5 Bark in /aː/ and 4.8 Bark in /uː/, see F. Al-Tamimi & Heselwood, 2011, based on F1 and F2 frequencies in Bark Table 5, pp. 179). The results presented in J. Al-Tamimi (2017b) showed that on average, Z2-Z1 in the pharyngealised context in JA was close to 5.5 Bark and that of MA close to 6.2 Bark at the Onset; at the midpoint, Z2-Z1 was close to 5.5 Bark in JA and 5.1 Bark in MA, with [+Low] and [+Back] vowels in both dialects having Z2-Z1 below the 4 Bark. Z3-Z2 provided an additional quantification of the difference between pharyngealised and non-pharyngealised contexts, and between the two dialects. JA seems to show increased Z3-Z2 difference in comparison to MA due to differences related to vowel density reported previously. The predictions advanced in J. Al-Tamimi (2017b) is that JA seems to show a constriction location around the upper-mid pharynx, whereas that of MA is between the mid-low pharynx.

4.5.2.4 Laryngeal changes

A second issue highlighted in J. Al-Tamimi (2017b) concerns the lack of examination of the acoustic correlates of the raised larynx and pressed/tense voice quality reported in the literature (see e.g., F. Al-Tamimi & Heselwood, 2011; Cantineau, 1960; Lehn, 1963; Zeroual et al., 2011, among others). In fact, Cantineau (1960, p. 23-24) was among the first to report that only pharyngealised consonants were produced with a "pressed voice". This "pressed voice" quality was not explored in a systematic way in previous literature and the I wanted to evaluate whether it relates to a "true" laryngeal change; a point explored further in J. Al-Tamimi (2017b, see as well: J. Al-Tamimi, 2014; J. Al-Tamimi, 2015; J. Al-Tamimi, 2016; J. Al-Tamimi, 2017a).

Looking at the literature on the various phonation types, it is clear that constricting the ventricular folds leads to various changes in the voice quality that is identified with various terms, such as a tense voice, a pressed voice, a laryngealised voice or an extreme creaky voice. Acoustically speaking, this constriction leads to an an overall lowered or flatter spectral tilt, due to the increased energy above the first harmonic (Halle & Stevens, 1969; Hanson, Stevens, Kuo, Chen, & Slifka, 2001; Klatt & Klatt, 1990; Laver, 1980, 1994; Moisik, 2013b; Moisik & Esling, 2010; Stevens, 2000; Sundberg & Askenfelt, 1981). Based on previous literature on phonation changes associated with [−ATR] vowels, or with those examining (epi-)laryngeal changes in the singing voice, J. Al-Tamimi (2017b) examined spectral tilt changes via the following acoustic metrics[17] (following Aralova, Grawunder, & Winter, 2011; Fulop, Kari, & Ladefoged, 1998; Garellek, 2012; Guion, Post, & Payne, 2004; Hanson et al., 2001; Kang & Ko, 2012; Keating, Garellek, & Kreiman, 2015; Klatt & Klatt, 1990; Kuang & Keating, 2012, 2014; Ladefoged & Maddieson, 1996; Story, 2019):

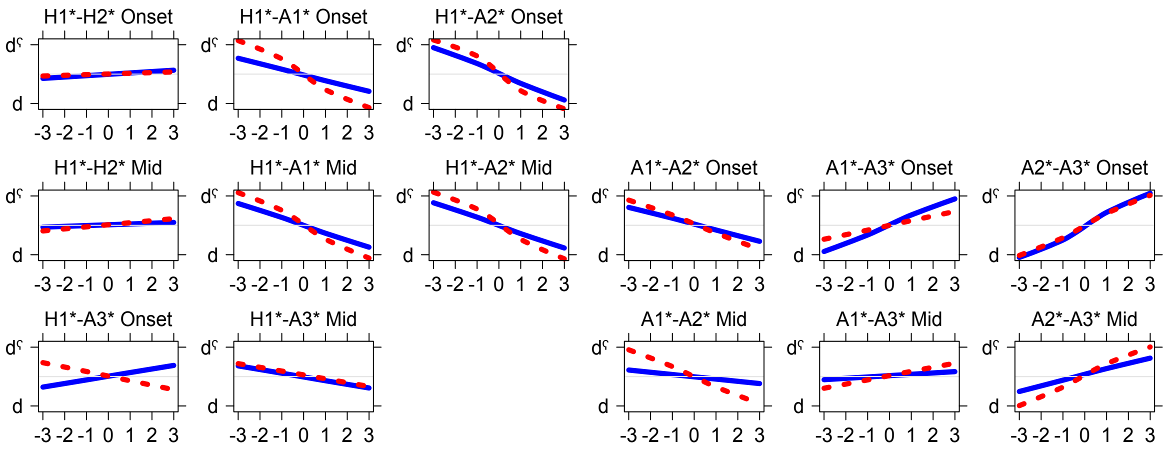
[17] $H1$ = Harmonic closest to $f0$; $H2$ = Harmonic closest to $2*f0$; $A1$-3 = Harmonic closest to F1-3; * = Normalised metrics following Iseli, Shue, and Alwan (2007)

3. 1) ↓*H1*-H2** ⇒ Main acoustic correlate differentiating breathy from tense, pressed, laryngealised or creaky voice
   2) ↓*H1*-A1** ⇒ Correlates with abruptness of vocal fold closure; increased F1 bandwidth, leading to decrease in spectral tilt
   3) ↓*H1*-A2** ⇒ Same as *H1*-A3**
   4) ↓*H1*-A3** ⇒ Correlates with abruptness of vocal fold closure, leading to increased energy around F2-F3 and a decrease in spectral tilt
   5) ↓*A1*-A2** ⇒ Main correlate for [−ATR] vowels; bandwidth of F1 and F2 increased due to pharyngeal and laryngeal constrictions
   6) ↑*A1*-A3** ⇒ Correlates with abruptness of vocal cord closure, hence decrease in energy around F3 and increase in the difference or with an enhanced voice (in singing) as a consequence of an extreme epilaryngeal constriction, leading to an increase in energy around F3 and hence a decrease in the difference
   7) ↑*A2*-A3** ⇒ Same as *A1*-A3**

Figure 4.26 presents the results of this study. Starting with spectral tilt measures (i.e., measures 1-4 from list 3) at both onset and midpoint of the vowels in both JA and MA, Figure 4.26a shows that the amplitude of *H1*-A1** and *H1*-A2** have falling sigmoid curves from /dˤ/ (at z-score -3) to /d/ (at z-score 3), indicating an overall lowering of these spectral tilt measures. This correlates well with the overall decrease in spectral tilt associated with tense or pressed voice reported in the literature. Interestingly, *H1*-H2**, which is often reported to be the main correlate for the phonation contrast across languages shows a near flat curve in both dialects. This is likely due to the fact that the tense voice quality seen here is a by-product of an epilaryngeal constriction in an articulation characterised by a secondary pharyngeal constriction; hence is secondary rather than being primary. Finally, *H1*-H3** at the onset shows an opposite direction depending on the dialect, with MA showing a falling curve; a raising one for JA. At the midpoint, both dialects show similarities in the direction of the curve. It is possible that the type of realisation of the epilaryngeal constriction and the tense voice quality is dialect-dependent. Next, looking at the acoustic correlates for [−ATR] vowels and for the high frequency component (measures 5-7 from list 3), Figure 4.26b shows both dialects to display similar curves, with that of *A2*-A3** at the onset displaying the most sigmoidal raising shape. *A1*-A2** show an overall decrease, and both *A1*-A3** and *A2*-A3** show an overall increase. These measures correlate again well with a tense voice quality induced by a partial epilaryngeal constriction due possibly to an abrupt closure of the glottis (Hanson et al., 2001).

4.5.2.5 The feature [+cet] - supra-laryngeal and laryngeal changes

As highlighted in the previous two sections, the results reported in J. Al-Tamimi (2017b) showed two important and novel aspects. On the one hand, supra-laryngeal changes quantified via absolute formant frequencies and bark-differences are clearly indicative of changes in the spectral domain as follows. A vowel following a pharyngealised consonant is associated with an ↑F1, a ↓F2, and a variable F3, with an ↑F3 in an /uː/ context, but with a ↓F3 in an /iː/ (and possibly /aː/) context. These results were previously reported by my own research and by others (as summarised above). The novelty lies in the use of the bark-difference metrics as the results showed a systematic ↑Z1-Z0, a ↓Z2-Z1, and a ↑Z3-Z2 pattern. These results clearly show a difference between "true" pharyngeal consonants and impact of pharyngealisation on vowels in



(a) Spectral Tilt (b) High Frequency components

Figure 4.26: Predicted Probabilities for spectral tilt measures in /dˤ/ vs /d/ in JA (blue solid) or MA (red dashed) curves across z-scores (-3 to 3): (a) Spectral tilt quantified via harmonic differences in relation to *H1**; (b) High Frequency components quantifying the Spectral tilt between Formant Harmonics (reproduced from J. Al-Tamimi, 2017b, pp. 22 and 23)

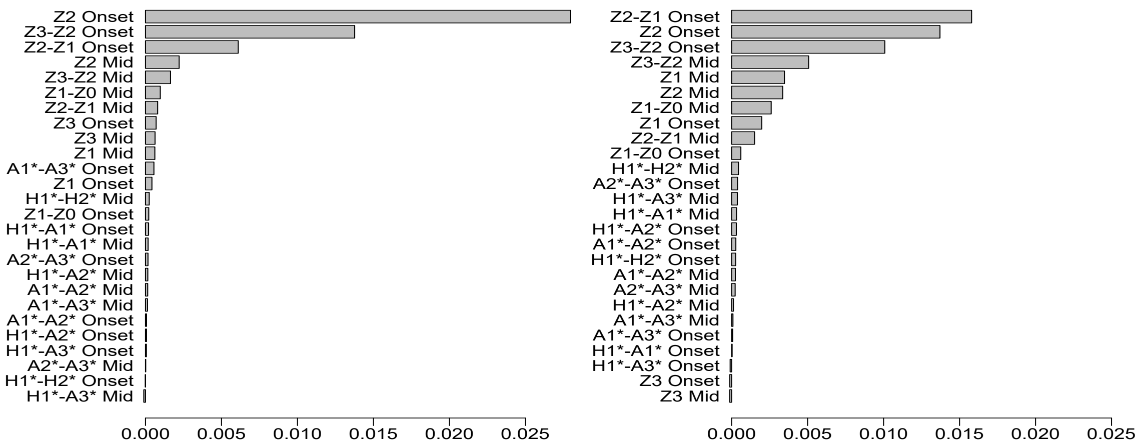
both dialects; the constriction location is thus located between the upper and mid-pharynx in both, but with a slighlty lower constriction location in MA. In addition, the results provided a psycho-acoustic account for the feature [+Flat] advocated by Jakobson (2002); Jakobson, Fant, and Halle (1952/1976) to correlate with the impact of the pharyngealised context on vowels and correlates well with the observed differences on Z2-21 and Z3-Z2. In terms of laryngeal changes, J. Al-Tamimi (2017b) reported for the first time on these systematic changes associated with pharyngealisation using spectral tilt and high frequency components. Overall, vowels following a pharyngealised consonant showed a decrease in spectral tilt via ↓*H1*-A1**, ↓*H1*-A2**, ↓*H1*-A3**, ↓*A1*-A2**, but with a decrease in high frequencies around F3 with an ↑*A1*-A3** and ↑*A2*-A3**. This specific spectral tilt profile is similar to what has been reported as acoustic consequences of a tense, pressed and/or laryngealised voice quality. The combined impact of supra-laryngeal and laryngeal changes is clearly indicative of an epilaryngeal constriction that accompanies pharyngealised consonants in both dialects.

Subsequent to this analysis and to evaluate whether supra-laryngeal and/or laryngeal changes provide a meaningful account of the patterns observed in the data, all of the 26 measures (13 acoustic measures obtained at both onset and offset) were submitted to a Random Forests analysis grown via the Conditional Inference Trees framework (Strobl et al., 2008, 2007; Strobl, Hothorn, & Zeileis, 2009; Strobl, Malley, & Tutz, 2009), which was run for each dialect separately and in seven different versions: 1) Absolute Formants (Form) + Bark-Difference (BkDiff) + Voice Quality (VQ); 2) Form + BkDiff; 3) Form; 4) BkDiff; 5) Form + VQ; 6) BkDiff + VQ and 7) VQ. The aim of these versions was to assess whether absolute formants, Bark-Difference and/or VQ are more predictive to the differences. The results presented in Table 4.4 show the performance of the classifiers. It is clear that the best performance is obtained when combining all 26 measures (i.e. Form + BkDiff + VQ) in both dialects, with extremely high classification rates in separating /d/ from /dˤ/. The rates are almost identical with either combination of supra-laryngeal changes (i.e., Form and/or BkDiff). Interestingly, Bark-difference measures on their own provide a mere increase by 0.8% in JA or by 0.3% in MA in comparison to absolute formants on their own; which highlights that Bark-Difference measures may be better suited to describing the patterns in the data. Finally, when looking at Voice Quality measures on their own, the rates of classification dropped to 70.2% in JA and to 75.8% in MA; these classification results are better than chance (set at 50%), but are clearly indicative of a secondary role of VQ measures to the contrast.

Table 4.4: Summary of predictive accuracy as classification rate for each of the random forests in JA and MA. Form = Absolute Formants; BkDiff = Bark-Difference formants; VQ = Voice Quality (reproduced from J. Al-Tamimi, 2017b, pp. 25)

	Form+BkDiff+VQ	Form+BkDiff	Form	BkDiff	Form+VQ	BkDiff+VQ	VQ
JA	93.5%	93.2%	92.1%	92.9%	92.2%	93.1%	70.2%
MA	91.1%	91.0%	90.5%	90.8%	90.6%	91.0%	75.8%

Figure 4.27 shows the results of the Variable Importance Scores obtained via Conditional Permutation (to account for collinearity). The results for JA (Figure 4.27a) show that Z2 (F2 in Bark), followed by Z3-Z2 and Z2-Z1 at the Onset, then at the midpoint are the best predictors; in MA (Figure 4.27b), it is Z2-Z1, followed by Z2 and Z3-Z2 at the Onset, then at the midpoint are the best predictors. Clearly all of the VQ measures have minimal contribution to the contrast, but they are not cancelled (except from the last three, which have negative scores or zeros).



(a) Variable Importance - JA (b) Variable Importance - MA

Figure 4.27: Variable Importance scores obtained via conditional permutation ordered from strongest to weakest predictor: (a) in JA; (b) in MA (reproduced from J. Al-Tamimi, 2017b, pp. 25)

The results reported in J. Al-Tamimi (2017b, see as well: J. Al-Tamimi, 2014; J. Al-Tamimi, 2015; J. Al-Tamimi, 2016;J. Al-Tamimi, 2017a) showed for the first time that the pharyngealised coronal voiced stop in both JA and MA are acoustically associated with a combined articulatory gesture: The have tongue retraction as a primary feature associated with a secondary laryngeal constriction causing a tense voice quality. These combined gestures are not independent from each other; they are combined and are due to an epilaryngeal constriction that we claim is partial rather than extreme as would be expected from "true" pharyngeals. In our case, and following LAM, we postulated that this combined gesture is due to either a laryngeal constriction and/or larynx raising which are the primary cause for tongue root and dorsum retraction. From a phonological point of view, and as highlighted in J. Al-Tamimi (2017b, Section 4.1, pp. 29-30), it is unclear if the feature [+RTR] that is widely postulated to be the main phonological feature to account for the impact of pharyngealisation on surrounding vowels could account for laryngeal changes, as it is primarily a lingual feature (Moisik et al., 2012; Moisik, Czaykowska-Higgins, & Esling, 2019). The phonological features [+Retracted] and [+CG] (for "Constricted Glottis") have been claimed to account for epilaryngeal changes as combined lingual and laryngeal features, which are the main features combined in the newly postulated feature [+cet] (Esling et al., 2019; Moisik et al., 2012, 2019; Moisik & Esling, 2011). Finally, accounting for both phonetic and phonological changes associated with this type of constriction allows to explain the supra-laryngeal and laryngeal changes as observed in this study.

4.5.3 Guttural consonants - From articulation to acoustics

4.5.3.1 Background

J. Al-Tamimi (2017b) was the first study to provide an acoustic evidence for a partial epilaryngeal constriction during the production of voiced pharyngealised coronal stops in both Jordanian and Moroccan Arabic dialects by systematically using supra-laryngeal and laryngeal measures to the contrast. This combination allows to evaluate the filter and the source, respectively, by using indirect acoustic measures. To further highlight the contribution of the epilarynx in the production of back consonants in Arabic, I conducted a relatively major study[18].

The main body of research within this section falls within these publications: J. Al-Tamimi (2018b, BAAP on using GAMMs with UTI data in guttural consonants in light of LAM), J. Al-Tamimi (2018a, 16th LabPhon on using GAMMs with UTI data in guttural consonants in lights of LAM), J. Al-Tamimi (2018c, Invited talk at Atelier « Temporalité et Séquentialité Dans Les Formes Sonores », Nantes, on using GAMMs with UTI data in guttural consonants in lights of LAM), J. Al-Tamimi (2019c, Invited talk at the SRPP, LPP, Paris, on using GAMMs with UTI data and initial acoustic results in guttural consonants in lights of LAM), J. Al-Tamimi (2019b, Invited talk at the LIMSI, Paris, on distinguishing between male and female speakers using guttural consonants), J. Al-Tamimi and Ferragne (2020a, BAAP on using deep learning and random forests to distinguish between guttural consonants using acoustics), Turton, Khattab, Alsharif, and Al-Tamimi (2020, BAAP UTI analysis on emphatic conditioning), J. Al-Tamimi and Ferragne (2020b, 17th LabPhon on using deep learning and random forests to distinguish between guttural consonants using acoustics), J. Al-Tamimi and Palo (2020, 12th ISSP on quantifying retraction using UTI data modelled via GAMMs), J. Al-Tamimi (2020a, Invited talk at the SSF, Cambridge, on using the role of the epilarynx in describing guttural consonants in lights of LAM, using UTI, Acoustics and EGG data), J. Al-Tamimi (2021a, 4th PaPE presenting the acoustic profile of gutturals using random forests), J. Al-Tamimi (2022l, Invited talk at the SRPP, LPP, Paris on a whole tongue approach to guttural consonants using GAMMs on UTI data), J. Al-Tamimi and Palo (2023, 20th ICPhS presenting on a whole tongue approach to guttural consonants using GAMMs on UTI data), J. Al-Tamimi and Palo (In preparation, Article Laboratory Phonology presenting on a whole tongue approach to guttural consonants using

[18]This research project was funded by a British Academy/Leverhulme small research grant, UK (SG160181; 2017-2019), a Leverhulme International Academic Fellowship, UK (IAF-2018-016), with continued partial support from the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083), contributing to the IdEx Université Paris Cité - ANR-18-IDEX-0001

GAMMs on UTI data in the light of LAM), J. Al-Tamimi (In preparation, Article JASA presenting results of laryngeal and supra-laryngeal activity in gutturals in the light of LAM).

As highlighted in the introduction (see Section 4.5.1, pp. 93), my study focused on the status of the guttural consonants in Levantine Arabic by examining how these sounds are produced, in terms of their front and back lingual configuration, in addition to laryngeal constriction and height, using a synchronised Ultrasound Tongue Imaging (UTI), ElectroGlottoGraphy (EGG; two channels for Larynx height and contact) and acoustic analyses of supra-laryngeal and laryngeal changes. Given the claims that gutturals form a natural class, due to their conditioned alternations, usage of specific morphophonological patterns and being produced in the same oro-sensory zone in the pharynx (J. McCarthy, 1994) it is assumed that they should show similarities based on the phonetic grounding of phonology hypothesis whereby all phonological features are phonetically implemented and are part of the grammar (Pierrehumbert, 2000). Members of the guttural natural class are primarily pharyngeals /ħ ʕ/ (or epilaryngeal /ʜ ʢ ʡ/) and uvulars (/χ ʁ q/) (J. McCarthy, 1994), but pharyngealised consonants (/tˤ dˤ ðˤ sˤ/) (Sylak-Glassman, 2014a), and/or Laryngeals (/ʔ h/) (Zawaydeh, 1999) are also considered to be part of this class. In fact, researchers showed that members of this class use a variable region in the pharynx and have different tongue configurations, degrees of retraction and/or backness, different positions of the larynx and/or of laryngeal constriction (F. Al-Tamimi & Heselwood, 2011; Bellem, 2007; Heselwood & Al-Tamimi, 2011; Zeroual, 2003; Zeroual & Clements, 2015). From a coarticulatory point of view, gutturals seem to favour regressive rather than progressive coarticulation, which leads to different feature spreading patterns depending on the category of the consonant; pharyngealised consonants show stronger regressive feature spreading than other members (Bellem, 2007; Hellmuth, 2013).

Although members of the guttural class seem to be implemented differently from a phonetic point of view, the aim of this study is to use the multiple combined articulatory (UTI and EGG) and acoustic measures to highlight how (dis-)similar the members of the guttural class are to each other and whether there are specific features that unit these consonants on either their realisation but also their impact on surrounding vowels. The predictions of LAM are important in this context and we use the specific physiological states described in Esling et al. (2019); Moisik et al. (2019) as the "Phonological Potentials" of lingual states (i.e., {tre} for tongue retraction (back and down), {tra} for tongue raising (back and up), {tfr} for tongue fronting and {tdb} for tongue double-bunching). These phonological potentials explain that particular states can be either synergistic or anti-synergistic with an epilaryngeal constriction. For instance, {tfr} and {tra} are anti-synergistic with an epilaryngeal constriction, whereas {tre} and {tdb} are synergistic (Moisik et al., 2019). The three states {tfr}, {tra} and {tre} are claimed to be anti-synergistic with each other (Moisik et al., 2019, pp. 11), i.e., one cannot have both {tfr} and {tra} at the same time. However, for specific contexts, e.g., uvular stops, the assumption is that both {tra} and {tre} are used: uvular stops are produced by a closure of the oropharyngeal isthmus: they have tongue raising {tra} as a primary state *and* tongue retraction {tre} as a secondary one (Esling et al., 2019, pp. 179-180). Tongue double bunching will have two associated states: tongue fronting {tfr} towards the palatal region and tongue retraction {tre} towards the pharynx (for more details, see Esling et al., 2019; Moisik et al., 2019).

When using the predictions of LAM and relying on the phonological potentials, it is evident that one can have categorical predictions for extreme or null epilaryngeal constriction, however, and following Esling (2005); Sylak-Glassman (2014a, 2014b), we claim that an epilaryngeal constriction needs to be more gradient, ranging from minimal (or null) at the top end, moving to partial, in the middle of the Laryngeal Vocal Tract and then to an extreme case towards the Lower part of the Laryngeal Vocal Tract. Indeed, Esling (2005, Figure 3, pp. 21) shows that when compared to the larynx in a neutral position (Esling, 2005, Figure 3 (a), pp. 21), two types of constrictions can be seen. The first is the case of an extreme epilaryngeal constriction, where there is an "*...almost complete laryngeal constriction, with a narrowed aryepiglottic passage, shortened vocal folds, extreme larynx raising, and extreme tongue retraction*" (Esling, 2005, legend (c) of Figure 3, pp. 21), whereas a second type of epilaryngeal constriction that is more partial with "*...partial aryepiglottic fold sphinctering, moderate larynx raising, and moderate tongue retraction*" (Esling, 2005, legend (b) of Figure 3, pp. 21).

Given this minute change, it is predicted that an extreme epilaryngeal constriction leads to a maximal tongue retraction {tre} and larynx raising {↑lx}, when it is located towards the lower pharynx as is the case with epilaryngeal consonants /ʜ ʢ ʡ/); it is null or minimal towards the uvula due to the primary {tra} and secondary {tre} in the case of uvular consonants (/χ ʁ q/, with no specific predictions for larynx height, with potentially different predictions for stops and fricatives, see Esling et al., 2019; Sylak-Glassman, 2014b); and it is variable in-between these two extremes due to variable degrees of larynx raising/constriction and tongue root/body retraction. This partial epilaryngeal constriction can be seen as correlating with cases of double bunched consonants (e.g., retroflex consonants /ɻ/ or even approximant realisation of the voiced pharyngeal /ʕ/), in addition to pharyngealised consonants. For laryngeals, the predictions are that if they are produced as glottal, they will only trigger larynx raising/constriction and not tongue root retraction due to them being produced with a narrower epilaryngeal tube (see Esling et al., 2019, pp. 50). However, if laryngeals are produced as "epiglottal", they will cause the tongue root/epiglottis to be retracted and the larynx to be raised, leading to an extreme epilaryngeal constriction (see Esling et al., 2019, pp. 51).

This study looked at this combined gesture from an articulatory (UTI and EGG) and acoustic point of view, to provide a clear articulatory to acoustic mapping of the changes in the Laryngeal Vocal Tract. The data were obtained from 10 Urban Levantine Arabic speakers (5 males and 5 females) coming from Lebanon (7: 3 from south, 2 from Beirut, 2 from North), Palestine (1), Jordan (1) and Syrian (1). Although participants come from four different countries, they all speak an urban variety of Arabic, identified primarily by their realisation of: /q/ as /ʔ/, the uvulars /χ ʁ/ as /x ɣ/, an approximant realisation of /ʕ/, the usage of /ʒ/ rather than /d͡ʒ/ and presence of /zˤ/ as a reflex of /ðˤ/ (Embarki, 2008, 2013; Versteegh, 2001). Speakers were asked to produce a list of real and nonce-words in the following frame /ˈʔVːˈCVː/, with Vː = symmetric /iː aː uː/; C = all possible consonants in Levantine and other Arabic varieties = /b t d m n r f θ ð s z ʃ ʒ l w j k g x ɣ q tˤ dˤ ðˤ sˤ zˤ lˤ ħ ʕ ʔ h/), with three symmetric repetitions. A third of the items was used as fillers, and 21 consonants were categorised into the following six classes (see list 4):

4. 1) **Plain** ⇒ /t d ð s z l/
   2) **Velar** ⇒ /k g x ɣ/
   3) **Uvular** ⇒ /q/
   4) **Pharyngealised** ⇒ /tˤ dˤ ðˤ sˤ zˤ lˤ/
   5) **Pharyngeal** ⇒ /ħ ʕ/
   6) **Glottal** ⇒ /h ʔ/

The six classes were then merged into two major classes: gutturals vs non-gutturals. Uvular, pharyngealised and pharyngeal were included within the guttural class, whereas plain (coronals), velars and glottal were included in the non-guttural. This division was informed by the auditory, the UTI and the acoustic results of this study. From an auditory point of view, all productions of the glottal stop /ʔ/ especially at the beginning of the /ˈʔVːˈCVː/ frame was more of a glottal rather than epiglottial. From UTI and acoustic analyses (see Sections 4.5.3.2 and 4.5.3.4), glottals did not exert any major influences on the tongue contour, nor had any major coarticulatory influences on the surrounding vowels. Hence, we took these results as a clear indication of glottals not belonging to the guttural natural class. We still examine their patterns to inform the theory. The following sections present the results for each of the UTI, EGG and acoustic measures, before concluding with a formal account to guttural consonants (see J. Al-Tamimi, In preparation; J. Al-Tamimi & Palo, 2023, In preparation, for further details on material used and data analyses).

4.5.3.2 Ultrasound Tongue Imaging

UTI data from 8 participants (4 males; 4 females) were analysed, using Articulate Assistant Advanced (AAA, version 2.18.04 A. A. Wrench, 2018). Before splining the data, the UTI video recordings were de-interlaced to 59.977 fps (frame length = 16 ms). The acoustic boundaries from the forced-aligned speech were used in subsequent analysis to guide landmarks selection. Given that we wanted to dynamically track the changes in tongue contours throughout the VCV sequence, we specified nine-time intervals (timeFrame) within a VCV sequence with a 25% interval shift, starting with the temporal location at 50% of the preceding vowel (V1) and ending at 50% of the following vowel (V2); the other seven intervals were equally spread across the remaining portion at 75% of V1; at 0%, 25%, 50%, 75%, and 100% of the medial consonant (C2), and at 25% of V2. At each time interval, we traced the full tongue contour, in an unrotated view, from the visible portions on the UTI videos, using a 42 fanline shaped window, which was limited between the hyoid and the mandible bones (for more details, see J. Al-Tamimi & Palo, 2023, In preparation). A total of 13698 tongue splines were automatically splined and then manually checked. The data from the 42 fanline coordinates were exported in Polar coordinates with the Rho values representing tongue height in mm (r) and each fanline representing an angle value in Radians (φ). The first 4 and last 4 fanlines that were hidden by the hyoid and the mandible bones were excluded from subsequent analyses.

Using an Auto-Regressive Generalised Additive Mixed Model (AR-GAMM), we modelled smooths for two time-series: the Angle (34 points) and timeFrame (9 points) as a function of the context by vowel interaction, adjusting for both speaker and items as factor smooths (see full model specification in J. Al-Tamimi & Palo, 2023, In preparation). From the model's predictions, we used three types of plots. The first is a 3D surface plot (Figure 4.28), the second is an averaged (across the five time intervals within C2 and across vowels) 2D spline difference plots between two contexts (Figure 4.29) and the third is a 3D difference plot between two contexts, across the 9 time intervals, and within each of the three vowels (Figure 4.30).

The results presented in Figure 4.28 show the tongue surfaces for the interaction between each of the six contexts and vowels (plain, velar, uvular, pharyngealised, pharyngeal and glottal in columns; /iː aː uː/ in rows). One observation to make is that when looking at the uvular, pharyngealised and pharyngeal contexts (gutturals), similar tongue contours are obtained with similar tongue portions impacted upon, especially in an /iː/ context.

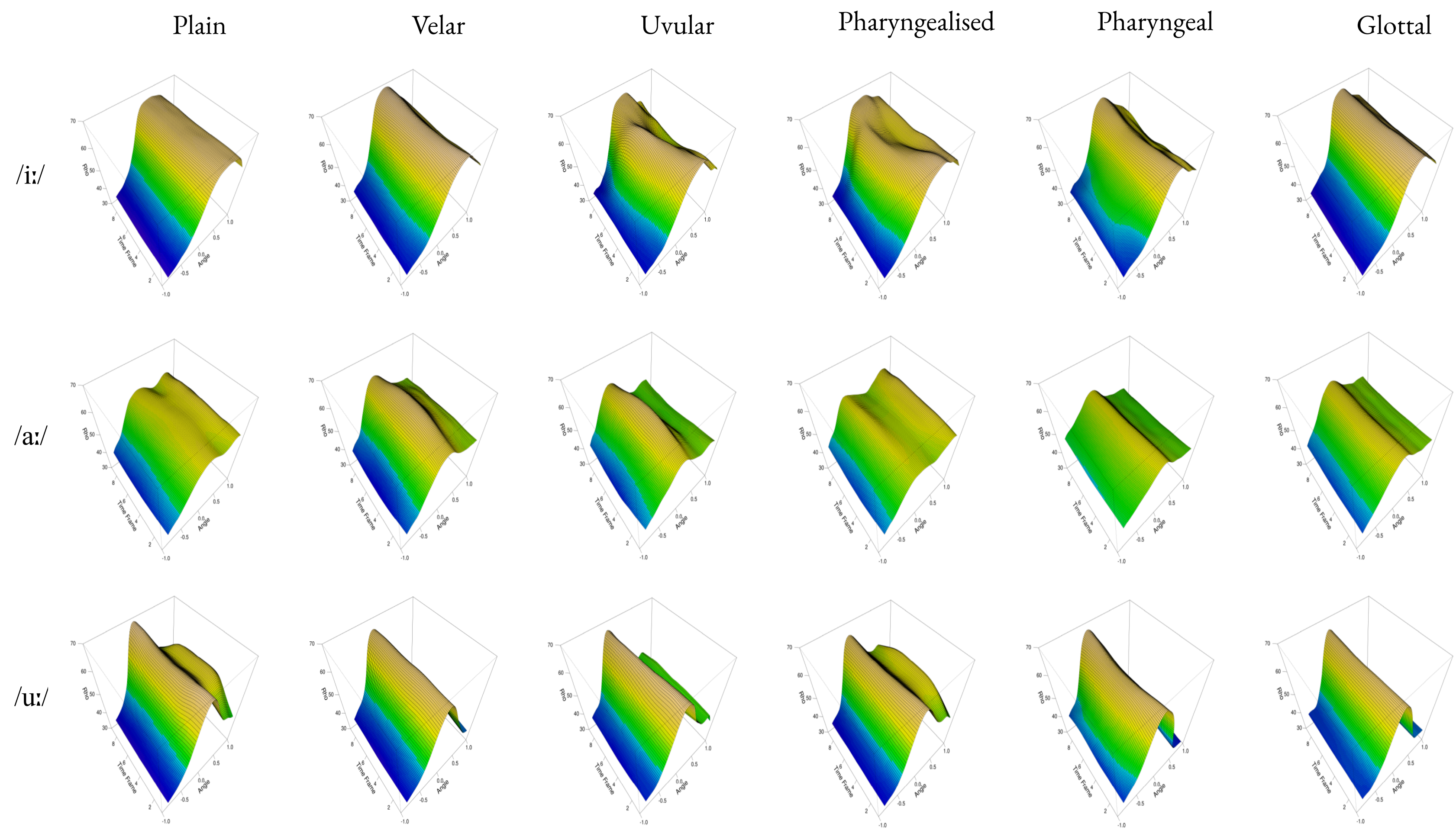


Figure 4.28: 3D surface contours for the six classes (Plain, Velar, Uvular, Pharyngealised, Pharyngeal and Glottal) in the vowels /iː aː uː/ (top, middle and bottom rows), according to the Angle (x-axis; bottom right to top right; tongue root to tip), timeFrame (y-axis; bottom left to top left; V1 50% to V2 50%) and Rho (z-axis; bottom to top). Tongue height is indicated by both height within the image and colour continuum blue-green-orange (reproduced from J. Al-Tamimi & Palo, 2023, In preparation)

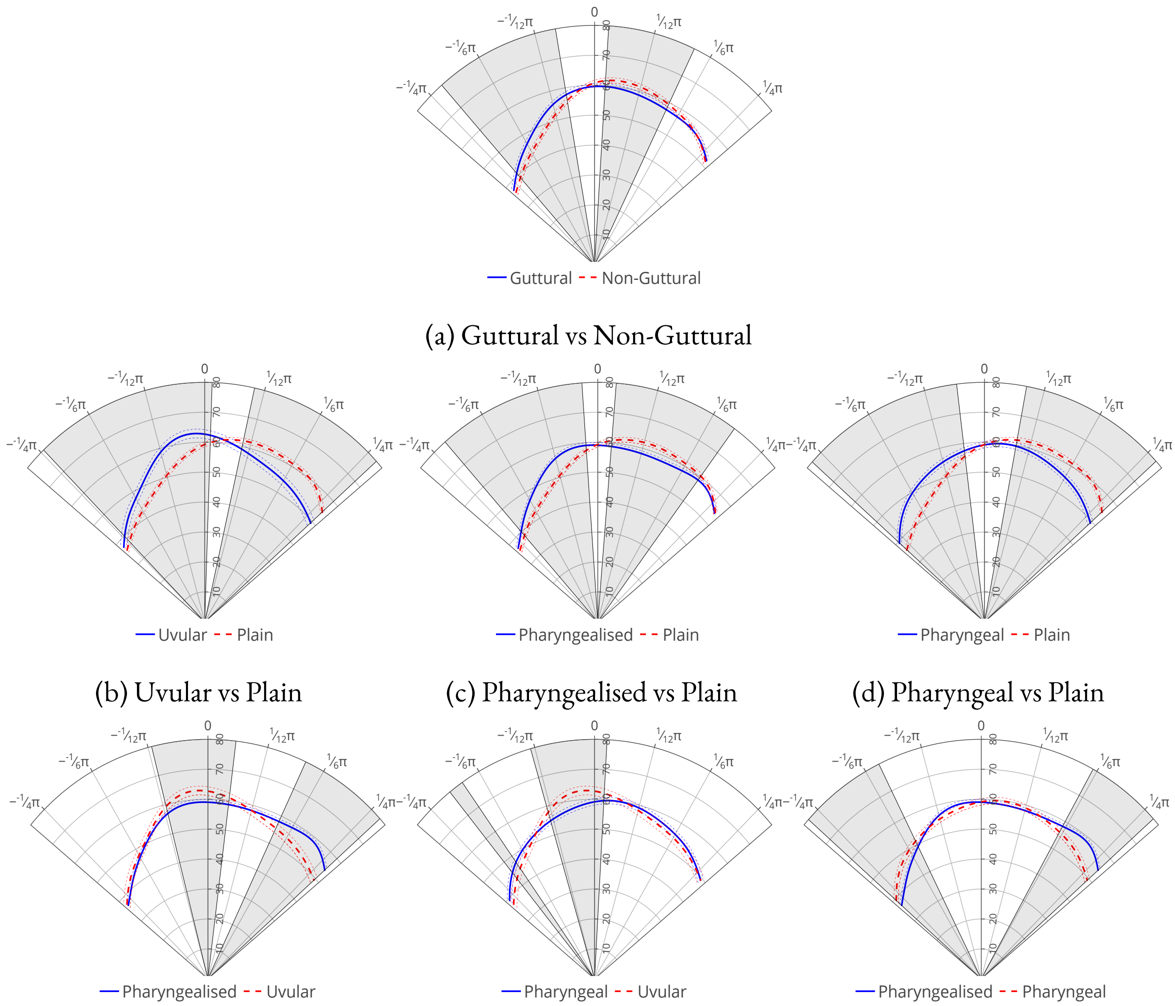
In an /iː/ context (top row), the three guttural contexts show tongue front and dorsum depression (see angle = 0 to +0.84) and retraction towards the tongue back/dorsum area (angle = 0 to -0.84), in addition to tongue root changes. This is also evident within /aː/ and /uː/ to a lesser extent. When comparing with the remaining three contexts (plain, velar and glottal), there is no evidence for tongue front and dorsum depression. Uvular and pharyngealised contexts are more similar to each other, with small differences within the tongue dorsum and root retraction, in addition to tongue height difference. Within /aː/ (middle row), similar patterns are observed between the three guttural contexts, with most of the changes observed towards the front, mid, back/dorsum and root of the tongue. Tongue body is lower than in /iː/ as expected given the [+Low] associated with the vowel /aː/, but shows minimal depression specifically in the pharyngealised context. The three other non-guttural contexts show different patterns to those observed in the gutturals. Finally, within /uː/ (bottom row), most of the changes are observed within the uvular and the pharyngeal contexts, with less differences within the pharyngealised context. The three non-guttural contexts are again different from the guttural. An interesting observation is that most of the changes throughout the VCV sequence are seen within the consonant itself, with changes observed at the vowels' edges, with less/no differences at the vowels' midpoint.

Gutturals show similarities in the type of tongue changes, especially from front to back, with differences in tongue height. The similarities observed in the gutturals are specific and seem to be compatible with the "double bunching" for a "pharyngeal" place suggested by Esling et al. (2019); Moisik et al. (2019), specifically for /iː uː/ contexts, with changes within /aː/ context located at the root of the tongue (for more details, see J. Al-Tamimi & Palo, 2023, In preparation).

To further evaluate how different guttural and non-guttural contexts are, we used a 2D difference plot presented in Figure 4.29, within the consonant itself. These are predicted splines averaged across the five time intervals within the C2 and across all vowels. These are used for demonstrating the type of changes seen towards either the front or back cavities (see J. Al-Tamimi & Palo, 2020, for a similar approach, see Heyne, Derrick, & Al-Tamimi, 2019). Figure 4.29a shows that overall, guttural consonants (i.e., uvular, pharyngealised and pharyngeal) show a more "retracted" tongue back and root and more depressed tongue body and dorsum than non-guttural consonants. Specifically, when comparing the differences between each of the guttural contexts to the plain, similar patterns emerge. The uvular context is different in both front and back areas with a more raised tongue back (Figure 4.29b), the pharyngealised shows tongue body depression and tongue retraction (Figure 4.29c), and the pharyngeal has a lowered tongue body/front and a bunched tongue back and root (Figure 4.29d).

When comparing within each of the guttural contexts, the pharyngealised context shows a more raised tongue front and tip and a more lowered tongue back than the uvular context (Figure 4.29e), which indicates that the pharyngealised context is articulatory different from the uvular. The pharyngeal context shows a much lowered tongue back and a more retracted tongue root than the uvular (Figure 4.29f). Finally, the pharyngealised context shows a more advanced tongue root and a more raised tongue front/tip than the pharyngeal (Figure 4.29g).

It is important to note that these results are averaged across all vowels and the five time frames within the C2. However, they provide a clear indication that in comparison to the plain context, there are similar patterns emerging, but with gradient, rather than categorical changes, towards the tongue back and front areas. Within guttural comparisons are again indicative of differences



(a) Guttural vs Non-Guttural

(b) Uvular vs Plain (c) Pharyngealised vs Plain (d) Pharyngeal vs Plain

(e) Pharyngealised vs Uvular (f) Pharyngeal vs Uvular (g) Pharyngealised vs Pharyngeal

Figure 4.29: Static 2D difference plots, based on averaged timeFrames and vowels across C2 with statistically significant regions (95% CI) highlighted in grey, with the Angle (x-axis; tongue root, left to tip, right), Rho (y-axis; bottom to top): (a) guttural vs non-guttural; (b) Uvular vs Plain; (c) Pharyngealised vs Plain; (d) Pharyngeal vs Plain; (e) Pharyngealised vs Uvular; (f) Pharyngeal vs Uvular; (g) Pharyngealised vs Pharyngeal (reproduced from J. Al-Tamimi & Palo, 2020, In preparation)

between the three contexts, with a potential gradient change. The pharyngeal context shows the most retracted tongue back; the uvular and pharyngealised show similarities in tongue back retraction, but with clear differences in tongue height.

To account for a more dynamic change across the contexts, Figure 4.30 shows a 3D difference plot between each of the guttural contexts and the plain context, with the angle on the x-axis (root left, tip right), the time interval on the y-axis (V1 50% bottom to V2 50% top), the Rho value on the z-axis (tongue height, darker colour indicates lowered value; lighter colour indicates increased value) and the estimated constriction location inspired by Carignan et al. (2020). These difference plots compare each of the members of the guttural natural class to the plain context in /iː aː uː/ contexts. Across all vowels and contexts, gutturals show an increase in the difference at

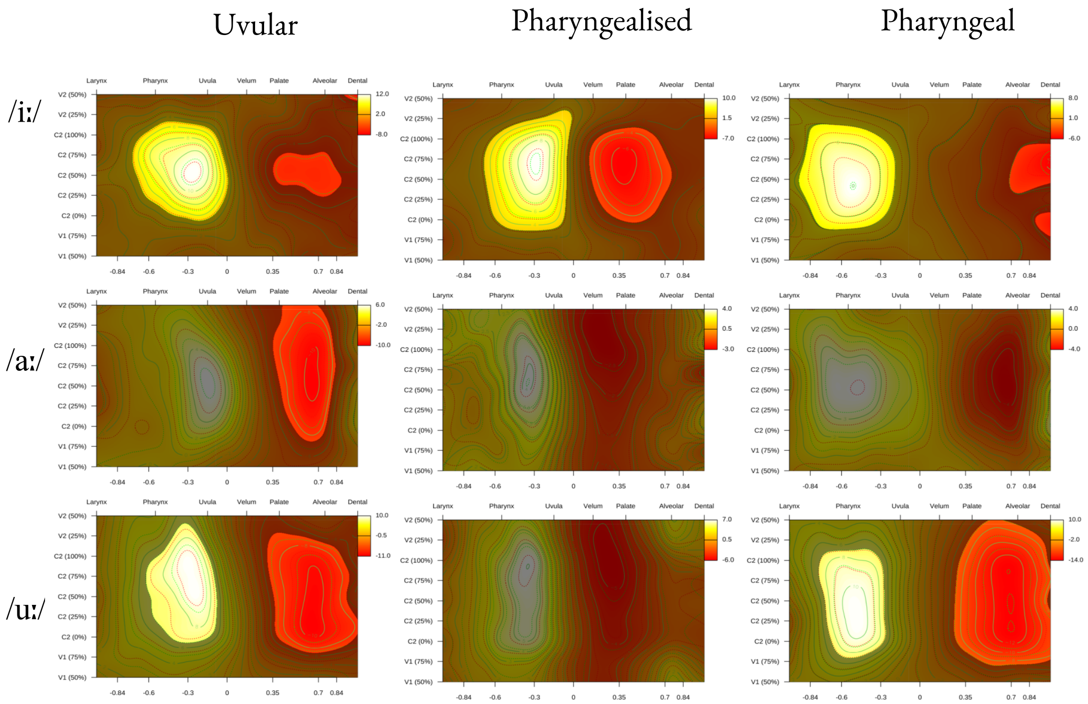


Figure 4.30: 3D difference plots, with statistically significant regions 95% CI highlighted. The plots show differences between each of the three classes (Uvular, Pharyngealised and Pharyngeal) and the plain class, in the vowels /iː aː uː/ (top, middle and bottom rows), according to the Angle (x-axis; tongue root to tip), timeFrame (y-axis; V1 50% to V2 50%), Rho (z-axis; tongue height difference indicated by lighter and darker colours, with lighter = increase in tongue height; darker = decrease), with estimated constriction location (secondary x-axis, top; reproduced from J. Al-Tamimi & Palo, 2023, In preparation)

the back of the tongue (lighter, yellower area), and a decrease towards the front of the tongue (darker, red area). Interestingly, within the vowel /iː/, the three contexts show this patterns, with a retracted tongue back and root that is more marked in the pharyngeal and uvular contexts, than in the pharyngealised. Within the uvular, the concentration of the difference is closer to the uvular (upper pharyngeal) place, whereas it is much lower for the pharyngeal towards the middle pharynx spreading into the lower pharynx. Within the pharyngealised, it starts from the upper pharynx spreading to middle pharynx. A similar picture can be seen for both /aː/ and /uː/. Interestingly, most of the observed differences are located throughout the C2 that goes into V2 (up to the middle). Within V1, the only differences observed are around and after 75% of the vowel. The results suggest overall that the constriction location of "gutturals" is similar and is located within an upper to low pharyngeal that shows a gradient constriction across the contexts. Uvulars show the highest location, followed by pharyngealised and then pharyngeals, with a possibility for a larynx position change, especially in the pharyngeal context (quantified via the estimated secondary constriction location).

The UTI results show an interesting pattern whereby the three members of the guttural class share similarities in the portions of the tongue that are impacted. These are located primarily within the consonant and spreading into the surrounding vowels. The tongue surface and difference plots show similar types of impacts that are located within the front and back areas.

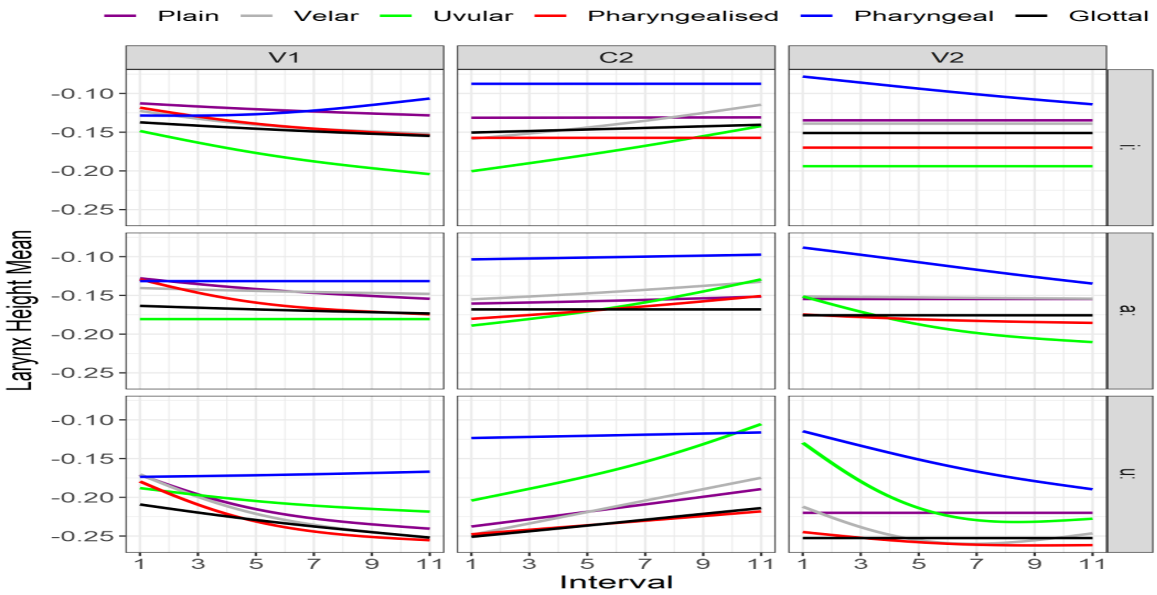
However, the results point to differences between the members of the guttural natural class. The uvular context shows partial "retraction" with a "raised" and "back" gesture following LAM (Esling et al., 2019) and is indeed correlated with a primary {tra} and a secondary {tre} (following Moisik et al., 2019). The pharyngealised context has an intermediate tongue "retraction", with a back and mid-down gesture in /iː aː/, and a back and mid-up gesture in /uː/. The uvular and pharyngealised contexts are different in their constriction location with differences mostly in the degree of tongue rising. Hence, we claim that pharyngealised has only a secondary retraction {tre} but not raising *{tra}. The pharyngeal context shows a near maximal "retraction" with a lowered tongue dorsum and tongue root changes, but with a fronted tongue position. The results point to both pharyngeal and to some extent the pharyngealised and uvular to show possible double bunching in an /iː uː/ contexts, with tongue fronting towards the palatal region and tongue retraction towards the pharynx, in addition to depression of the tongue dorsum in the three contexts; with the following two phonological potential states {tre} and {tdb} (following Esling et al., 2019; Moisik et al., 2019).

4.5.3.3 Electroglottography

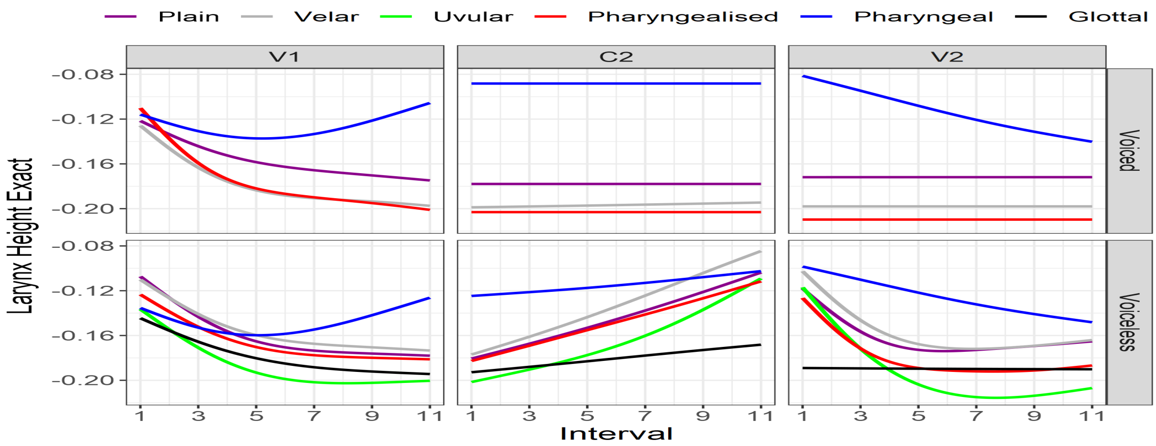
The UTI results presented in the previous section showed how the three contexts: uvular, pharyngealised and pharyngeal, shared similar patterns with respect to the portion of the tongue impacted upon. Clearly, the three contexts show similar lingual gestures, especially within the consonant itself (C2), but also within the following vowel (V2). One of the major claims of LAM is that an epilaryngeal constriction is a combined *lingual* and *laryngeal* gesture. So far, we only looked at the articulatory changes associated with the lingual gesture[19]. In this section, we extend this to the articulatory changes associated with the laryngeal gesture, using non-invasive and an indirect method, quantified via EGG. This part is still work in progress, but initial results were presented in: J. Al-Tamimi (2020a, Invited talk at the SSF, Cambridge, on using the role of the epilarynx in describing guttural consonants in lights of LAM, using UTI, Acoustics and EGG data), J. Al-Tamimi (In preparation, Article JASA presenting results of laryngeal and supra-laryngeal activity in gutturals in the light of LAM). When examining the articulatory correlates for tense, pressed, laryngealised voice and/or constricted glottis, using the EGG, the following measures have been used (for a review, see Kuang & Keating, 2012, 2014):

5.
1) Contact Quotient (CQ; % contact) ⇒ ⇑CQ due to increase in contact in glottis
2) Peak Increase in Contact (PIC; amplitude of positive peak of dEGG Michaud, 2004) ⇒ ⇓PIC due to faster/abrupt closer of glottis, leading to increase in energy in higher frequencies
3) Peak Decease in Contact (PDC; amplitude of negative peak of dEGG Kuang & Keating, 2012, 2014) ⇒ ⇓PDC
4) Speed Quotient (SQ; ratio between contacting/closing and decontacting/openning phases Holmberg, Hillman, & Perkell, 1988) ⇒ ⇓SQ due to shorter contacting/closing phase (CP) compared to longer decontacting/opening phase (OP) in creaky and tense voice
5) Larynx Height (LH) ⇒ ⇑LH due to epilaryngeal constriction, leading to creaky/laryngealised/tense voice

[19] As highlighted above, the use of the estimated constriction location from UTI, allowed to evaluate the changes within the laryngeal area, where pharyngeals were predicted to show raised larynx posture. However, the results are non-conclusive as they are based on model's predictions



(a) VCV by vowel



(b) VCV by voicing

Figure 4.31: Larynx height results across the VCV sequence, with V1, C2 and V2 (in columns) and time intervals on the x-axis (time 1 = onset, 6 = midpoint and 11 = offset): (a) by vowel /iː aː uː/ (in rows); (b) by voicing (in rows) (reproduced from J. Al-Tamimi, 2020a, In preparation)

Based on the list of measures obtained from the EGG signal (see List 5), it is to be hypothesised that a constricted epilarynx would yield: ⇑CQ, ⇑LH and ⇑OP, ⇓PIC, ⇓PDC, ⇓SQ, and ⇓CP. Of course these are the main expected patterns, however, these become more complex to quantify due to how the EGG signal is obtained, to the variable manners of articulation, vowels, and voicing states. In addition, Kuang and Keating (2012, 2014), explained that when dealing with minute differences in Voice Quality states, such as tense vs lax voice quality, changes observed on the EGG measures will be very small. Our aim in this section is to document the minute differences observed between the three contexts which are part of the guttural class.

In terms of Larynx height, which was obtained from the second channel of the EGG, Figure 4.31[20] presents the results of the Larynx Height across the VCV sequence. Starting with

[20]It should be noted that the onset of V1 shows at raised larynx in all cases given that the initial consonant was

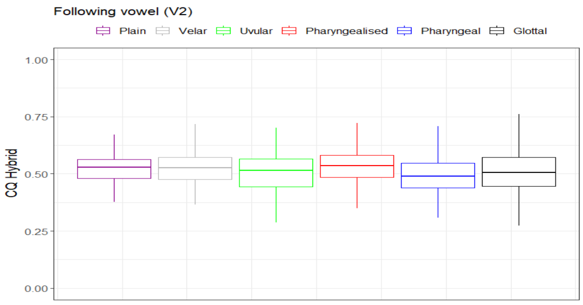
the interaction between the V1, C2 and V2 in the three vowels /iː aː uː/, Figure 4.31a shows an overall raised larynx in /iː aː/ and lowered in /uː/ across all contexts. For gutturals, pharyngeal (in blue) shows an overall raised larynx throughout the VCV in all vowel contexts. Interestingly, uvular stop (in green) shows a decrease in larynx height throughout V1, then an increase throughout C2, reaching its maximum around the release (interval 9-10) and offset of the consonant (interval 11), regardless of the surrounding vowel quality. The pharyngealised context (in red) shows a similar pattern, with a stable/decreased larynx height in the surrounding vowels, but with a steady increase within the consonant, especially within /aː uː/. Looking at the interaction between the VCV sequence and the voicing status (Figure 4.31b), a clearer pattern emerges. The Voiced pharyngeal /ʕ/ (in blue) shows an overall increased larynx height throughout the VCV sequence, with its maximum located within the C2 and onset of the V2. The Voiced pharyngealised consonants /dˤ ðˤ zˤ lˤ/ (in red), show an overall decreased larynx height throughout the VCV sequence, due to participants trying to maintain voicing as much as they can in these productions, leading to a secondary larynx lowering (Westbury, 1983). In the Voiceless sets; the Voiceless pharyngeal /ħ/ (in blue) shows a fall-raising, raising and falling patterns in the VCV sequence. The Voiceless pharyngealised consonants /tˤ sˤ/ (in red), show a falling, raising and falling larynx height throughout the VCV sequence. Finally, the uvular stop /q/ (in green) shows a similar pattern to that of the pharyngealised (falling, raising and falling). The three guttural consonants have their maximum larynx height towards the offset of C2/onset of the V2. Interestingly, when compared to the glottal consonants (in black), the three members of the guttural natural class show raised larynx especially in the second half of the C2 up to the first half of V2. Larynx Height quantified via the second channel of the EGG machine captured these minute changes. It is unclear if the overall changes throughout the VCV sequence should be taken as "template" to how the laryngeal activity is represented in the speech of our participants or if changes within the individual components should be used to assess laryngeal changes. For instance, looking at the Voiceless velar consonants /k x/ (in grey), one can easily identify an extremely raised larynx especially at the release of the consonant (intervals 8-11, within C2), which is much higher than that seen for the Voiceless pharyngeal; is this indicative of an epilaryngeal constriction simply due to the extreme raised larynx? Of course not! An epilaryngeal constriction has two major changes: laryngeal and supralaryngeal changes[21].

These novel results show that the guttural consonants have an overall raised larynx, with the pharyngeal showing the highest larynx position, followed by the uvular and then the pharyngealised. This suggests that the laryngeal gesture in the three contexts is similar, albeit with different degrees of raising.

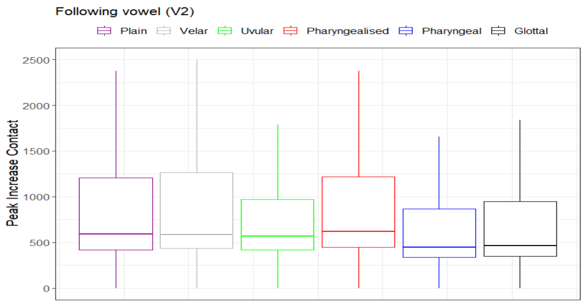
As highlighted so far, most of the differences seem to operate at the onset of the V2. To evaluate the strength of these differences, Figure 4.32 shows various EGG measures obtained from EGGWorks, in addition to results obtained from a Principal Component Analysis (PCA) and from Random Forests. Figure 4.32a shows that the highest CQ is within the pharyngealised context, with the lowest in the pharyngeal. This is a clear indication that the pharyngealised consonants induce a constriction in the glottis and not specifically a raised larynx position. The PIC results (Figure 4.32b) shows that the uvular, pharyngeal and glottal have an overall lower

always a /ʔ/ with a default raised larynx.

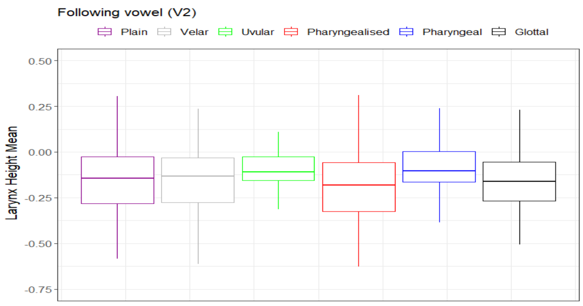
[21] It is likely that the release of /k/ is produced in an ejective-like manner, inducing an increase larynx height; for the velar /x/, one can question whether it should be included as part of the "uvulars" regardless of whether it is produced as a velar or uvular /χ/. These two observations can possibly explain the increase in larynx height observed in this study.



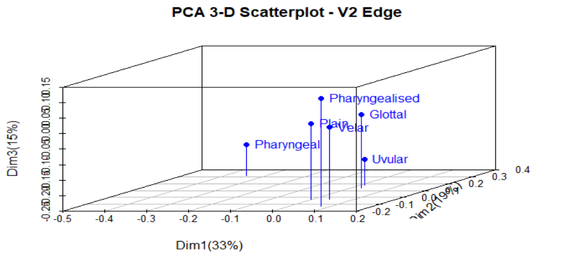
(a) CQ hybrid



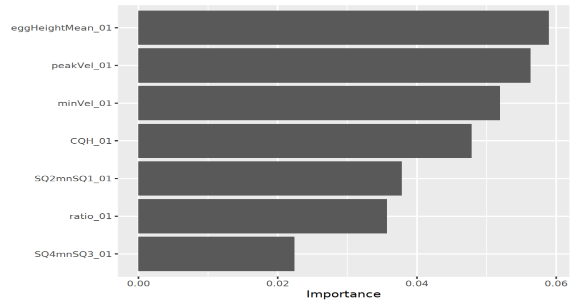
(b) Peak Increase in Contact



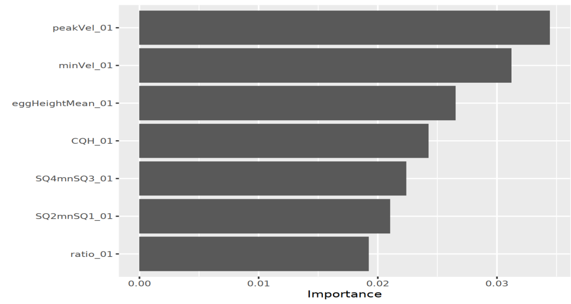
(c) Larynx Height



(d) PCA six contexts



(e) Variable Importance - Six contexts



(f) Variable Importance - Guttural vs Non-Guttural

Figure 4.32: Various averaged results for EGG at onset of V2: (a) CQ hybrid; (b) Peak Increase in Contact; (c) Larynx Height; (d) 3D PCA using all EGG predictors; (e) Variable Importance Scores across the six contexts; (f) Variable Importance Scores comparing guttural vs non-gutturals (reproduced from J. Al-Tamimi, 2020a, In preparation)

PIC, likely indicative of a faster/abrupt closer of the glottis; the pharyngealised has a relatively higher PIC values. The averaged results for the Larynx Height (Figure 4.32c) confirm again that both uvular and pharyngeal consonants induce an raised larynx at the onset of the V2; the pharyngealised induce a lowered larynx. It should be noted however that the error bars for the pharyngealised context span a wide range; due to variable larynx heights influenced by the Voicing state and the surrounding vowels (see Figure 4.31b presented previously).

We then assessed the impact of the combination of the seven EGG measures obtained here: CQ, PIC, PDC, CS, OS, SQ and LH using PCA and RF. Figure 4.32d shows a 3D PCA plot. The first three dimension of the PCA explained around 67% of variance, with mixture of correlates. The results point to the uvular, and glottal to be clustered together with the pharyngeal context being separate from everyone else. The plain, velar and pharyngealised are clustered together, however, the pharyngealised seems to be closer to the glottal context on the third dimension. Using Random Forests as a classification tool, our results show a rate of 50% separating the six contexts; this increases to 71% when comparing the guttural to the non-guttural classes. This rate is relatively important but is clearly indicative of laryngeal changes to be secondary. Variable Importance scores on the six contexts (Figure 4.32e) suggests that the Larynx Height is the main predictor to distinguish the consonants, followed by the PIC (peakVel), PDC (minVel) and CQ. For the guttural vs non-guttural (Figure 4.32f), the main correlate is the PIC (peakVel), PDC (minVel), followed by LH.

The patterns reported here will be explored further as it is clear that the Voicing state impacts on some of the patterns reported. However, these results point for the first time to systematic laryngeal changes associated with gutturals. These need to be considered as secondary due to the relatively low classification rates obtained. These results seem to correlate with [CG] (for "Constricted Glottis") as a secondary feature, with variable {↑lx} results.

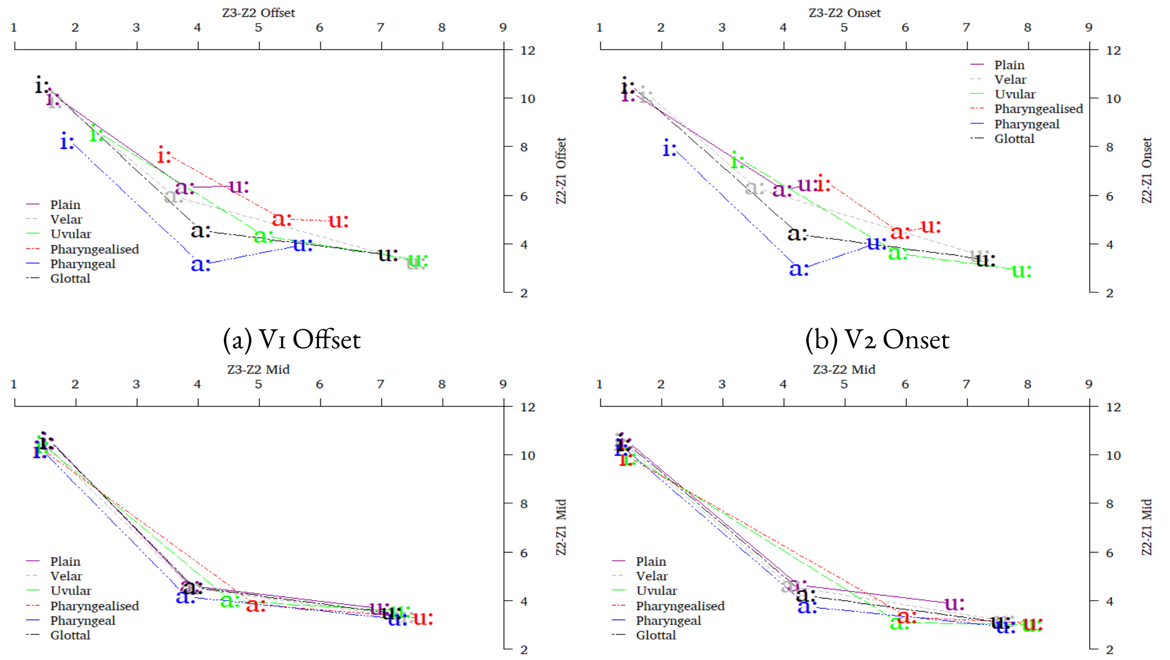
4.5.3.4 Acoustics - Laryngeal and Supra-Laryngeal

So far, we looked at how guttural consonants show similarities in the patterns reported using UTI and EGG. The UTI results pointed to general retraction of the tongue dorsum, back, and root, with depression of the tongue dorsum/front especially in /iː uː/ environments. From the EGG, the Voiceless gutturals showed raised Larynx Height from the middle of the C2 to the midpoint of the V2. The CQ and PIC results were also suggestive of specific patterns related to increased CQ (but decreased PIC) in the pharyngealised; the reverse for uvular and pharyngeal.

Given these patterns and to correlate with the findings reported in J. Al-Tamimi (2017b), we explored the data further from an acoustic point view, by using VoiceSauce with Praat's algorithm to obtain formant frequencies as Bark-Difference metrics, in addition to Voice Quality measures used in the psychoacoustic model of voice quality (Garellek, 2019, 2020; Garellek, Samlan, Gerratt, & Kreiman, 2016; Kreiman, Lee, Garellek, Samlan, & Gerratt, 2021). The results presented in this section were available in the following: J. Al-Tamimi (2019c, Invited talk at the SRPP, LPP, Paris, on using GAMMs with UTI data and initial acoustic results in guttural consonants in lights of LAM), J. Al-Tamimi (2019b, Invited talk at the LIMSI, Paris, on distinguishing between male and female speakers using guttural consonants), J. Al-Tamimi and Ferragne (2020a, BAAP on using deep learning and random forests to distinguish between guttural consonants using acoustics), J. Al-Tamimi and Ferragne (2020b, 17^{th} LabPhon on using deep learning and random forests to distinguish between guttural consonants using acoustics), J. Al-Tamimi (2020a, Invited talk at the SSF, Cambridge, on using the role of the epilarynx in describing guttural consonants in lights of LAM, using UTI, Acoustics and EGG data), J. Al-Tamimi (2021a, 4^{th} PaPE presenting the acoustic profile of gutturals using random forests), and J. Al-Tamimi (In preparation, Article JASA presenting results of laryngeal and supra-laryngeal activity in gutturals in the light of LAM).

Figure 4.33 shows a scatterplot for each of the V1 (offset and midpoint) and V2 (onset and midpoint) using Bark-Difference metrics that were identified as the most predictive in J. Al-Tamimi (2017b). Looking at the vowels' midpoint (Figures 4.33c for V1 and 4.33d for V2), there are minor differences between the pharyngeal, plain, velar, and glottal, while both uvular and pharyngealised show reduced Z3-Z2; they are both correlating with a more back production and with spectral divergence that is more marked within V2).

At the vowel edges (Figures 4.33a for V1 and 4.33b for V2), pharyngeal consonants show the



(a) V1 Offset (b) V2 Onset

(c) V1 midpoint (d) V2 midpoint

Figure 4.33: Bark-difference patterns using Z3-Z2 (x-axis) and Z2-Z1 (y-axis) across the six contexts at vowel edges (top row) and midpoint (bottom row) and within V1 (left column) or V2 (right column): (a) V1 at the offset; (b) V2 at the onset; (c) V1 at the midpoint, (d) V2 at the midpoint (reproduced from J. Al-Tamimi, 2019c, In preparation)

most retracted quality with a lower Z2-Z1, pharyngealised shows the least retracted quality with higher Z2-Z1, especially within V1 (offset), but with a much lower Z2-Z1 at the V2 onset. The uvular shows intermediate retraction via median Z2-Z1 and similar backness/divergence to that of the pharyngealised context. These formant frequency patterns are a clear indication of a difference in the constriction location across the three contexts, with pharyngeals showing a much lower location, uvular a higher up and pharyngealised in-between; results correlating with the predictions and results obtained previously in J. Al-Tamimi (2017b).

Next, we explored two machine learning algorithms: Convolutional Neural Networks (CNNs) and Random Forests via ExtraTrees, to understand the patterns in the data. Starting with the CNNs, J. Al-Tamimi and Ferragne (2020a, 2020b) used each of the /ˈʔVːˈCVː/ words and generated spectrograms for each, which were converted to 8-bit greyscale images and resized to 257 (vertical) by 800 (horizontal) pixels (1 pixel = 1.15 ms and 35.71 Hz, for more details, see Ferragne, Gendrot, & Pellegrini, 2019). The spectrograms were then fed into the CNN, which tries to identify specific recurrent patterns in these images. After training, tuning and validating the models, we obtained classification accuracy results. Our model achieved a very high classification accuracy: Within the six classes, we achieved an 86.9% accuracy and an 87% on discriminating between the two classes: gutturals vs non-gutturals. The confusions in the former were generally within the gutturals themselves with some instances confused with the other categories (especially the plain and glottal). To further evaluate the results, an occlusion analysis was performed, using multiple mask sizes and strides: (a) 25% (51-40 and 26-20 pixels); (b) 50% (51-80 and 26-40 pixels) and (c) 100% (default; 51-160 and 26-80 pixels).

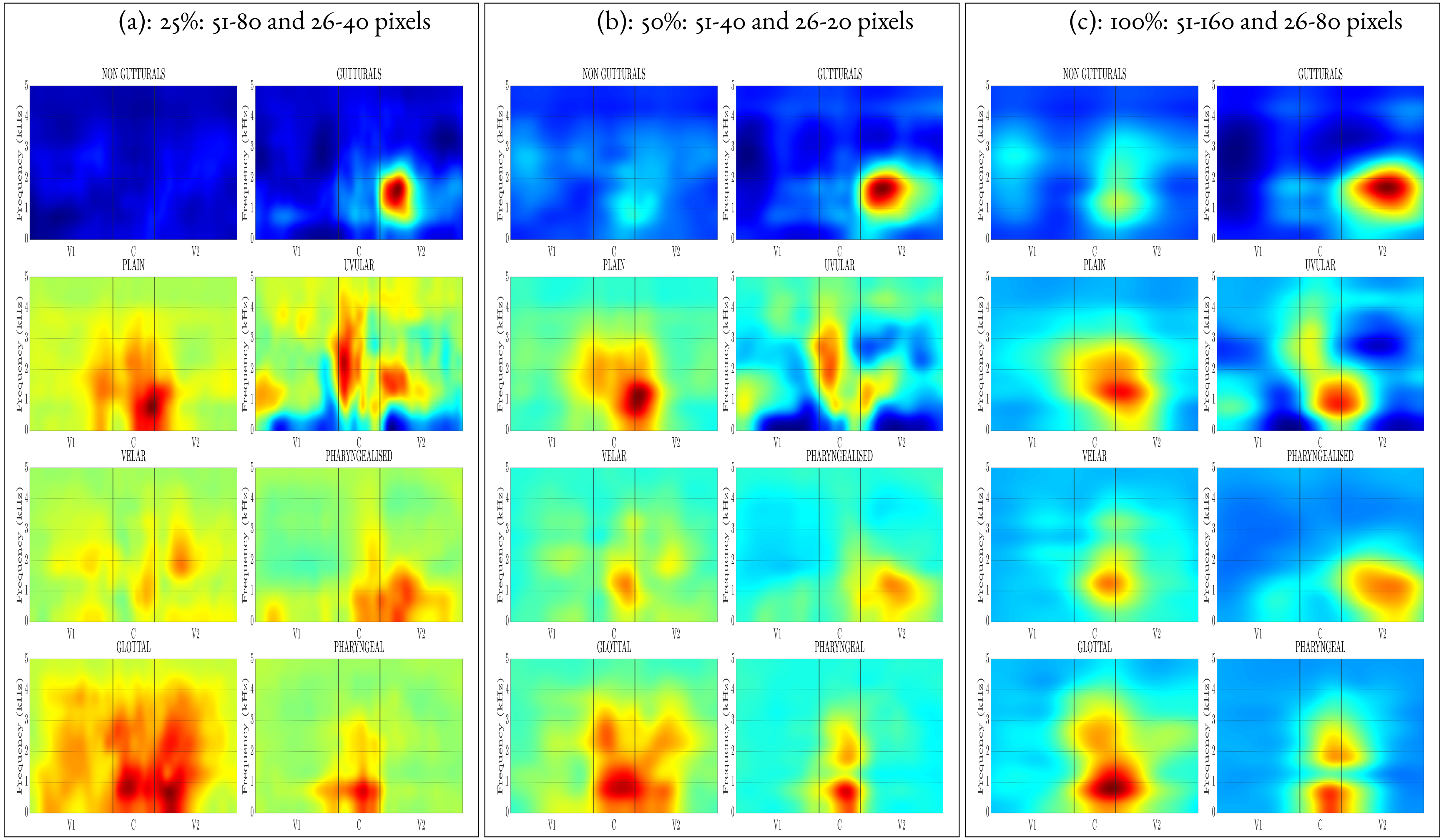


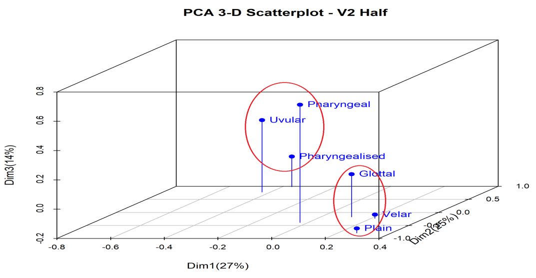
Figure 4.34: Regions of interest identified based on occlusion analysis on the greyscale spectrogram images for each /ˈʔV:ˈCV:/ word fed to a Convolutional Neural Network (CNN), with multiple mask sizes and strides: (a) 25% (51-40 and 26-20 pixels); (b) 50% (51-80 and 26-40 pixels) and (c) 100% (default; 51-160 and 26-80 pixels), for either the guttural vs non-guttural classes or each of the six categories. Activation colour: blue = least, yellow = partial, red = maximal (reproduced from J. Al-Tamimi & Ferragne, 2020b)

Figure 4.34 shows the results for each of the two classes (guttural vs non-guttural) or the six classes (Activation colour: blue = least, yellow = partial, red = maximal). When comparing the occlusion analyses across the three types of mask sizes and strides, it is clear that any of the three provide very similar patterns, however, the mask size and stride of 25% (51-40 and 26-20 pixels) showed the most precise activations. Non-gutturals show a mixture of regions in V1, C2 and V2; within gutturals, uvular and pharyngealised show similar activation regions, whereas pharyngeal is mostly within C2. When evaluating the differences between the two classes, it is evident that the guttural shows a predominance in activations located in low-mid frequency around 1-3 kHz within V2 (onset to mid). These results point to the members of the guttural class to share acoustic similarities located towards the onset to midpoint of the V2.

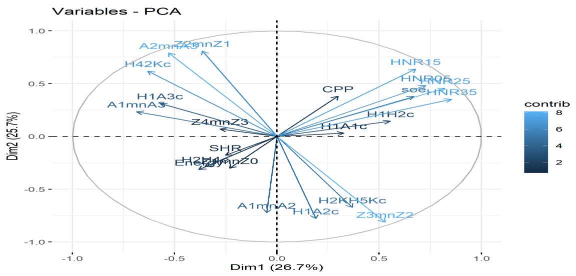
Given than the CNN results show clear patterns located within the V2, we evaluated this further using PCA and Random Forests grown via ExtraTrees. Figure 4.35 shows the results presented in J. Al-Tamimi (2021a, see as well J. Al-Tamimi, In preparation). Using the combination of 22 Acoustic measures (4 bark-difference, 10 voice quality and 8 energy/noise correlates), we first used PCA to evaluate the clustering of the data. Figure 4.35a shows a 3D of the PCA, where the first 3-dimensions explained around 65% of the variance. It is clear that the two classes are well separated, but not perfectly. Within gutturals, uvular and pharyngealised are clustered together, with pharyngeal separate but close to them. Within non-gutturals, the plain, velar and glottal are clustered together; glottal is close to pharyngeal. Figure 4.35b shows the contribution of each of the predictors. There were variable contributions from predictors, with the strongest being spectral compactness and divergence (Z2-Z1 and Z3-Z2), noise measures (HNR, 0.5 to 3.5kHz), energy above F1 and around F3 ($A2^*$-$A3^*$ and $H4^*$-$H2\text{kHz}^*$). Moving next to the Random Forests results, and when comparing classifications for both V1 and V2 using the six categories and the two classes (guttural vs non-guttural), the rates presented in Figure 4.35c show the highest rates within V2, with 87% for the two classes; 69% for the six. Using both Bark-difference and Voice Quality measures showed the highest rates; with 80% achieved for Voice Quality only on the two classes. This is interesting as the rates here are much higher for Voice Quality than those reported in J. Al-Tamimi (2017b), which were close to 72% for an urban eastern dialect, on pharyngealisation vs plain only. Figure 4.35d shows the Variable Importance Scores for the top 10 predictors. The most influential predictors are: spectral compactness/divergence, openness, and pharyngeal location (Z2-Z1, Z3-Z2, Z1-Z0 and Z4-Z3). The energy above F1 and around F3 ($A2^*$-$A3^*$, $H4^*$-$H2\text{kHz}^*$, $A1^*$-$A3^*$) were secondary. Finally, Figure 4.35e shows how each of the top 10 predictors are different across gutturals and non-gutturals.

The results point to both supra-laryngeal and laryngeal changes to explain the patterns in the data. For supra-laryngeal, it is a primary spectral compactness (⇓Z2-Z1), divergence (⇑Z3-Z2), openness (⇑Z1-Z0) and pharyngeal constriction (⇓Z4-Z3), and for laryngeal, it is secondary decrease in spectral tilt with increased energy above F1 and around F3 (⇓$A2^*$-$A3^*$, ⇓$H4^*$-$H2\text{kHz}^*$, ⇓$A1^*$-$A3^*$, ⇓$H1^*$-$A3^*$) and decreased energy around F2 (⇑$H1^*$-$A2^*$ and ⇑$A1^*$-$A2^*$). Interestingly, these primary and secondary differences located around 1-3kHz match those observed with the CNN results presented above.

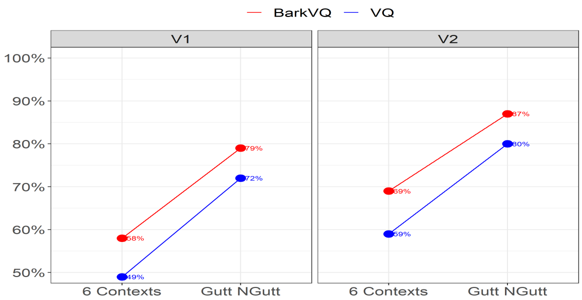
The results combining UTI, EGG and acoustics provide complementarity in the patterns reported and are correlated with each other. They are indicative of a primary retraction of the tongue back and root, depression of the tongue dorsum (quantified via UTI), with a raised larynx



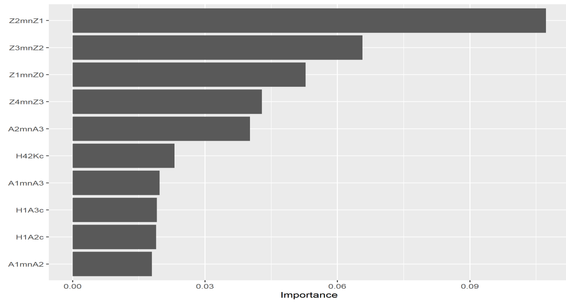
(a) 3D PCA - classes



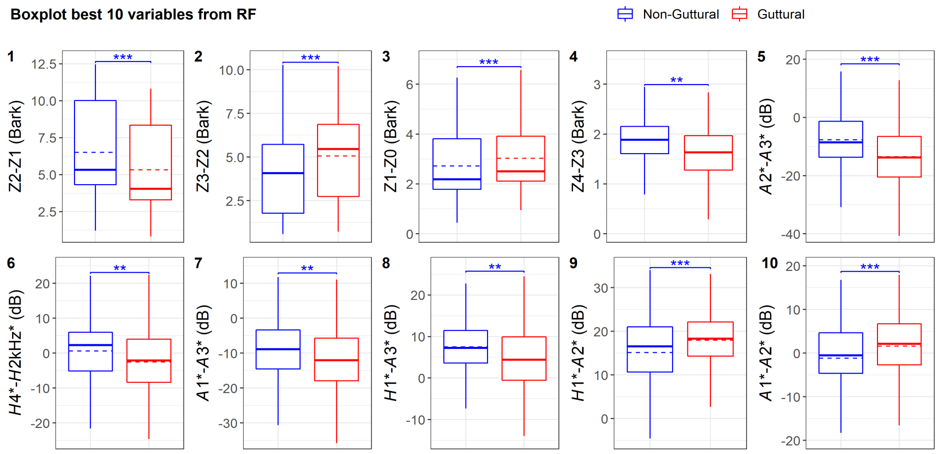
(b) 2D PCA - predictors



(c) Classification within V1 and V2



(d) Variable Importance Scores - V2



(e) 10 best predictors

Figure 4.35: Acoustic results for guttural vs non-guttural based on the first half of the V2: (a) 3D PCA showing clustering of each of the six classes; (b) 2D PCA showing the variable loadings on the first two dimensions; (c) Classification accuracy across V1 and V2; (d) Variable Importance Scores for the first half of V2; (e) Box-plots and statistical significance of the top 10 predictors to the guttural vs non-guttural contrast (reproduced from J. Al-Tamimi, 2021a, In preparation)

especially in the Voiceless gutturals from the middle of the C2 up to the middle of the V2, with a more constricted larynx in the case of the pharyngealised and with an abrupt/fast closure of the larynx in the case of the uvular and pharyngeal (from the EGG). From an acoustic point of view, the complementarity between the CNN and the Random Forests results are an indication of robustness of the reported patterns. Clearly, the first half of the V2 shows most of the impacts of the gutturals, with progressive coarticulation being more visible. Voice Quality measures used in this study allowed to differentiate the two classes with 80% accuracy; this increases to 87% when they are combined with Bark-Difference measures. Again, supra-laryngeal changes are primary with spectral compactness and divergence as main correlates. This is the first time we report on the measure Z4-Z3 to be an important correlate for the guttural consonants with a ⇓Z4-Z3 correlating with [+back] and/or [+retracted] especially in upper-pharyngeal/uvular and extreme epilaryngeal constriction (Fant, 1960; Johnson, Sherman, & Sherman, 2012; Story, 2019). Laryngeal changes quantified via acoustic analyses are again indicative of specific changes associated with gutturals. These are to be seen as gradient rather than categorical.

4.5.3.5 Formal account

The results of this study contributed articulatory and acoustic evidence of how guttural and non-guttural consonants behave in this variety of Arabic. We examined both supra-laryngeal (UTI + acoustics) and laryngeal (EGG, partially UTI and acoustics), which allowed the examination of lingual and laryngeal changes associated with the two classes. It is clear from the results that gutturals affect the tongue contour in a relatively similar fashion across all of its members, which we identified as being uvular, pharyngealised and pharyngeal, with glottals being identified as not belonging to gutturals due to previous literature failing to identify similarities with gutturals, and by identifying that glottals do not excerpt any major impact on tongue contours, which correlates well with the predictions of LAM to constider them as true glottals. The similarities and differences between members of the guttural class can be formalised by combining the predictions of the "Phonological Potentials" of lingual states (Esling et al., 2019; Moisik et al., 2019) and by using the formal accounts identified in Esling et al. (2019); Moisik et al. (2012, 2019); Moisik and Esling (2011). From supra-laryngeal and lingual changes point of view (i.e., based on UTI and acoustic results), it is clear that the three members use various lingual states, which are either similar or different to each other. Clearly, the uvular stop uses both {tra} as primary and {tre} as secondary, which follows the predictions of the Phonological version of LAM (see Esling et al., 2019, pp. 179-180). Pharyngeal has a clear primary {tre} and secondary {tfr} due to the fact that it is produced with tongue double-bunching {tdb}. The auditory quality and acoustic results of the vowel /aː/ surrounding the pharyngeal /ħ ʕ/ in Levantine Arabic shows patterns similar to that of an [æː], which can be used to identify a pharyngeal rather than an epilaryngeal articulation for these consonants Moisik (2013a, Section 7.3.1); Moisik et al. (2019, pp. 24) and Esling et al. (2019, pp. 171). Finally, pharyngealised consonants have a primary {tfr} (due to them being primary coronals) and secondary {tre}; which resembles the cases of true double bunched consonants (e.g., retroflex consonants /ɻ/) and matches what Jakobson (2002); Jakobson et al. (1952/1976) identified as a similar acoustic impression of flatness associated with labialised, retroflex, and pharyngealised consonants, and which were assigned the feature [+Flat].

From Laryngeal changes point of view (EGG, partially UTI and acoustic results), pharyngeal consonants in Levantine Arabic are associated with an increase in Larynx Height {↑lx} with an overall decrease in spectral tilt and with an increase in energy in higher frequencies. Uvular stop shows a similar patterns of a raised Larynx {↑lx}, with a decrease in spectral tilt and an increase in energy in higher frequencies. The two contexts show small differences as the raised larynx posture is inherent to the Voiced and Voiceless pharyngeals and is seen throughout the VCV sequence, whereas within the uvular stop, this raised Larynx position is only seen at the release of the consonant. Based on the PIC results, both uvular and pharyngeals show an overall lower value likely indicative of a faster/abrupt closer of the glottis possibly due to the High Larynx position. For pharyngealised, the picture is more complex. An increase in Larynx Height {↑lx} is only seen in the Voiceless set at the release, where the larynx follows a {↓lx} in V1, {↑lx} in C2, and a {↓lx} in

V2. The relatively high CQ and high PIC values are indicative of a more constricted glottis; results which match those previously identified in J. Al-Tamimi (2017b). From an acoustic point of view, pharyngealised contexts induce an overall decrease in spectral tilt and an increase in energy in higher frequencies. The acoustic profile of the members of the gutturals as a natural class is similar, albeit with small differences to be accounted for.

We used the feature [±cet] (for "constricted epilaryngeal tube" Moisik et al., 2012; Moisik & Esling, 2011) to account for the specific changes identified above. The feature [±cet] is used to describe the combined gesture of tongue body and root retraction in addition to voice quality changes causing a more laryngealised, pressed, tense, or creaky voice caused by an epilaryngeal constriction. The feature [cet] is composed of the double features: [retracted] and [constricted glottis]. The former highlights the fact that the vowel is [+Open] and [+Back] and correlates well with the [RTR]. The latter correlates well with laryngeal changes associated with the larynx state (and assumed to encompass Larynx Height). The List 6 provides a feature account for gutturals using [±cet] taking into account the UTI, EGG and acoustic results of the current study.

6. 1) Uvular ⇒ **[−cet]** ([+Retracted], [+Raised] and [−Open], [+Constricted Glottis])

   2) Pharyngealised ⇒ **[+cet]** ([+Retracted], [−Raised] and [+Open], [+Constricted Glottis])

   3) Pharyngeal ⇒ **[+cet]** ([+Retracted], [−Raised] and [+Open], [+Constricted Glottis])

As highlighted in this list, it is clear that the binary feature specification causes issues. For instance, the three members of the guttural class are specified for the feature [+retracted] due to either tongue body and/or root retraction; the feature [+constricted glottis] highlights that the three contexts have Larynx raising (especially for the Voiceless sets) and/or constriction. The specifics of [±cet] set follow from the predictions of LAM given that uvulars are to be assigned a [−cet], whereas pharyngeal [+cet]. Pharyngealised have clearly a different profile to pharyngeals or uvulars, and based on our own predictions, should receive the feature [+cet].

However, and as advocated in our previous accounts, we follow a gradient privative features account advocated by J. Beckman et al. (2013) and employed by J. Al-Tamimi and Khattab (2018a, 2018b). Here we claim that the feature [cet] is active in Arabic gutturals and will receive different numeric notations explaining the gradient nature of epilaryngeal constriction (see List 7: 0 = inactive; 1 = minimal; 9 = maximal).

7. 1) Uvulars ⇒ **[2cet]** ([4Front], [2Retracted], [0Open], [5Constricted Glottis], [5Raised Larynx])

   2) Pharyngealised ⇒ **[5cet]** ([8Front], [5Retracted], [3Open], [7constricted glottis], [2Raised Larynx])

   3) Pharyngeal ⇒ **[7cet]** ([7Front], [7Retracted], [6Open], [5Constricted Glottis], [9Raised Larynx])

This gradient feature specification allows to account for the minute differences observed in our study, by identifying that our pharyngeals are true pharyngeals and not epilaryngeals (hence the [7cet]); that uvulars induce a minimal epilaryngeal constriction (hence the [2cet]) and that pharyngealisation is in-between (with a [5cet]). The [Front] and [Retracted] features are used to highlight the potential double-bunching identified in our UTI data, especially for uvulars and pharyngealised and for the [æ]-like quality of the vowels surrounding our pharyngeal consonants. Accounting for these changes in the light of the Phonological Potentials of LAM will be explored in future research.

4.5.3.6 Extensions to guttural and epilaryngeal constriction

The role of epilaryngeal constrictions and changes in the source associated with guttural consonants continued to be a major area of research in which I continued exploring. In the following sections, I provide a summary of two complementary studies either from my own research or as part of PhD supervision[22].

4.5.3.6.1 Dialect and Speaker Identification

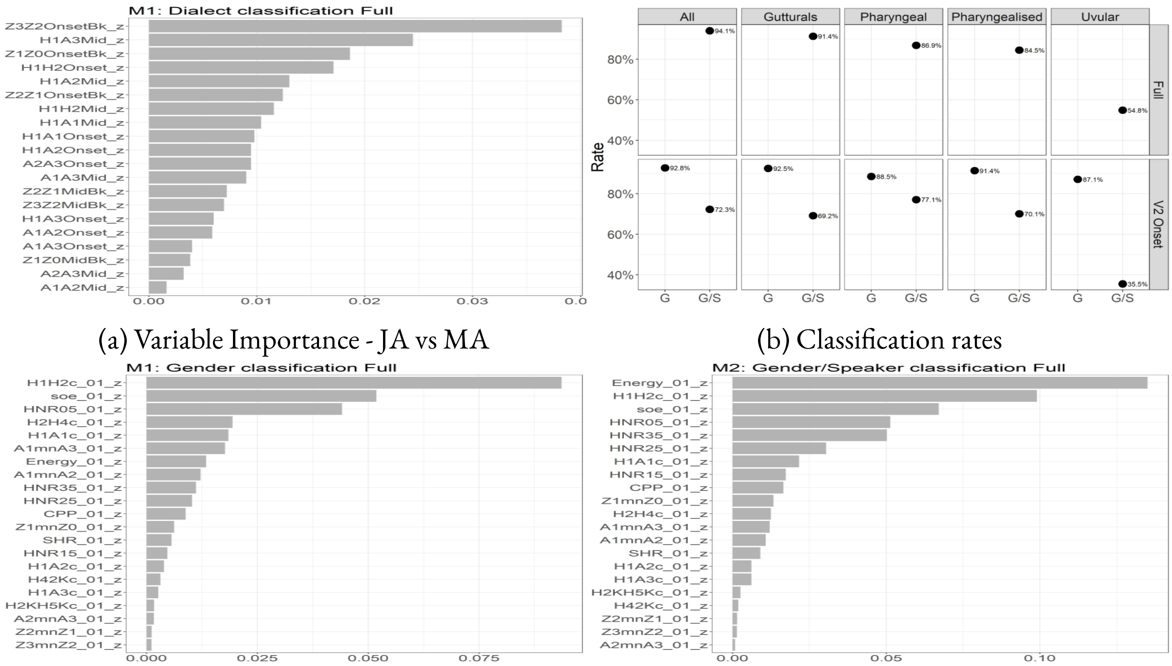
Previous research highlighted that cross-dialectal differences are well marked and can allow to distinguish Arabic dialects. Bellem (2007) explained that gutturals are usually different between Arabic dialects; with Bedouin dialects showing more guttural quality than Urban dialects. However, less research was undertaken to quantify this type of difference and it is unclear what was meant by this more or less guttural quality. We saw earlier that the constriction in pharyngealised consonants is usually different and vary between post-velar to mid-pharyngeal (J. Al-Tamimi, 2017b; Bellem, 2007; Khattab et al., 2006). Within pharyngeals, place and manner differences are important cross-dialectally, with approximants forming the majority of cases, but also stop and constricted in the voiced variant (Heselwood, 2007). Hence we can predict that some dialects will show an epilaryngeal constriction (lower pharyngeal), such as in Bedouin dialects, Western Dialects, or Iraqi dialects (J. Al-Tamimi, 2017b; Alsiraih, 2013; Heselwood, 2007) and pharyngeal (mid-pharyngeal) constriction that correlates well with an approximant realisation in more Urban, and Eastern dialects (J. Al-Tamimi, 2017b; Heselwood, 2007). These differences can also vary with sociolinguistic, gender and regional differences.

In relation to gender differences, it is well-known that physiological differences between males and females are mostly related to difference in the length and morphology of the vocal tract and size of vocal folds impacting on formant frequencies and $f0$ patterns (see Hanson, 1997; Hanson & Chuang, 1999). Female speakers are more likely to have an incomplete closure of the vocal folds leading to an increase in airflow (Hanson, 1997), and to an increase in perceived breathiness (Hanson, 1997; Klatt & Klatt, 1990) with a flatter/steeper spectrum and lower spectral tilt than male speakers. In Arabic, researchers showed that female speakers tend to have a "softer" realisation of gutturals. Pharyngealised consonants tend to be produced as "de-pharyngealised" and vowels are more palatalised, with a possible loss of the secondary constriction (Haeri, 1991, 1992, 1994; Royal, 1985), with longer VOT patterns in pharyngealised consonants (Khattab et al., 2006).

Given the cross-dialectal and sex-specific difference, the work presented in J. Al-Tamimi (2019c) within this area looked specifically at the case of pharyngealised consonants allowing the distinction between Jordanian and Moroccan Arabic (by revisiting the patterns reported in J. Al-Tamimi, 2017b) and by examining how guttural consonants can support gender specific differences in Levantine Arabic (by revisiting the patterns reported in J. Al-Tamimi, 2021a).

Starting with the first experiment, Random Forests classification results allowed to separate the vowels produced in the pharyngealised coronal stop in Jordanian Arabic from those in

[22] In addition to the work by my PhD student Alsiraih (2013), I am currently supervising another PhD, as part of a continued supervision agreement with Newcastle University, UK, on UTI analysis of the guttural natural class in Levantine Arabic using a complementary approach with specific landmarks, VC to CV coarticulatory patterns and using classification techniques (see Alsharif, In preparation)



(a) Variable Importance - JA vs MA
(b) Classification rates
(c) Variable Importance - Gender
(d) Variable Importance - Gender/speaker

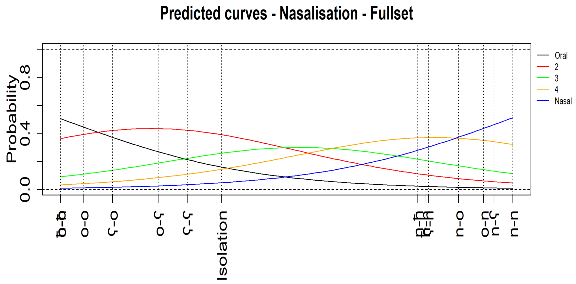
Figure 4.36: Dialect and gender classification results: (a) Variable Importance score in the discrimination between JA and MA; (b) Classification rates in gender or gender/speaker discrimination; (c) Variable Importance score in gender discrimination; (d) Variable Importance score in gender/speaker discrimination (reproduced from J. Al-Tamimi, 2019c)

Moroccan Arabic with a relatively high accuracy of 83.9%. 90.8% of JA vowels were classified correctly, whereas only 73.8% of Moroccan Arabic vowels were classified correctly. Figure 4.36a shows the Variable Importance Scores out of this classification task and the results point to a mixture of bark-differences and harmonic differences as main correlates. Clearly, the most important voice quality measures were harmonic differences in low and high frequencies accounting for the differences; these are related to creaky voice (*H1*-H2**) and abrupt closure of glottis (*H1*-A2**, *H1*-A3**).

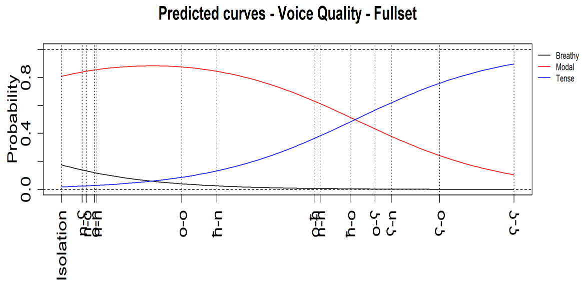
We move next to gender discrimination. Multiple models run on Gender or Speaker by Gender grouping and the results are displayed in Figure 4.36b. The results point to the highest rate to be with Gender on either Full intervals or V2 Onset only; gutturals as a class led to increased discrimination between males and females. Gutturals yielded the second highest rates on Gender (V2 Onset) or Full intervals for Speaker/Gender. As can be seen from the figure, smaller reductions in the rates are observed when using Pharyngeal or Pharyngealised contexts; the Uvular context leads to the lowest rates in Speaker/Gender in either Full or V2 Onset. The Variable Importance scores for gender (Figure 4.36c) or gender/speaker (Figure 4.36d) show that voice quality metrics are the best to allow distinction between participants' productions, leading to a clear indication of individual variability and stability in how sounds are produced, and this is evidenced either in the local (V2 Onset) or long domain coarticulatory patterns. Clearly, the (*H1*-H2**) metrics widely reported to allow distinction between males and females (Simpson, 2012) is the top predictor in our case followed by various HNR measures in addition to strength of excitation (soe), which allows to differentiate between stop/approximant production, but also strong or weak productions in general (Mittal, Yegnanarayana, & Bhaskararao, 2014; Vijayan & Murty, 2014).

4.5.3.6.2 Nasalisation and epilaryngeal constriction?

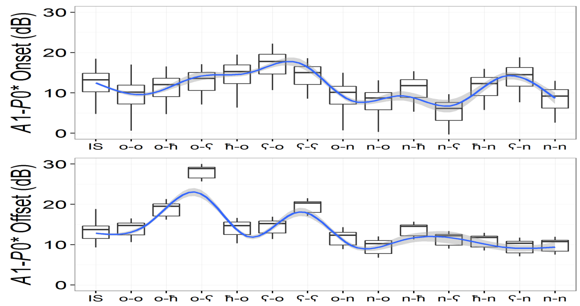
While being interested in the role of Voice Quality in guttural consonants, the work done by Wasan Alsiraih during her PhD, who was my first PhD student at Newcastle University working with me on the interaction between pharyngeal consonants and nasalisation in Iraqi Arabic (Alsiraih, 2013), with a paper based on her findings (Khattab, Al-Tamimi, & Alsiraih, 2018). Our research showed that due to the extreme epilaryngeal constriction in Iraqi Arabic (see as well Hassan & Esling, 2011), epilaryngeals in are produced with an increased percept of nasality, in the vicinity of nasal consonants. This original study provided empirical evidence for this interaction using both an impressionistic auditory analysis and acoustic analyses of nasality and voice quality features to evaluate whether the percept of nasality seen here is influenced by an increase in breathiness; in the productions of the male subjects participating in this study. The results of this study are presented in Figure 4.37.



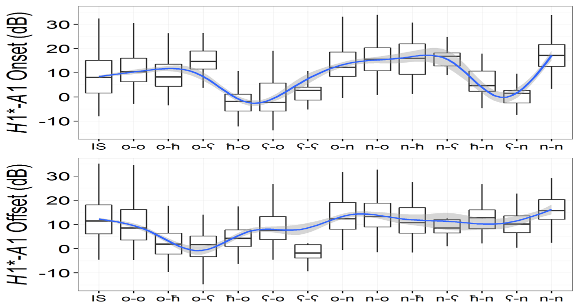
(a) CLMM - Nasality rating



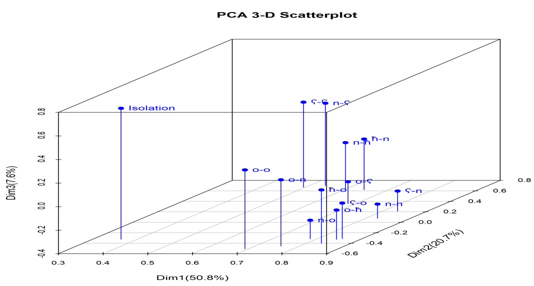
(b) CLMM - VQ rating



(c) $A1^*$-$P0^*$ at onset/offset



(d) $H1^*$-$A1^*$ at onset/offset



(e) PCA loading contexts

Figure 4.37: Results of nasalisation and epilaryngeal constriction in Iraqi Arabic pharyngeals: (a) Predicted curves for nasality rating; (b) Predicted curves for VQ rating; (c) $A1^*$-$P0^*$ at onset/offset of vowels; (d) $H1^*$-$A1^*$ at onset/offset of vowels; (e) 3D PCA scatterplot of contexts (reproduced from Khattab et al., 2018, pp. 327, 329, 331, 336 and 338)

First the productions of the dataset were evaluated by six phonetically trained raters (including the first two authors; with four non-native speakers of Arabic, but who had training on Voice Quality assessment). Figures 4.37a and 4.37b show the predicted curves based on a CLMM (Cumulative Logit Mixed-effects Models) of the ratings 1 to 5 (for nasality) or 1 to 3 (for Voice Quality). Figure 4.37a shows that the /n-n/ context received the highest ratings for nasality (i.e., ratings 4 or 5), which was followed by: /n-ʕ/, /o-n/, /n-o/, /ʕ-n/, /ħ-n/ and /n-ħ/[23]. When using the Isolation context (i.e., vowels produced in isolation) or any of the combination of oral consonants or /ħ ʕ/ in combination with an oral context, an increase in ratings of oral (i.e., ratings 1 or 2) were observed. Looking at the Figure 4.37b, it is clear that when the vowel was in an /ʕ-ʕ/, the percept of "tense" voice received over 80%. Most of the nasal contexts (i.e., nasals at the onset or coda of syllable in combination with an oral, nasal or pharyngeal) received ratings as "Modal". Only a few vowels in isolation received 20% ratings of "Breathy". This auditory analysis confirmed that /ħ ʕ/ in Iraqi Arabic can be safely described as more epilaryngeal than pharyngeal, however, Alsiraih (2013) showed that the realisation of these consonants varied between compressed, pressed, fricative, stop or a tight approximant (similar to the category reported in Heselwood, 2007).

From an acoustic point of view, and using multiple nasality and voice quality measures, we report only on the two widely used measures $A1^*$-$P0^*$ at onset/offset (Figure 4.37c) and $H1^*$-$A1^*$ at onset/offset (Figure 4.37d). Pharyngeal consonants are associated with an increase in nasality due to an epilaryngeal constriction, which seems to impact heavily on the higher frequencies. Figure 4.37c shows that when combined with a nasal context, our epilaryngeal consonants show a decrease in $A1^*$-$P0^*$ similar to when nasals are on their own with values lower than 10dB (following M. Y. Chen, 1995, 1997). When they are in an oral context, $A1^*$-$P0^*$ is over 10dB. This first result is indicative or an increase in the acoustic correlate of nasality and correlates with the auditory results (seen in Figure 4.37a). Moving on to Figure 4.37d, the results point to the pharyngeals in the vicinity of oral contexts to show a decrease in spectral tilt quantified via the measure $H1^*$-$A1^*$, to values around 0dB or even below. When pharyngeals are in a nasal context or when comparing oral-oral contexts, an increase in $H1^*$-$A1^*$ at onset/offset to 10 or 20 dB is observed; which correlates well with either modal or breathy voice. Clearly, our pharyngeals are associated with increase in tense/creaky voice and they induce an increase in nasality in the vicinity of nasals.

Finally, to make sense of the complex acoustic patterns reported in Khattab et al. (2018), we used Principal Component Analysis. Figure 4.37e shows how the different contexts are clustered in this 3D PCA. The results point to our pharyngeal contexts to be closer to the nasal contexts (e.g., /ʕ-ʕ/ close to /n-ʕ/; /ħ-n/ close to /n-ħ/; /ʕ-n/ close to /n-n/) or to oral contexts (e.g., /ʕ-o/ close to /ħ-o/; /ħ-n/ close to /n-ħ/, etc.).

The results of this study provided for the first time an empirical evidence based on auditory judgments of six judges and from an acoustic point of view that the percept of nasality is increased when pharyngeal consonants /ħ ʕ/ are associated with nasal contexts. They lead to an increase of nasality over and above the actual nasal contexts (see the results for the 22% subset of the data in Figures 5 and 6 from Khattab et al., 2018).

[23] n = nasal consonant; o = oral consonant.

4.5.4 Conclusion

It is clear from the result of this section that guttural consonants have a special status in Arabic (and likely to be the case in other languages). My research within this area put emphasis on going beyond the main articulatory and/or acoustic correlate and to explore the role of secondary correlates. It is important to note that Arabic does not have a phonation contrast to signal differences between tense vs modal voice or breathy vs creaky voice as is the case in many African and South-Asian languages. However, the use of phonation to differentiate categories that are primarily distinguished on the basis of supra-laryngeal constrictions has not been used extensively in the literature, albeit in speaker and/or dialect identification (as an example, see Gendrot, Ferragne, & Pellegrini, 2019).

In this major research area, I showed how FPD can help in classifying and distinguishing specific contexts, and provided an empirical evidence for the legitimacy of gutturals as a natural class. The use of Voice Quality measures is important as it highlights how individuals systematically use their voice characteristics to guide differentiating between phonological categories in their languages. Of course, Voice Quality traits are individual and shape the identity of a speaker; this will be explored further in the next section.

4.6 Automatic methods and cognitive disorders - Speech as a biomarker?

The body of research presented in the previous sections highlighted the importance of FPD and its role in describing the gradient nature of phonological categories. The aim was to identify how phonetic implementation of phonological features can inform the rich nature of our mental representations. By doing so, I aimed at quantifying how FPD can be explored, in the majority of my research from an acoustic point of view, with more recent articulatory work to allow for a clear articulatory-to-acoustic mapping.

It is clear that my research within FPD puts emphasis on phonological categories and so far relied heavily on the use of semi-automatic and manual segmentations with semi-automatic analyses of acoustic features. In each case, the use of multiple correlates led to the development of specific profiles related to the specific research question asked in each publication, but all required semi-automatic segmentation and analyses. In the next sections, I'll present some research activities that tried to go beyond this and were aimed for a more applicative use of acoustic analyses, for supporting diagnosis of cognitive disorders and their impact on speech as produced by various clinical populations.

In this body of research, the main question we are trying to ask is the following: Can automatic methods allow for speech to be used as a potential bio-marker to support diagnosis of cognitive disorders? We know that speech carries a lot of detail on how we speak, with speed of speech; pronunciation; (filled-)pauses; errors and disfluencies; misarticulations, sociophonetic information, etc.. We know as well from previous research that we have a rich representation in our mental lexicon, where speaker-idiosyncrasies and fine-phonetic detail about speech sounds are stored. We are able to recognise and recall speakers with specific coarticulatory and/or voice patterns (Foulkes & Docherty, 2006; Hawkins, 2010; McAleer & Belin, 2018; Pierrehumbert, 2001; Zellou, 2017a). If these details are well stored in our brain, the assumption is that our own productions will signal these idiosyncrasies. Our aim then is to assess whether automated methods of speech can help to unravel this and to identify specific patterns and if these can be used as a potential bio-marker to support diagnosis in cognitive disorders. Of course the use of automated method is used here to allow for an analyst-independent tool in supporting diagnosis.

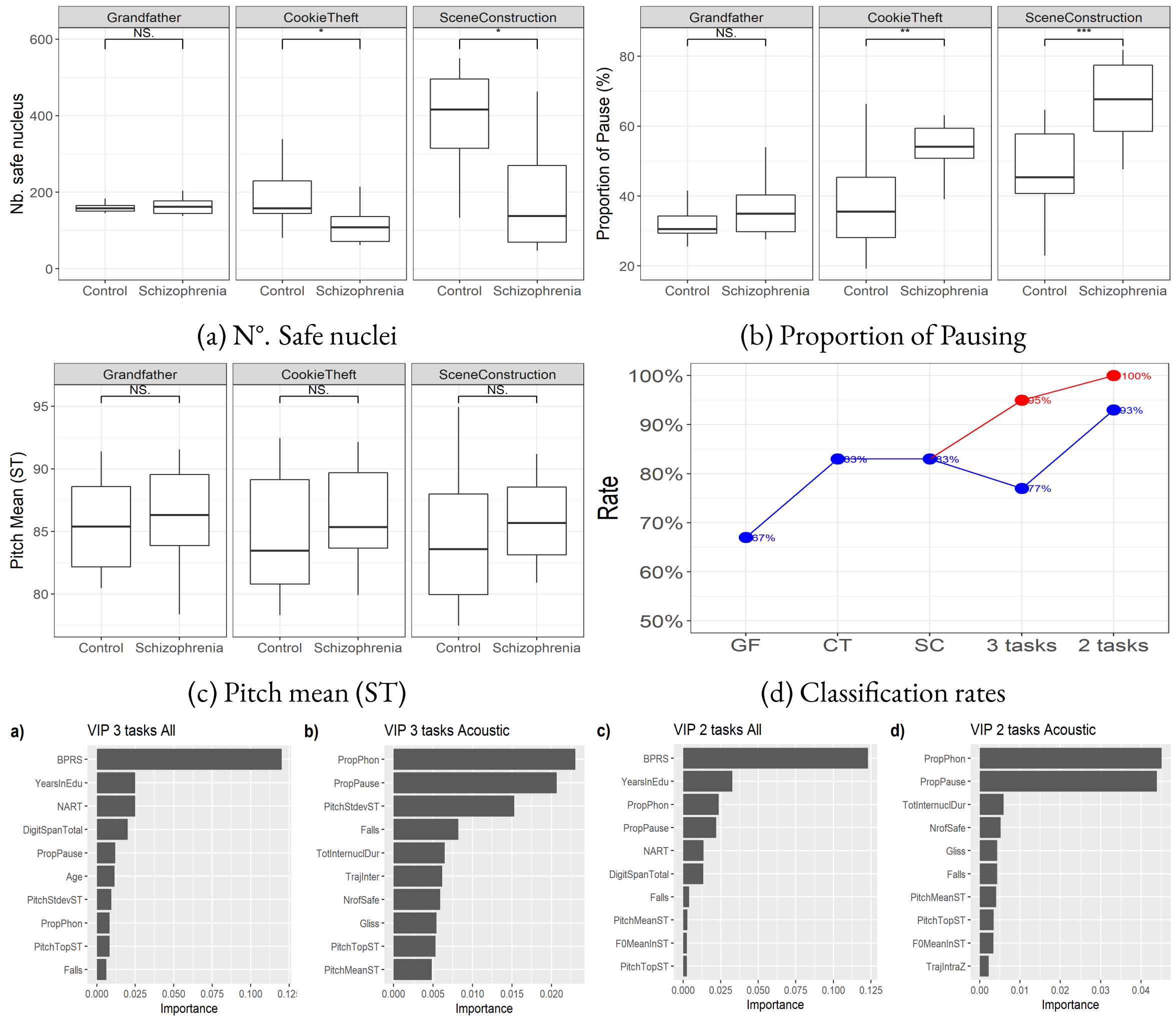
4.6.1 Schizophrenia

This research activity started off by asking the question of whether automated prosodic profiles allow to provide a clear indication of pathological voices and how different they are from neurotypical healthy controls (HC). Here we explored this question in the productions of participants identified with moderate Schizophrenia (SZ), in collaborations with Will Jones in Newcastle, UK and Wolfram Hinzen in Barcelona, Spain. This led to the following: W. Jones and Al-Tamimi (2020, BAAP conference presentation); W. Jones, Al-Tamimi, and Hinzen (2021, 4^{th} PaPE); J. Al-Tamimi (2022a, Invited speaker talk Symposium Labex EFL) and J. Al-Tamimi, Jones, and Hinzen (In preparation, Journal article to submit to JASA).

Previous literature on the SZ profile highlighted an impacted prosody, which was described as "Flat affect"; changes were observed in various domains, including increase in pausing, reduction in speech rate, pitch variation and changes and in the intensity of the voice (Cohen, Mitchell, & Elvevåg, 2014; Dickey et al., 2012; Hoekert, Kahn, Pijineborg, & Aleman, 2007; Martínez-Sánchez et al., 2015). However, and due to the lack of established biomarkers for monitoring or diagnosing schizophrenia and heterogeneity in schizophrenia types, variability is clearly marked in the identified patterns (Parola, Simonsen, Bliksted, & Fusaroli, 2019). This was especially the case due to variation in type of speech used.

Our study looked specifically at this by quantifying the prosodic profiles of our participants. The data comes from 12 HC and 12 SZ participants (7 males and 5 females in each group), who were age-matched, with a range between 22-68 (SZ → Mean = 45.2, SD = 9.8; HC → Mean = 44.6, SD = 15.1; SZ: β = 0.583, t(22) = 0.109, p = 0.914). We collected three speech tasks: a **Constrained** task as read speech of the grandfather passage, GF; **Partially constrained** as the "Cookie Theft" Picture description, CT and **Unconstrained** as a scene construction of a derelict-building, SC. We used Prosogram (Mertens, 2004), a plugin to Praat (Boersma & Weenink, 2013), to first automatically obtain an acoustic segmentation into acoustic syllables and nuclei and then a full prosodic profile for each participant and speech task. This Profile, included 27 out of the of 42 predictors (2 of these contained zeroes due to automatic computations), such as: temporal (speech rate, proportion of phonation/pausing, mean/SD nucleus duration) and f0 measures (f0 mean, median, SD, range), in addition to Global (number of safe nuclei, total duration of pausing or intervals) and Specific (falls, raises, trajectories, normalised and non-normalised metrics).

Figure 4.38 shows a snapshot of some patterns observed in the data. Clearly, SZ participants produced less nuclei than HC (Figure 4.38a) and with an increase in the proportion of pausing (Figure 4.38b) especially in the two less controlled tasks, i.e., CT and the SC, with no observed differences in the GF passage. Many of the other measures showed non-statistically significant differences across the two groups (see Figure 4.38c for an example of systematic non-significant increase in the Pitch mean in SZ) . We then decided to include all 27 predictors, in addition to demographics and psychological ratings in our Random Forests classification algorithm. The results show that when using acoustic metrics alone (Figure 4.38d, red curves), the best rates are achieved across CT, SC, and two less constrained tasks, with the best rate at 93% in the 2 tasks. The main predictors used by the algorithm were %phonation/pausing; pitch measures with mean,



(a) N°. Safe nuclei (b) Proportion of Pausing

(c) Pitch mean (ST) (d) Classification rates

(e) Variable Importance - 3 Tasks (f) Variable Importance - 2 Tasks

Figure 4.38: Results of of the Prosodic Profile in HC vs SZ across the speech tasks: (a) N°. Safe nuclei; (b) Proportion of Pausing; (c) Pitch mean (ST); (d) Classification rates with acoustics only in red and acoustics plus psychological measures in blue; (e) Variable Importance for the 3 Tasks (GF + CT + SC); (f) Variable Importance for the 2 Tasks (CT + SC); GF = Grandfather; CT = Cookie Theft; SC = Scene Construction (reproduced from W. Jones et al., 2021)

trajectories, falls/rises and duration of nuclei (see Figures 4.38e, b and 4.38f, d). When combining psychological measures with acoustic ones (Figure 4.38d, blue curves), no improvements are achieved in the individual tasks over acoustics alone, however, the rates increase to 95% in 3 tasks and 100% in 2 tasks. The main predictors used by the algorithm were the BPRS (Brief Psychiatric Rating Scale Overall & Gorham, 1962); other psychological measures and % phonation/pause.

This study highlighted that the speech task has a major impact on identifying clear patterns in SZ. When looking at the unconstrained tasks, more differences emerged between the groups. Combining psychological measures to acoustic ones increased classification accuracy to 95-100%, when combining the tasks. Finally, using acoustic measures alone on the two unconstrained tasks allowed the classification algorithm to achieve 93% accuracy in separating the two groups. These results are promising and more exploration into the role of automated prosodic profiles is carried out with another group of patients with cognitive disorders.

4.6.2 Alzheimer

In[24] the same vain of using automated methods to support diagnosis, we continued our collaboration by exploring whether automatic prosodic and voice quality profiles allow for separating various groups of patients. We look in the next sections at two case studies: the first quantifying prosody and Voice Quality in Alzheimer Disease (AD) and its various pre-clinical and early stages (Section 4.6.2.2), on comparing performance between automatically obtained speech-based and language-based feature sets (Section 4.6.2.3) in addition to current work I am engaged in with an M2 student on using openSMILE vs Prosogram in guiding diagnosis of AD (Section 4.6.2.5).

4.6.2.1 Introduction

Alzheimer Disease (AD) is one of the major diseases in the Neurocognitive Disorders Spectrum, following the classification of DSM-5 Association (2013), impacting on various levels, including learning & memory, language (both receptive and expressive) and social cognition. AD is one of the major causes of deaths in elderly, with no drugs to cure or prevent it, although there have been recent developments of promising drugs targeted specifically for early stages (van Dyck et al., 2022). Thus, early detection becomes an important factor to slow its progression. Automated language and speech analyses of outputs from patients with AD have recently emerged as potential bio-markers to distinguish them from matching healthy controls, with majority of studies being conducted on English (Haider, Luz, Fromm, & MacWhinney, 2021). The development of dementia caused by AD is gradual, and cognitive changes are detectable years, and sometimes decades, before dementia is diagnosed. Following Reisberg and Gauthier (2008), AD can be classified as "Mild", "Moderate" and "Severe", which can be diagnosed via fMRI scans and through various neuropsychological tests. They are characterised by the presence of "severe" memory loss. In addition, there is a pre-clinical stage, described as Mild Cognitive Impairment (MCI), which is described as a precursor to AD and Vascular Dementia; diagnosed via interviews and questionnaires with presence of "moderate" memory loss. Finally, Subjective Cognitive Decline (SCD) is described as a potential precursor to AD and is associated with ageing from cognitively "normal" participants. SCD can have co-morbidity with other conditions, e.g., MCI, dementia and depression. It is diagnosed via interviews and questionnaires with presence of "mild" memory loss. The List 8 describes the major areas impacted upon by AD.

8. 1) Language (Lindsay, Tröger, & König, 2021) ⇒ Decline in
   - Semantic ⇒ Specific-word finding; loss of comprehension; incorrect words; loss of verbal fluency
   - Pragmatic ⇒ Speaking at an inappropriate time, repetition and digression from topic
   - Syntactic ⇒ reduced complexity and underlying structure of language and sometimes grouped together with grammaticality

   2) Speech (Martínez-Nicolás, Llorente, Martínez-Sánchez, & Meilán, 2021; Meilán et al., 2014; Meilán, Martínez-Sánchez, Carro, Sánchez, & Pérez, 2012; Shah et al., 2021; Themistocleous, Eckerström, & Kokkinakis, 2020) ⇒ AD patients exhibit

[24] This research attempts to identify prosodic and voice quality profiles of the earliest signals of dementia. It is partially motivated after the loss of my dad to vascular dementia, which remained misdiagnosed for many years! Hence I aim to contribute to this field of research by using my expertise in automated methods, speech processing and machine learning as a way to guide diagnosis.

- Longer and more frequent hesitations,
- Lower speech and articulation rates,
- Longer pauses
- An overall increase in fo and reduction in mean periods
- Increased breathiness, signalled by a higher Shimmer, a lower Cepstral Peak Prominence and lower periodicity

4.6.2.2 ACE - Prosodic and Voice Quality profiles in the various stages of AD

The details presented in the introduction showed that both language and speech changes associated with AD are often identified in the various stages of the disease, and are caused by cognitive decline. In our research, we ask the question on whether these changes are clearly identified in the pre-clinical (MCI) or precursors (SCD) to AD. We know well that normal ageing processes will show similarities to the above reported patterns, with small differences. In fact, Ivanova, Martínez-Nicolás, and Meilán (2023, pp. 3) explained that ageing processes will cause the speech of healthy older adults to be slower, with increase in the duration of pausing, reduction in speech rate, with an increase in fo fluctuations, with decrease in spectral energy and frequently perceived with a rough voice quality. Clearly these changes in heathy older adults mean that systematically comparing their productions to those of patients identified with dementia becomes important. Our aim here then is to evaluate potential changes associated with the various stages of AD in comparison to those of matched healthy older adults.

The data for the present study comes from audio-recorded interviews from a cohort of participants at the Alzheimer Centre Educativo (ACE) foundation in Barcelona, Spain. There was a total of 119 participants (71 males; 48 females; similar distribution across groups), belonging to 5 groups: HC (n = 19; age = 67.4±7.8), SCD, (n = 30; age = 69.1±7.8), nMCI (n = 16; age = 76.3±6.4), aMCI (n = 23; age = 74.0±8.0), and mild AD (n = 31; age = 78.5±8.0). Using the Clinical Dementia Rating, the HC and SCD had a rating of 0 (=no dementia), the nMCI and aMCI a rating of 0.5 (=questionable) and for AD, they had a rating of 1 (26 participants; Mild) or 2 (5 participants; Moderate). Various neuropsychological measures were used to evaluate the participants, in addition to three types of speech tasks: **Constrained** ⇒ Questions and answers; present, past and future tense; **Partially constrained** ⇒ Picture description and **Unconstrained** ⇒ free speech as scene construction (for more details, see J. Al-Tamimi, 2023; J. Al-Tamimi, Lofgren, et al., In preparation; Lofgren et al., Under Review).

We used Prosogram (Mertens, 2004) to obtain an automatically segmented speech signal into acoustic syllables and nuclei. Subsequently, two profiles were obtained: a Prosodic and a Voice Quality profile (using a modified version of Prosogram, implementing most of the acoustic measures defined in the "Acoustic Breathiness Index" Barsties v. Latoszek, Maryn, Gerrits, & De Bodt, 2017). For the prosodic profile, we used 33 predictors (32 predictors out of the 42 generated by Prosogram + the rate of pausing defined as duration of pausing/spoken output). For the Voice Quality profile, we used 34 predictors spanning temporal (mean period duration), fo fluctuations (Jitter), amplitude fluctuations (Shimmer), HNR (several bands), spectral slope/tilt, CPP (Cepstral Peak Prominence), Hammarberg Index, Energy components (various ranges), spectral slope and tilt, band energy differences and Glottal-to-Noise-Excitation (at 3500 and 4500 Hz) (J. Al-Tamimi, Lofgren, et al., In preparation, ; see J. Al-Tamimi, 2022h for the full Praat script available at `https://jalalal-tamimi.github.io/Praat-VQ-Measurements/`).

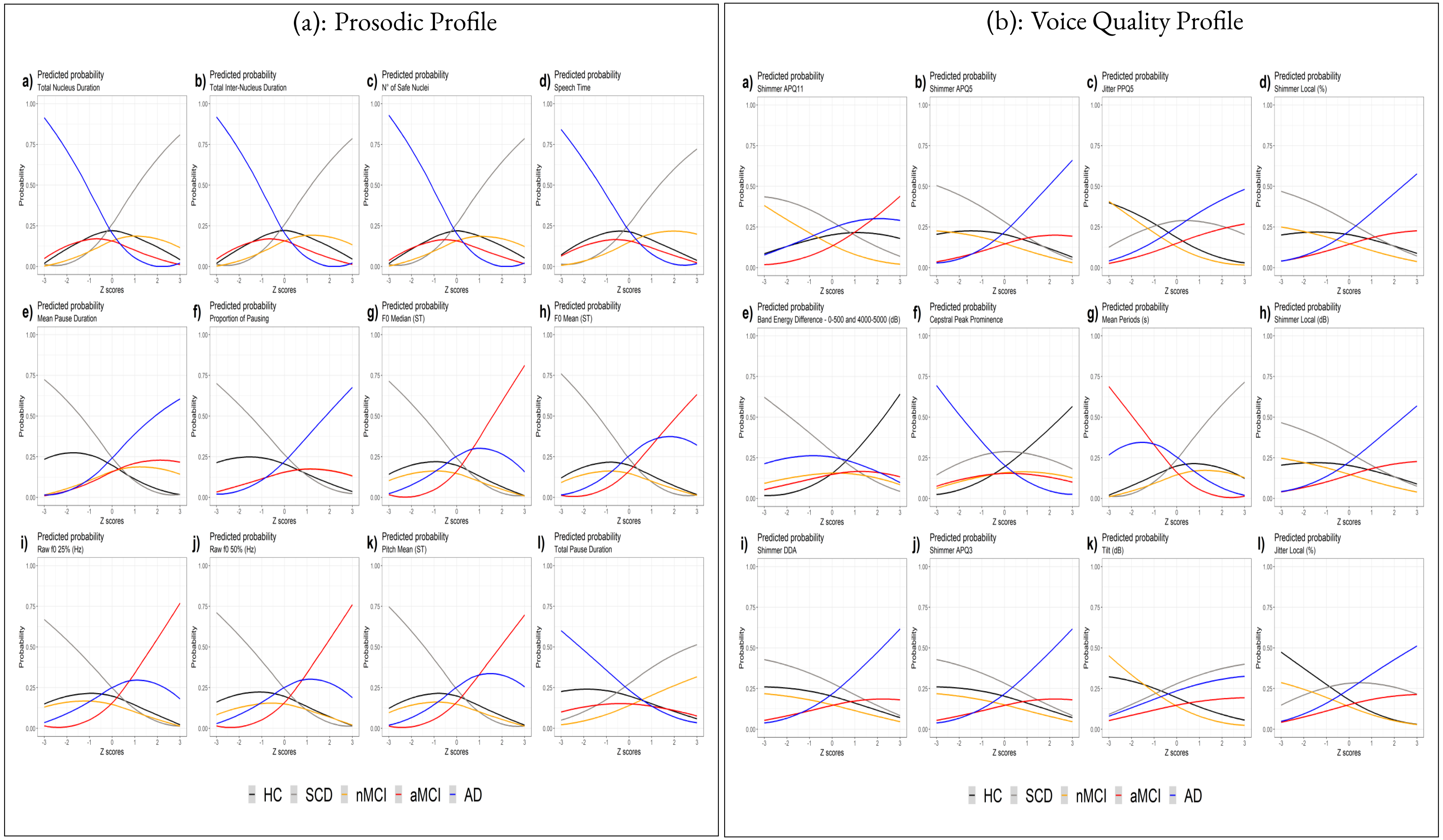


Figure 4.39: Prosodic and Voice Quality Profiles for the various stages of Alzheimer disease with HC, SCD, aMCI, nMCI and AD: (a) Prosody (Model fit: a to f have p<0.05; g to l p>0.05 and <0.1); (b) Voice Quality (Model fit: a to c have p<0.05; d to l p>0.05 and <0.1) (reproduced from J. Al-Tamimi, 2023; J. Al-Tamimi, Lofgren, et al., In preparation)

A predictive modelling approach was used here. First a Multinomial Logit Regression (MLR) was used, which included the four covariates (age, gender, language and years in education) to adjust for each coefficient. All acoustic predictors were z-scored to put all predictors on the same level. Then using model comparison, we retained predictors that had a $p<0.099$ to allow for both significant predictors and those displaying a tendency to work together in our subsequent analyses (Prosody = 12/33; VQ = 12/34). We obtained predictions from the MLR in probabilities (range -3 to +3 z-scores; covering 99.9% of distribution). The results presented in Figure 4.39 show the predicted probabilities for each of the Prosodic (Figure 4.39, a) or the Voice Quality (Figure 4.39, b) Profiles, for the retained predictors, ordered from the one contributing the most to the least, in each case. Given that the predictors are on the z-score, these reflect standardised β. The increase in a standardised β leads to an increase in the probability that a particular group is associated with an increased risk to show a positive association with the predictor in question. For instance, within the Prosodic Profile, the total nucleus duration (Figure 4.39, a, subplot a) shows that AD is associated with the lowest probability, followed by aMCI, HC and nMCI show similar patterns and finally, SCD shows the highest probability. This indicates that SCD are associated with an increase in the the total nucleus duration, with both AD and aMCI showing a decrease in the output; HC and nMCI are in-between.

Table 4.5: Summary of the patterns observed for Prosodic and VQ measures. ↑ = increase; ↓ = decrease; → = marginal change. Energy = energy in higher frequencies (reproduced from J. Al-Tamimi, 2023; J. Al-Tamimi, Lofgren, et al., In preparation)

	Type of measure	HC	SCD	nMCI	aMCI	AD
Prosodic	Spoken output	→	↑	↓→	↓	↓
	Pausing	→	↓	↑→	↑	↑
	Fo	→	↓	↓→	↑	↑
VQ	Shimmer	→	↓	↓→	↑	↑
	Jitter	→	↑	↓→	↑	↑
	Energy	→	↑	↑→	↓	↓

To make sense of the results and as a way to develop a specific profile for each stage, Table 4.5 presents the profile of our participants spanning prosodic and voice quality changes, with specific dimensions, with the HC taken as the reference for our comparison[25]. It is clear that our AD group shows an decrease in both the spoken output, and the energy in high frequencies, with an increase in pausing, fo, shimmer and jitter. These are indicative of slowness in articulation and increase in breathiness over and above that observed in HC. Interestingly aMCI shows an exact profile to AD; it is evidenced here that aMCIs are at a higher risk of conversion to AD due to the similarities in their profiles and in their diagnosis. The nMCI are different from both AD and aMCI, and are closer to HC. Finally, our results point to the SCD group to show a specific profile, with an increase in the spoken output, in the jitter and energy in high frequencies, with a decrease in pausing, fo and the shimmer. It is likely that our SCD group are using these specific patterns as a compensatory strategy to differentiate themselves from HC.

Based on our Prosodic and Voice Quality Profiles, our aim was to evaluate the discriminatory

[25]It should be reminded that our HC group is composed of healthy older adults. We did not compare the profiles of our HC group to younger adults to establish any specific differences related to ageing processes.

power of the patterns. Using machine learning (specifically, Extremely Randomised Random Forests), our results presented in Figure 4.40 show that none of the predictors (including prosody, voice quality, neuropsychological or any combinations) allowed for a clear discrimination between the five groups. This is likely due to the fact that the neuropsychological measures have a low sensitivity when it comes to separating the various stages of AD and pre-AD. When performing binary classifications (between groups and when combining them into healthy: HC and SCD vs clinical: nMCI, aMCI and AD), it was evident that clearer patterns emerged. Interestingly, the neuropsychological measures allowed for a perfect separation between Healthy vs Clinical, or between HC vs AD, SCD vs AD or SCD vs aMCI. When comparing SCD to nMCI using the neuropsychological measures, the rate was around 80% which was boosted to around 90% by combining any of the prosodic or voice quality measures to the neuropsychological measures. Voice Quality measures were clearly better at separating the HC vs SCD groups, at 82% in comparison to the 50% obtained with the neuropsychological measures.

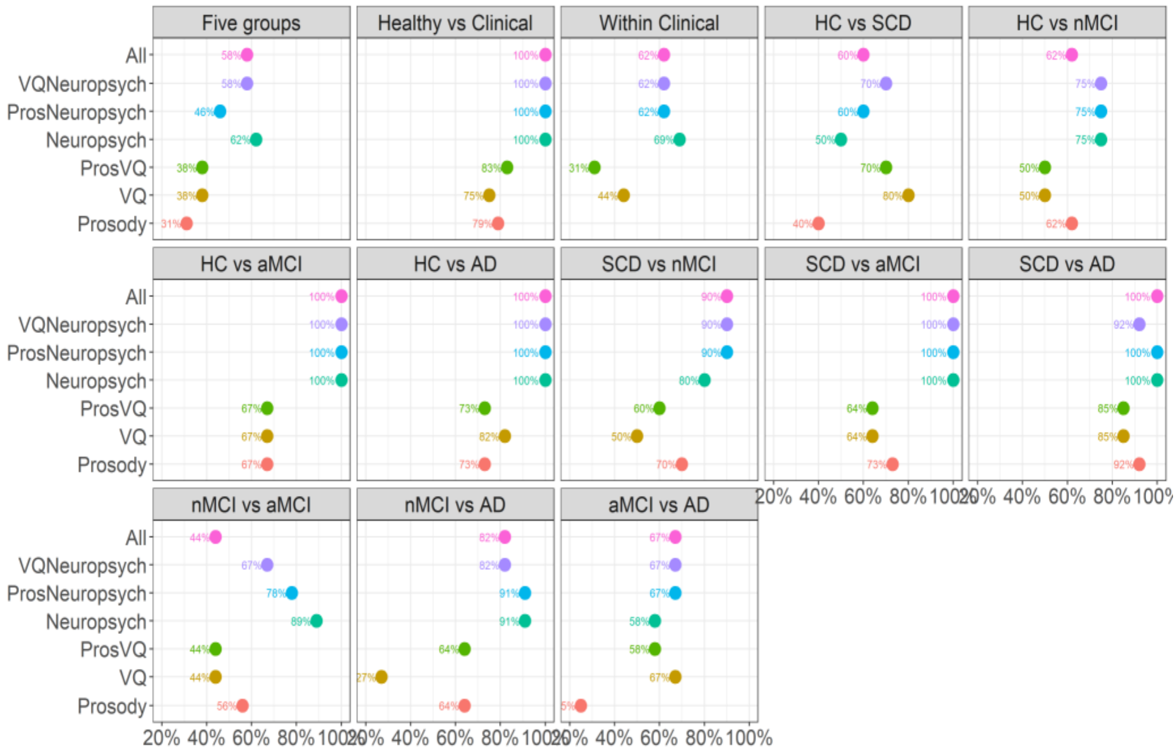


Figure 4.40: Classification rates for the Prosody, VQ, Prosody + VQ, Neuropsych, Prosody + Neuropsych, VQ + Neuropsych and the combination of all metrics (all) on the y-axis within the various types of comparisons (reproduced from J. Al-Tamimi, 2023; J. Al-Tamimi, Lofgren, et al., In preparation)

These results are interesting as they show that the small scale automated measures obtained from Prosogram (for both Prosodic and Voice Quality profiles), provide a useful account for the patterns observed in the data. Clearly, the various stages of AD are different from AD, with aMCI showing the closest patterns to those with AD; nMCI on the other hand shows differences in comparison to nMCI or AD. Finally, the SCD group who are healthy older participants subjectively complaining of Cognitive Decline, and who obtain neuropsychological scores similar to the neurotypical healthy older participants (our HC group) show clear voice quality patterns that can be used to allow differentiating them from the HC group. In fact, when using the MLR predictions on the neuropsychological measures (see supplementary material from J. Al-Tamimi, Lofgren, et al., In preparation), it is clear that the SCD group show opposite patterns to those seen in the AD group and their scores are often higher than those of the HC group; again likely to be a compensatory mechanism. In terms of Voice Quality measures, it is expected that some laryngeal changes are to be associated with normal ageing processes, however, Mayo-Yáñez and Cabo-Varela (2020, pp. 6, citing Woo, Casper, Colton, & Brewer, 1992) explain that "*...dysphonia in the elderly should be considered abnormal, and sometimes it represents the initial symptoms of a degenerative disease*". Our results here are indicative of the SCD group to be at a higher risk of developing Type 1 dysphonia/hoarse voice (following the classifications of S. R. Schwartz et al., 2009; Yanagihara, 1967) and are associated with partial breathy phonation (Barsties v. Latoszek et al., 2017). For more details, see J. Al-Tamimi (2023); J. Al-Tamimi, Lofgren, et al. (In preparation).

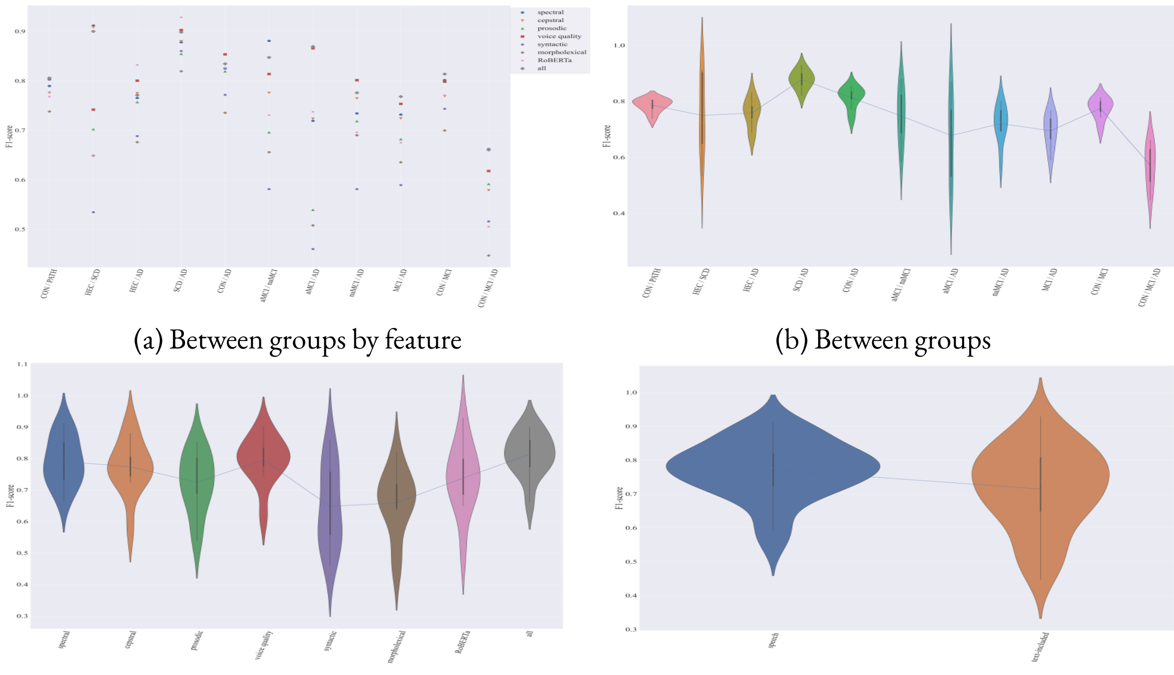
4.6.2.3 ACE - Language or Speech?

He et al. (2023) used a more data driven approach and aimed at identifying which automated approach yields the best discrimination between groups. Using the Scene Construction task on the same dataset and trying to separate the five groups, He et al. (2023) worked with the following Speech and language features (see List 9).

9. 1) Speech ⇒ **Automatic** acoustic features
    - openSMILE (Eyben, 2016; Eyben, Wöllmer, & Schuller, 2010) ⇒ ComParE16 feature set (Schuller et al., 2016) comprising 6373 predictors, divided into. spectral, Cepstral (ASR), voice quality, prosody
    - Prosogram ⇒ 42 prosodic predictors (temporal and *f*0)

   2) Language ⇒ From pre-transcribed material (3 minutes)
    - Syntactic ⇒ **Manual**: 21 predictors, including number of utterances, utterance length, type of articles, anomalies, etc. (following Chapin, Clarke, Garrard, & Hinzen, 2022)
    - Morpho-lexical ⇒ **Automatic**: 154 measures, including ratios of different word classes and morphological variants of words
    - Semantic properties ⇒ **Automatic** (on first 510 tokens): pretrained contextual word embeddings, using the robust optimized version of BERT; RoBERTa

ComParE16 feature set reduced from 6373 predictors to 1500, then used machine learning, in the form of Extremely Randomised Random Forests, 10-folds cross validation, with tuning of hyperparameters. Figure 4.41 shows the results of the classifications across various domains. Comparing between the groups and feature sets (Figure 4.41a), it is evident that there were clear variations in the performance of each group of feature sets. For instance, the aMCI vs AD comparison shows the best performance using Voice Quality and All (VQ ⇒ Accuracy = 89.8%, F1 score = 0.866; All ⇒ Accuracy = 90.5%, F1 score = 0.869), while the other domains showed variable performance with the F1 score ranging between 0.4 and 0.7. This figure shows that again the task separating all groups is a difficult task (see Figure 4.41a, CON / MCI / AD), as this comparison receives the lowest F1 scores. Figure 4.41b shows a violin plot with the averaged F1 scores. The results show that the best performance with a robust effect is achieved between SCD and AD: ≈ 89% (F1 score = 0.878) and between Control vs AD = 84% (F1 score = 0.812). Moving on to the effect of the feature set used, comparing performance across speech and language domains, Figure 4.41c highlights the advantage of speech over language measures. The best



(a) Between groups by feature

(b) Between groups

(c) Between feature sets

(d) Between Domains

Figure 4.41: Results of the classifications, with the F1-score across the various modalities: (a) Between groups by feature; (b) Between groups; (c) Between feature sets; (d) Between Domains (Speech vs text-based); CON = Control (HEC and SCD); PATH = Pathological (aMCI, nMCI and AD), naMCI = nMCI, (reproduced from He et al., 2023)

performance is achieved when all features are combined: F1 score = 0.813, VQ = 0.796; spectral = 0.791; Cepstral = 0.773; embeddings from RoBERTa = 0.738; prosodic = 0.724; morpho-lexical = 0.659 and manual syntactic features = 0.648. And when combining all speech measures and all text-based measures together (Figure 4.41d), the results show on average, that Speech measures have an F1 score = 0.78, whereas Text-Based measures achieve an F1 score = 0.72.

These results point to an advantage for automatically obtained speech-based measures, especially Voice Quality over automatically quantified language-based measures based on manually pre-transcribed data. The results of this study report similar patterns to the previous study, whereby the SCD group displays an overall increase in various metrics, including Voice Quality and Cepstral metrics over and above HC or AD, as a possible compensatory strategy.

4.6.2.4 Correlation between prosodic measures and Aβ+-amyloid depositions?

An extension to this work looked at the atypical cortical abnormalities identified in adults with a positive Aβ-amyloid depositions, which is known to be an early neuropathological marker of AD. The work presented in He et al. (Under Review) investigated large-scale cortical gradients of functional connectivity with resting state fMRI, which capture the hierarchical integration of cortical functions, together with automatically extracted acoustic-prosodic features from spontaneous speech, in cognitively unimpaired older adults (total N = 188) with and without Aβ positivity. The results showed a significant alteration of the cortical hierarchies in the Aβ+ group, with less variation in comparison to the group-level template and with an increase in quantitative prosodic measures (e.g., N°. Of Nuclei, Speech time, Nucleus duration and Inter-Nucleus Duration). Figure 4.42 show the correlation plot for the five prosodic measures that retained for subsequent analyses after FDR correction. The results show a statistically significant correlation between four (out of the five automatically computed prosodic measures) and the third gradient (G3), with an increase in speech quantity. This suggests that unimpaired older adults with a positive Aβ-amyloid depositions show neuropsychological alterations in working memory functions (for more details, see He et al., Under Review)

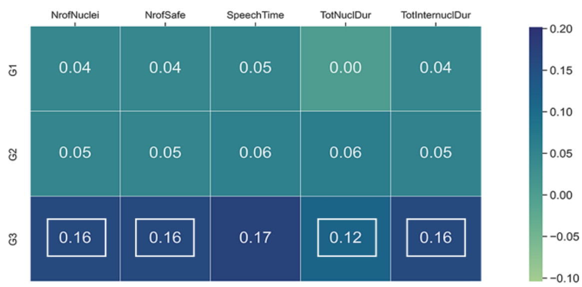


Figure 4.42: Coefficients of the partial Pearson correlation between the five Prosodic measures and the three gradient fits; white boxes indicate significant correlation (reproduced from He et al., Under Review, Figure 4)

The results of this study confirm that long before the clinical stage and objective cognitive impairment, increased risk of cognitive decline as indexed by Aβ-amyloid accumulation is marked by neurofunctional changes in the cortical hierarchy, which are related to automatically extractable speech patterns. We continue working on this dataset to evaluate additional automatically extracted but explainable measures, possibly using Voice Quality correlates.

4.6.2.5 The ADReSS Challenge

The results reported in J. Al-Tamimi (2023); J. Al-Tamimi, Lofgren, et al. (In preparation) highlighted an important role for the Prosodic and Voice Quality profiles in differentiating the various stages of AD; the precursor to AD (SCD), the preclinical (nMCI and aMCI) and mild to moderate AD. We used Prosogram to compute the Prosodic and the Voice Quality profiles (using a modified version of Prosogram). In our subsequent work, and in collaboration with Eric Jordan, during his M2 internship working with me, we wanted to evaluate the robustness of the patterns when comparing the performance of Prosogram to openSMILE (Eyben et al., 2010).

In this task, we used the ADReSS Challenge dataset (Luz, Haider, de la Fuente, Fromm, & MacWhinney, 2020), which aimed at developing and testing automated approaches in signal processing and Large Language Models (LLMs) to support diagnosis of AD. The dataset comes from the DementiaBank (Becker, Boiler, Lopez, Saxton, & McGonigle, 1994)[26] and is composed of a total of 156 participants (70 males and 86 females; aged between 50-80 years old). 50% of participants are neurotypical healthy elderly and 50% are diagnosed with AD, with no specific details as to the severity of the disease (the MMSE scores point to various stages ranging from mild to severe Becker et al., 1994).

Jordan and Al-Tamimi (2023) used various algorithms: Prosogram (with 42 Prosodic Profile measures); GeMAPS (containing up to 62 measures: spectral, temporal, energy and voice quality,

[26]We acknowledge these two grants supporting the work within the DementiaBank: NIA AG03705 and AG05133; in addition to funding from the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083), contributing to the IdEx Université Paris Cité - ANR-18-IDEX-0001; the High Power Computing (HPC) facilities: CNRS/TGIR HUMA-NUM, IN2P3 and GENCI-IDRIS (Grant 2022-AD010613733)

see Eyben et al., 2016) and ComParE16 (up to 6373 measures: spectral, auditory, PLP, MFCCs, cepstral, temporal, energy, voice quality Schuller et al., 2016). Both GeMAPS and ComParE16 contain LLDs (Low Level Descriptors) and Functionals (derived) measures, which were used subsequently. We worked on various combinations of algorithms, including computations on the normalised chunks or raw data, averaged measures across frames, using default segmentations from openSMILE or from Prosogram. The Figure 4.43 shows a summary of the performance of the various combinations, with the accuracy and the F1 scores obtained from the Extra Trees Random Forests, correcting for both collinearity and with Recursive Feature Elimination. Overall, it is clear that the Prosogram measures outperform the GeMAPS or ComParE16 using original segmentations (normalised) or automatic segmentation from Prosogram (automatic). Luz et al. (2020) reported a benchmark rate of 62.5% using Random Forests; our rates using the GeMAPS or ComParE16 hover around this; the GeMAPS or ComParE16 using openSMILE's segmentation on the normalised audio chunks obtain rates close to 65%; with Prosogram's Prosodic profile we reach 72.9%, which is a relatively encouraging rate on this very challenging dataset, although researchers contributing to the volume Haider et al. (2021) using speech and/or LLMs with more sophisticated algorithms achieved rates close to 83%.

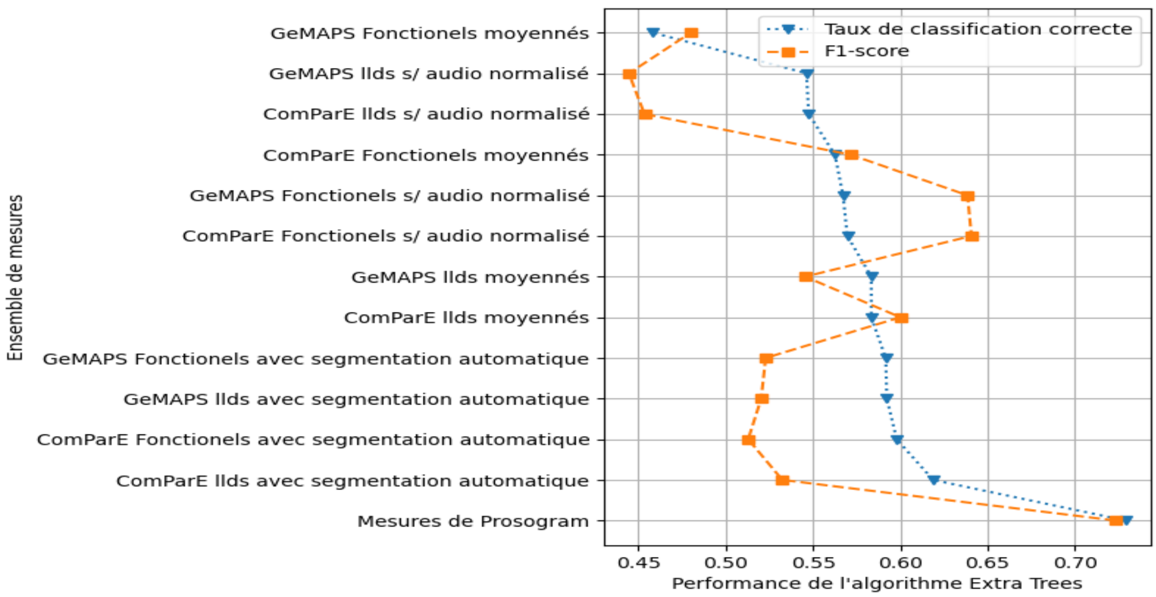


Figure 4.43: Accuracy and F1 measure of the various signal processing algorithms using the Extra Trees Random Forests (reproduced from Jordan & Al-Tamimi, 2023)

It is clear that ageing processes lead to changes in the speech produced by neurotypical healthy elderly, which is likely to show similarities with some patterns caused by AD. However, the speech of participants identified with various levels of AD contain specific markers, and more research is needed to identify speech and/or language features as potential bio-makers to guide diagnosis of this challenging disease.

4.6.3 Conclusion

The research conducted within this section highlighted the fact that FPD quantified via automatic signal processing algorithms is a viable solution at least in supporting diagnosis of cognitive disorders. The results presented in J. Al-Tamimi (2022a, 2023); J. Al-Tamimi, Lofgren, et al. (In preparation); Jordan and Al-Tamimi (2023) showed a clear advantage to the use of the Prosodic Profile, especially temporal domain and quantity of spoken outputs as the main features quantifying differences between HC and SZ or AD on the one hand, and HC vs AD or HC vs SCD on the other hand. Specific to AD, temporal measures are the most explored features and the most prevalent in the speech of participants identified with AD as they can be "*[...]can be significantly altered as a result of deficits in executive functions, working and episodic memory, mental manipulation, timing control, attention, language decision-making, semantic retrieval, discourse planning and topic adjustment (De Looze et al., 2021; Sluis et al., 2020)*" Ivanova et al. (2023, pp. 13). Pausing is another prominent feature in the speech of AD; its length becomes more important but not necessarily more frequent, which is clearly more marked in more cognitively challenging tasks, which require more demands on planning.

In addition, the results of He et al. (2023) showed a clear advantage for automatically obtained speech-based features (especially Voice Quality and Cepstral domains) over automatically quantified language-based measures. The language-based measures used in this study are based on manually transcribed data and we wonder whether the lower rates are obtained due to the data being manually transcribed? Hence, is it likely that using automatic language-based metrics from automatic transcriptions would work better? This is an open question which will be explored in future research.

The results reported in He et al. (Under Review) are promising as they show a potential for a compensatory strategy by cognitively unimpaired older adults identified with a positive Aβ-amyloid depositions in producing more spoken output, which also correlates with neuropsychological measures and with specific brain functions. Clearly, signs of cognitive decline associated with AD can emerge decades before actual clinical diagnosis and our results are a first step towards finding speech biomarkers to guide diagnosis of the disease.

In conclusion, my contribution within this field aimed at identifying linguistically interpretable measures to guide interpretation of the findings. The various feature sets within openSMILE provide automated measures that are sometimes widely used in ASR, in emotion detection, and as we saw here, in supporting diagnosis. Prosogram provides linguistically explainable metrics with cost at predictions; openSMILE provide strong predictions, with least explainability. It is clear that Combining various sources of information, such as acoustic + language + (neuro-)psychological measures is a promising approach guiding diagnosis and monitoring progression of the disease.

4.7 Arabic WebMAUS: Romanisation system and Forced-alignment

Since starting my work on describing Arabic phonetics and phonology, it was always a challenge to obtain accurately transcribed and time-aligned segmented speech. In most of the research described above, manual transcription and segmentation of speech was obtained. However, this was extremely time consuming as it required minute and fine detailed segmentation. The work described within Section 4.3 (minus that for child speech; Section 4.3.4) used a semi-automatic segmentation system that identifies consonantal and vocalic portions (transcribed as C and V, respectively) based on $f0$ and intensity changes (see Farinas, Rouas, Pellegrino, & André-Obrecht, 2005, for more details). However, due to the need for a more advanced system that allows to obtain a more robust segmentation of the speech stream in Arabic, I worked on the development of a forced-alignment system for Arabic that started in 2015.

First, using `PraatAlign` (Lubbers & Torreira, 2013/2018), a plugin to Praat that used `MAUS-universal language` (Schiel, 2015, 2018), I worked on modifying specific scripts to make the plugin work on Arabic. This was essential in adapting the system to use a first version of a romanisation system that was easy enough that allowed any researcher to transcribe Arabic scripts using Roman symbols. Due to the fact that the universal system relied heavily on acoustic models available from the 15 languages that were part of the MAUS forced-aligner, the success rate was around 70-80% reaching 100% accuracy for specific phonemes (e.g., nasals and lateral). For pharyngeal, and pharyngealised consonants in Arabic, the rates were close to a mere 50%. Since then, various researchers seeked support for their data to be forced-aligned and shared their primary data with me in the aim to build a robust forced-alignment system in the future.

This led to a formal collaboration with the MAUS team (see J. Al-Tamimi, Schiel, et al., 2022). First, an easier and more language-specific romanisation system for Arabic was developed, which was followed by creating the first macro, dialect-independent Arabic-WebMAUS service, which was fully integrated within the MAUS services (`https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface`: first release on 17-12-2021; version 2 updated to include additional resources on 15-04-2022; Schiel & Al-Tamimi, 2021).

For a full integration within the MAUS services, it was essential to devise a romanisation system that will work within the WebMAUS services. The romanisation system, named as `Al-Tamimi Romanisation System; ATR` is shown in Tables 4.6 and 4.7. ATR contains 98 phonemes that had a 1-to-1 match between a produced sound and a symbol to transcribe it, using ASCII characters. This covered all possible sounds available in the various Arabic varieties (and the standard form) used for the development of the first version of the Arabic-WebMAUS service (Lebanese, Levantine, Bahraini, Saudi, Jordanian Arabic). Based on previous literature, we decided to have different acoustic models for each of the singleton and geminate consonants (Table 4.6) and for short and long vowels (Table 4.7) as my own previous research has shown a clear quantitative and qualitative difference between these categories (see Section 4.3 for the former; Section 4.2 for the latter).

The second part of the development concerned creating 105 acoustic models, for the 98 phonemes and for an additional 7 silence and noise models. Out of the 98 phonemes, we trained 69 models for which there were at least 75 occurrences in our datasets; 29 models were cloned either from other languages or from singleton to geminate (or vice-versa, see J. Al-Tamimi, Schiel, et al., 2022, Section 3, pp. 7272, for more details). The training set in version 2 comprises 6610 recordings, from 94 speakers, with a total duration of 16h10min and 509804 labelled segments.

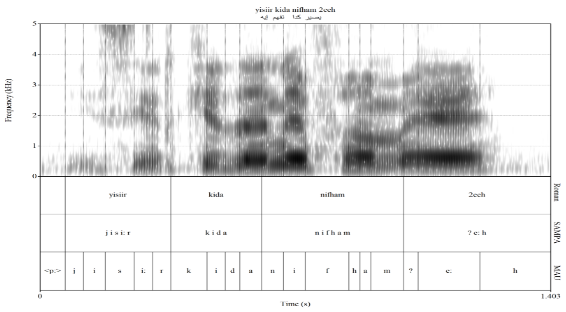
After training the acoustic models, two systems were developed and integrated within the WebMAUS services: A WebMAUS and a WebMINNI. The former allowed for forced-alignment of pre-transcribed speech using either the ATR convention or X-SAMPA); the latter allowed for an automatic Speech-to-Text transcription (see Figures 4.44). Figure 4.44a shows the Arabic WebMAUS performance on the utterance "yisiir kida nifham 2eeh" ("How do we understand this") as produced in Jiddah Arabic (a sample not used in the original training set). The performance of the forced-alignment (tiers 2 and 3) are impressive and are around 90% at 5ms

Table 4.6: List of singleton vs geminate consonants used in the ATR system, with each phoneme transcribed in IPA Unicode, ATR and X-SAMPA (reproduced from J. Al-Tamimi, Schiel, et al., 2022, pp. 7271-7272)

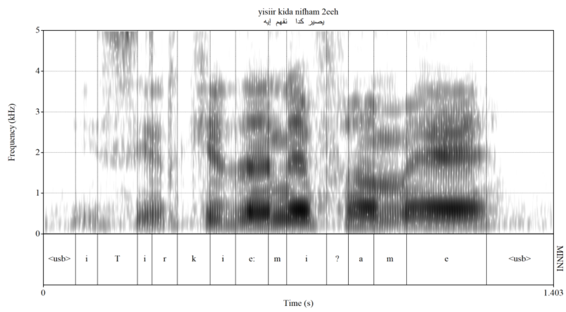
Singleton			Geminate		
IPA	ATR System	X-SAMPA	IPA	ATR System	X-SAMPA
ʔ	2	?	ʔʔ	22	??
b	b	b	bb	bb	bb
t	t	t	tt	tt	tt
θ	t\	T	θθ	t\t\	TT
ʒ	j	Z	ʒʒ	jj	ZZ
ħ	H	X\	ħħ	HH	X\
x	x	x	xx	xx	xx
χ	X	X	χχ	XX	XX
d	d	d	dd	dd	dd
ð	d\	D	ðð	d\d\	DD
r	r	r	rr	rr	rr
z	z	z	zz	zz	zz
s	s	s	ss	ss	ss
ʃ	s\	S	ʃʃ	s\s\	SS
sˤ	S	s_?\	sˤsˤ	SS	s_?\s_?\
dˤ	D	d_?\	dˤdˤ	DD	d_?\d_?\
tˤ	T	t_?\	tˤtˤ	TT	t_?\t_?\
ðˤ	D\	D_?\	ðˤðˤ	D\D\	D_?\D_?\
zˤ	z\	z_?\	zˤzˤ	ZZ	z_?\z_?\
lˤ	L	l_?\	lˤlˤ	LL	l_?\l_?\
ʕ	3	?\	ʕʕ	33	?\
ɣ	G	G	ɣɣ	GG	GG
ʁ	G\	G\	ʁʁ	G\G\	G\G\
f	f	f	ff	ff	ff
q	q	q	qq	qq	qq
ɡ	g	g	ɡɡ	gg	gg
k	k	k	kk	kk	kk
l	l	l	ll	ll	ll
m	m	m	mm	mm	mm
n	n	n	nn	nn	nn
h	h	h	hh	hh	hh
w	w	w	ww	ww	ww
j	y	j	jj	yy	jj
ʧ	ch	tS	ʧʧ	chch	tStS
ʤ	dj	dZ	ʤʤ	djdj	dZdZ
v	v	v	vv	vv	vv
p	p	p	pp	pp	pp

Table 4.7: List of short vs long vowels used in the ATR system, with each phoneme transcribed in IPA Unicode, ATR and X-SAMPA (reproduced from J. Al-Tamimi, Schiel, et al., 2022, pp. 7271-7272)

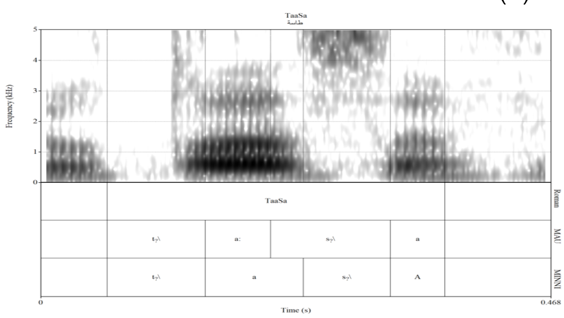
Short			Long		
IPA	ATR System	X-SAMPA	IPA	ATR System	X-SAMPA
i	i	i	iː	ii	i:
ɪ	I	I	ɪː	II	I:
e	e	e	eː	ee	e:
ɛ	E	E	ɛː	EE	E:
æ	ae	{	æː	aeae	{:
a	a	a	aː	aa	a:
ɑ	A	A	ɑː	AA	A:
ɔ	O	O	ɔː	OO	O:
o	o	o	oː	oo	o:
u	u	u	uː	uu	u:
ʊ	U	U	ʊː	UU	U:
ə	@	@	əː	@@	@:



(a) WebMAUS



(b) WebMINNI



(c) WebMAUS + WebMINNI

Figure 4.44: Results of the Arabic WebMAUS on a Saudi Arabic example: (a) Alignment of the utterance "How do we understand this" in Saudi Arabic; (b) WebMINNI transcriptions of the utterance "How do we understand this" and (c) Alignment (tier 2) and WebMINNI (tier 3) of the word "cup" (reproduced from the poster of J. Al-Tamimi, Schiel, et al., 2022)

boundary positioning that is seen in other platforms. Figure 4.44b shows the Arabic WebMINNI performance on the same utterance; the input was the audio file without any prior transcriptions. Of course, the system is not perfect, but interestingly, vowels seem to be identified more accurately than consonants, and in many instances, guttural sounds are better identified. Figure 4.44c presents a comparison of the performance of the WebMAUS and the WebMINNI services on the same word "TaaSa" ("cup"). All segments were correctly transcribed using the WebMINNI with almost exact boundaries, in comparison with the WebMAUS service. The work on adding additional resources continues; see Section 7.6.

4.8 Other collaborative work

This section summarises the various collaborations I have engaged in showcasing my contributions to each, which will be presented in chronological order.

4.8.1 Cross-linguistic child directed speech

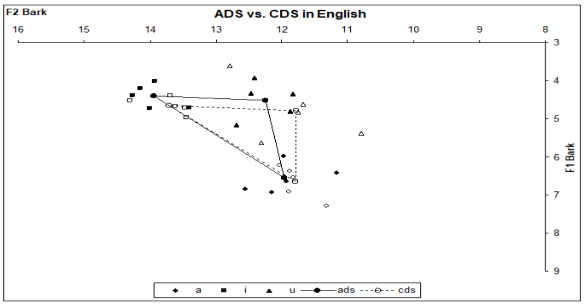
A first collaborative work emerged in 2007 during which I used my expertise in acoustic analyses of vowels and quantifying vowel space sizes, by exploring vowel dispersion differences in Adult-Directed-Speech (ADS) vs Child-Directed-Speech (CDS), which was published in Dodane and Al-Tamimi (2007, 16^{th}ICPhS). Our aim was to evaluate the claims advanced in P. K. Kuhl et al. (1997), who demonstrated that vowels embedded in specific CVC words in CDS are produced as more extreme than in ADS, with a clear hyperarticulated forms for the point vowels /i/, /a/ and /u/ in English, Swedish and Russian. CDS showed a clear expansion of the vowel triangles on the F1 and F2 dimensions.

In Dodane and Al-Tamimi (2007), we aimed to examine if these patterns are to be observed in languages with differing rhythmic structures (English, French and Japanese) producing more spontaneous and less restricted environment. Our results (see Figure 4.45) showed that there was no clear expansion of the vowel spaces in CDS when compared to ADS; in CDS, a more open production of the vowels /i/ and /u/ in English (Figure 4.45a) and in the three vowels in French and Japanese (Figures 4.45b and 4.45c), with increased F1; changes on F2 were mostly caused by coarticulatory changes associated with surrounding consonants. Our results showed that due to the raised fo associated with the productions of the mums in CDS, that the distance Z1-Zo was much lower than in ADS, which was more marked across the three English vowels and only in French /i/ (Figures 4.45d, 4.45e and 4.45f).

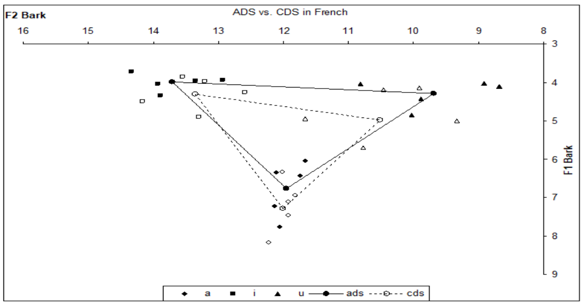
Our conclusion is that the main expansion effect observed in CDS is lost when using spontaneous interactions, however, CDS speech is associated with more open productions and not more peripheral positions, except in French. Hence, we partially corroborated the original findings.

4.8.2 Acoustics of the contested fifth liquid in Malayalam

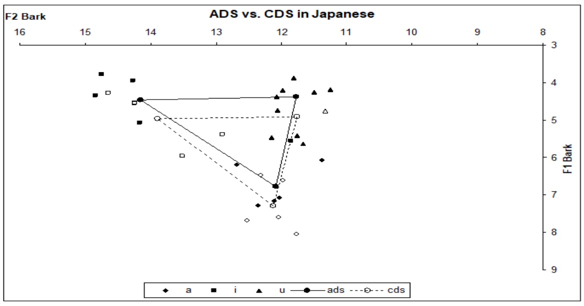
Another collaborative work emerged with a previous PhD student with whom I was a PhD panel member. The work published in Punnoose, Khattab, and Al-Tamimi (2013) looked at the acoustic correlates of the fifth liquid in Malayalam to try and separate it from the laterals and rhotics. Malayalam has five non-geminate liquids: Two laterals /l ɭ/, two rhotics /ɾ r/ and a fifth liquid /ɻ/. This fifth liquid is challenging as its descriptions in Dravidian languages varies from "voiced retroflex palatal fricativised lateral", to "voiced sublamino-palatal approximant", "simply a continuant", "simply a third rhotic produced as a central retroflex approximant" (Punnoose et



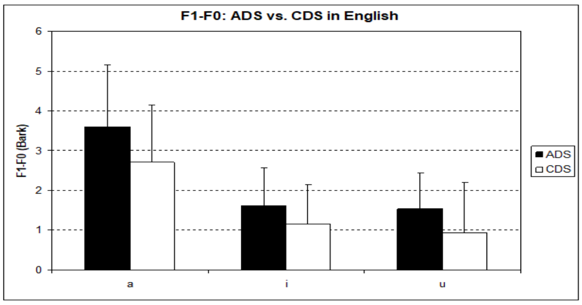
(a) ADS vs CDS in English



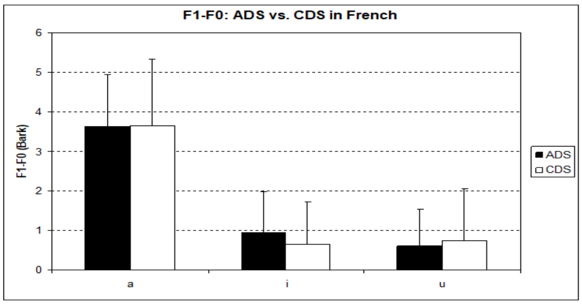
(b) ADS vs CDS in French



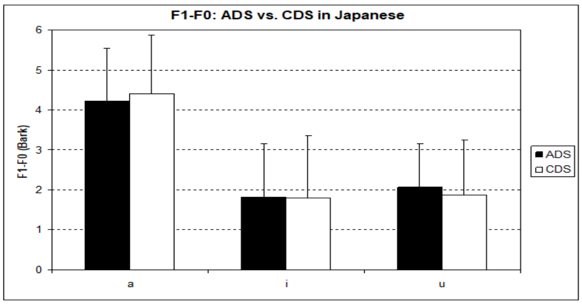
(c) ADS vs CDS in Japanese



(d) Z1-Z0 in English



(e) Z1-Z0 in French



(f) Z1-Z0 in Japanese

Figure 4.45: Results of ADS vs CDS in English (a and d), French (b and e) and Japanese (c and f): (a) Vowel spaces in English; (b) Vowel spaces in French; (c) Vowel spaces in Japanese; (d) Z1-Z0 in English; (e) Z1-Z0 in French; (f) Z1-Z0 in Japanese (reproduced from the article and poster of Dodane & Al-Tamimi, 2007, pp. 1575 and 1576)

al., 2013, pp. 274-275). Scobbie, Punnoose, and Khattab (2013) conducted a preliminary articulatory study on one Malayalam speaker, who found that /ɻ/ is produced as a retroflex approximant, with /ɻ/ displaying an advanced root and a raised tongue body (similar to /l ɾ/) as opposed to a retracted tongue root and lowered tongue body for /ɭ r//. My contribution to this study specifically looked at the acoustic patterns for F1, F2, F3 and F4 and in the statistical design.

Our results presented in Figure 4.46 show the pattern observed for the fifth liquid and how it differs from the other four categories. It is interesting to note that /ɻ/ showed similarities and differences with all four other categories. In fact, /ɻ/ showed a distinct profile: a high F2 like that of the alveolar lateral and tap /l ɾ/, suggesting clear resonance, with a relatively low F3 and F4 due to the retracted post-alveolar articulation, similar to that of /ɭ r/. An approximation between F2, F3 and F4 is observed /ɻ/, which can be indicative of a constriction in the pharyngeal/epilaryngeal area, similar to that we previously observed for Arabic gutturals with a potential for a double-bunched

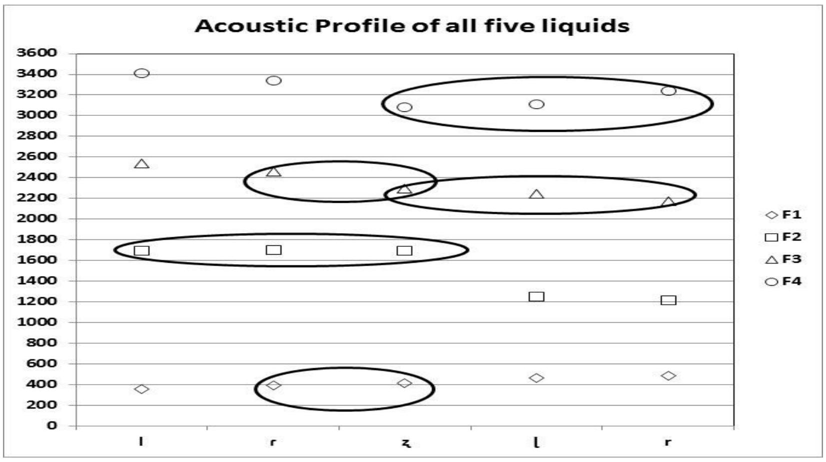


Figure 4.46: Average F1, F2, F3 and F4 mid-point values of the five liquids in Malayalam. Ellipses group formant results that show no significant difference between /ɻ/ and another liquid (reproduced from Punnoose et al., 2013, pp. 284)

type of articulation (following Esling et al., 2019; Moisik et al., 2019). Clearly, this profile led to specific patterns observed for the fifth liquid that made it sharing specific patterns from all other four categories. Coarticulatory patterns on the surrounding vowels showed a mixture of behaviour (for more details, see Punnoose et al., 2013).

4.8.3 "You came TO DIE?!"

I have contributed to the research project "You came TO DIE?! Perceptual adaptation to regional accents as a new lens on the puzzle of spoken word recognition"; grouping various researchers in Australia, New Zealand, and the United Kingdom and led by Catherine Best, Jen Hay, Bronwen Evans, Jason Shaw, Paul Foulkes and Gerry Docherty. The main aim of the project was to evaluate how short-term exposure to an unfamiliar accent impacts on the categorisation of that accent's vowels and consonants. The project looked at British English (represented by the varieties of Newcastle, York and London), Australian English (in Sydney) and New Zealand English (in Christchurch). I led on the data collection from 24 speakers (6 males and 6 females; two sites at Newcastle and York), which were then used for the perception experiments. Our results were published in three conferences proceedings that I co-authored. Shaw et al. (2014, 15th AISST) showed that consonants' categorisation in an unfamiliar accent is impacted even after a short-term exposure. Best, Shaw, Mulak, et al. (2015, 18th ICPhS) showed that a brief exposure of Australian listeners to English vowels as produced in London and Yorkshire from a read story leads them to assimilate the vowels to their own categories contributing to the perception of vowel variation. Finally, Best, Shaw, Docherty, et al. (2015, Interspeech) showed that when Australian listeners were exposed to the full vowel system of Newcastle English based on nonce-words, this lead to more similar vowels to be categorised as close to those in Australian vowels, whereas an increased perceptual differentiation between different vowels in the two accents led to heightened sensitivity to the fine-grained discrepancies. The results of the studies were situated within the episodic and abstractionist approaches to speech perception.

4.8.4 Research on L2 Phonology - Identification learners' varieties

4.8.4.1 Introduction

I got interested and engaged in this research activity after working with eight PhD students (six finished, one current, one starting in October 2023). The bulk of research I have engaged in looked specifically at how the speech of learners of an L2 (mainly English as an L2, with L2-English and L3-French) differ from the speech of first language users (English as an L1, and/or French as an L1) in terms of native-likeness and target-likeness and how in turn, they differ from their L1 (being Arabic, Thai, Mandarin, etc..). We were also interested in the role of input (second vs foreign language, type of instructions, being traditional or automated, etc.) on the patterns produced by the learners. A major component in these research activities is how structural differences between the target language and the mother tongue impact on the learning processes. We particularly look at the stability and variation in learning vocalic, consonantal and prosodic contrasts, results of which are situated within usage-based and exemplar-based approaches in L2 phonology, following: the "Speech Learning Model" (SLM Flege, 1995), the "Perceptual Assimilation Model" (PAM Best, 1993, 1995; Best & Tyler, 2007), the "Second Language Linguistic Perception" (L2LP Escudero, 2005) and the Hybrid model "Second Language Speech" (Colantoni, Steele, & Escudero, 2015).

One major point in the L2 research we have been interested in is how close the learners' productions are from the target (i.e., target-like; phonological) or from the native speaker's productions (native-like; variable speech). When examining learners productions on a scale going from target-likeness to native-likeness, one can easily observe that some learners achieve a target-like production, with partial native-like productions (mostly within L2 prosody research) or those being close to native-like productions, with least target-like (e.g., learners of English spoken in regions other than SSBE). Clearly, gradiency in L2 phonology is important to allow for an accurate evaluation of the learner's variety and how close/far it is from the target language. The varying levels of the learners will lead to some of them showing a clear assimilation of their production to the target language, whereas others will show deviations that are closer to their native language. Looking at this dichotomy leads to issues in quantifying and evaluating the productions of learners who are in the middle: those producing enough distinct forms from their native languages, which are still far from those of the target language. We look at this as a case of gradient productions, because in many instances, these productions will be identified as very close to being target-like and/or native-like but are different from those in the native languages. In what follows, I briefly describe in a chronological order the various activities conducted within this area with my PhD students, with details of publications emerging from these activities.

4.8.4.2 Summary of work with PhD students

Y. Li (2015) evaluated in her PhD the impact of audio only vs audio-visual information on performance of Mandarin L2 learners of English during the production and perception of English dental fricatives /θ/-/s/ and /ð/-/z/. 29 out of the 42 participants had clear difficulties in producing and perceiving these contrasts, and they were enrolled in a 9-week audiovisual perception training. When comparing the performance of the experimental group to that of a control, during and at the end of the training, the experimental group showed an improvement in their perceptual identification and production of the contrasts. This study highlighted the benefits of the audiovisual perceptual training on performance.

Kitikanan and Al-Tamimi (2012); Kitikanan, Al-Tamimi, and Young-Scholten (2012) looked first at the earliest stages of the voiceless fricative acquisition among Thai learners of Mandarin Chinese. The learners participated in a 14-hours teaching of Mandarin over a period of one-month during which they received two hours of tuition, twice a week. The results of these studies, which combined both auditory and acoustic analyses (spectral peak, and spectral moments) showed that the Thai learners in comparison to native Mandarin speakers produced the similar sounds (i.e., labiodental and alveolar) with 100% accuracy from an auditory and acoustic analysis point of view; the new sounds, such as retroflex and alveolo-palatal fricatives, were variable; and all velar fricatives were realised as glottal fricatives.

In the same vein, Kitikanan (2016) explored in her PhD the role of input and extralinguistic factors (e.g., age, motivation, etc.) affecting the learning of British English Fricatives by Thai learners. The results showed that motivation was one of the major extralinguistic factors predicting learners' productions, and that both new and phonologically similar sounds were variable in how they were acquired by learners. Using a production task, followed by an auditory-perceptual judgment by L1 English listeners' impression of the target-likeness and native-likeness of English fricatives as produced by Thai learners, Kitikanan (2016) evaluated similarities and differences between fricatives in both languages.

Specifically, Kitikanan, Al-Tamimi, and Khattab (2015, 18^{th} ICPhS) looked at how /s/, a phonologically similar fricative between English and Thai was produced by L1 English speakers (20 speakers, 10 males and 10 females; various accents), Thai L2 learners of English (50 speakers; 23 males and 27 females; less than 10 months in the UK) and L1 Thai speakers' productions (20 speakers; 10 males and 10 females). The acoustic results pointed to differences between L1 English and L1 Thai productions of the /s/ sound, whereby the former had lower spectral peak and centre of gravity than the latter, indicating a potential difference in place of articulation between a more alveolar place for English and a more dental place for Thai. Interestingly, L2 English productions were always in the middle between the two extremes and across gender. Clearly, L2 learners are starting from their L1's productions, signalled by an acoustically higher spectral peak and centroid, but they are diverging from their L1 to try and reach an acoustic quality similar to their L2, but without perfectly producing an /s/ similar to that of the targetted L2. The L2 /s/ sound was judged at 97.8% as the "same" sound as that found in L1 English by the 20 English judges (10 males and 10 females), indicating that auditory-perceptual analysis failed to differentiate it from the L1 English or L1 Thai productions. Clearly, phonologically similar sounds judged as target-like are not systematically produced in an identical manner by leaners, which is likely to induce the non-native-like percept of the productions of learners. This is an important aspect to consider to support learners achieving a target-like and a native-like auditory and acoustic quality.

One of the challenges to learners is acquiring both the segmental and suprasegmental systems of the target language. For the latter, Y. Chen (2020) in her PhD thesis explored the effect of explicit instructions of prosody of Chinese learners of British English, in an intonation training programme (over five weeks, 2.5 hours for each of six sessions, with 60-75 mins of explicit instruction followed by 60-75 mins of self -paced practice). The training programme focused on the informational, grammatical, and pragmatic functions of intonation following the British school of intonation, focusing on the 3Ts (tonicity, tonality or tone) that are properly controlled at the phonetic level. The results of the thesis were presented in Y. Chen, Khattab, and Al-Tamimi (2016, 8_{th} New Sounds) and as a 5-pages paper in Y. Chen, Khattab, and Al-Tamimi (2019, 19^{th} ICPhS). In fact, one of the main hypotheses advanced by Y. Chen et al. (2019, 19^{th}

ICPhS, see Y. Chen, 2020 for more details) was that the more explicit the training is, the more accurate responses on can obtain. In this study, two types of explicit instructions were administered: one using audio information only using Audacity (audacity group) and one using audio-visual information using Praat, who were able to view the annotated pitch curves of the sample recordings, produce their own voices and listen back to their own productions to compare with the sample recordings.

Y. Chen et al. (2019) presented the results of the training study, where 60 Chinese EFL learners were randomly divided into three groups: a control group only receiving the main instructions, and two experimental groups (one Audacity and one Praat). Comparing their performance to that of 10 native self-claimed RP speakers at the pre-test (on week 1), it was clear that the natives outperformed the learners on Accent, Phrasing and Tone and that the three experimental groups did not differ from each other in their performance (see Y. Chen et al., 2019, Figure 1, pp. 2288). Comparing the performance of the two experimental groups to that of the controls at the post-test (week 5) and at the delayed post-test (week 14), the results showed an overall improvement for the two experimental groups over the three levels of prosodic training, in comparison to the controls. The two experimental groups improved at the delayed post-test, in comparison to the controls, who actually either declined and/or stabilised. The results point to the role of explicit instructions at consolidating knowledge that was generalised at the delayed post-test. Interestingly, and contra to our predictions, the two experimental groups improved in a similar fashion, suggesting that audio-visual training does not outperform auditory-alone training for the teaching and learning of intonation meanings; the improvement is due to the explicit instructions of our training.

In her PhD thesis, Ehbara (2021) evaluated the impact of teaching methods on the acquisition of English phonemes by Libyan infants (aged 7 years old). This work was also presented at a conference in Ehbara and Al-Tamimi (2018, BAAP) and then as a book chapter in Ehbara, Young-Scholten, and Al-Tamimi (2021). The main aims of the thesis were to compare the impact of 1) presentation method, that is perception-only vs. perception and production practice, and 2) input type, that is native English CAL (Computer Assisted Learning) vs. Arabic-accented using traditional class teaching. The study focused on children aged 7 years old with the aim to evaluate whether age of acquisition has any impact on learning a foreign language. In addition, chances of finding child participants with no prior instruction in English are higher than those for adults, whose aspects of L2 language can potentially be fossilised. Hence, this study focused on 58 Libyan-Arabic children aged between 6;11 and 8;0 (+2 months in the delayed post-test). All were reported to have had no prior instruction in English before the training. They were divided into three groups: Traditional instructions (11 boys and 9 girls), CAL Listen-only (10 boys and 8 girls) and CAL Listen and speak (9 boys and 11 girls).

The three groups followed a three-weeks, five hours per week, training programme. The traditional group followed a three-week "regular" teaching in the classroom and the teacher was instructed to include the test words and the same pictures as the experimental conditions within the teaching activities (e.g., grammar tasks or alphabet drills, but no electronic devices). For the Listen and Speak group, each word was allocated a three-minute listening practice and one minute speaking practice. For the Listen-only group, each word was allocated an additional listening minute instead of the minute speaking practice; i.e., totalling 4 minutes listening. This was to eliminate the impact of differences in time allocated for each word between the Listen and Speak and Listen Only conditions. The Listen and Speak and Listen Only groups used native speaker input provided by a programme called the Digital Literacy Instructor (DigLin). Given that none of the children had prior instructions in English, it was impossible to perform a pre-test. To evaluate the performance at the end of the training, three tasks were used to elicit data: picture-naming, read aloud and delayed repetition. Visual stimuli (photographs) for the delayed post-test were different from those for the post-test and the training visual materials to avoid task familiarity. In addition, children were instructed to produce Arabic plosives to acoustically compare differences between their productions in both English and Arabic.

The results presented in Ehbara et al. (2021) highlighted that both experimental groups outperformed, albeit non-statistically significantly, the traditional group. The Listen and Speak group outperformed the Listen-only group only in affricates. The Listen-only group was outperformed by the traditional classroom instructions, possibly due to type of input and medium of instruction. VOT results showed a clear similarity between L1 Arabic and L2 English productions by children in the three groups; likely suggesting that children did not acquire the most important correlates to differentiating the two languages as the productions of the Voiced set in both languages was with a negative VOT and with a marginal positive VOT for the Voiceless set in English (see Ehbara et al., 2021, Figure 2, pp. 33). Overall, the conclusions from this study were that the output practice has an advantage over delayed production after only three weeks of training particularly in less marked sounds. The results also showed that the short time during which children acquired this new language, i.e., a three-week training with 15-hours total instructions may not be enough to warrant an accent-free pronunciation when learning English before the age of puberty.

Looking at the learners' variety in learning the prosodic cues in the L2, Kamphikul (2022) in his PhD thesis examined the prosodic differences between Thai learners of L2 English (20 speakers; 10 males and 10 females) in comparison with those of L1 English (20 speakers; 10 males and 10 females; SSBE). The aim of the thesis was to document the differences between the two groups in marking narrow focus by using variable accent types, and deaccenting, in addition to providing acoustic evidence for how each group marks narrow focus, tonal alignment and scaling of rising accents associated with marking focus in question, following the predictions of the Autosegmental-Metrical (AM) approach to intonational analysis (M. E. Beckman & Pierrehumbert, 1986; Ladd, 2008; Pierrehumbert, 1980). The results of this thesis pointed to similarities between the two groups, whereby native speakers and Thai learners used rising accents to mark narrow focus, in addition to deaccenting and rephrasing. However, they differed in the type of acoustic correlates used to express focus, in tonal alignment and scaling of rising accents which they used to mark focus[27].

As highlighted previous, learners find it often challenging to acquire the segmental and suprasegmental systems of the target language, and in many instances, the various training methods employed can increase the target-likeness and native-likeness of the learners' productions in their L2. For segmental and using the predictions of the high variability phonetic

[27]Although not on L2 speech, H. A. Moussa (2019, see as well H. Moussa & Al-Tamimi, 2018, In preparation) documented how native speakers of Jeddah Arabic (20 speakers; 10 males and 10 females) marked narrow and broad focus by using the predictions of the Autosegmental-Metrical (AM) approach to intonational analysis, combined with various temporal and spectral acoustic measures. The results pointed to clear differences in how native speakers marked both narrow and broad focus in terms of differing durations, *f*0 peaks and excursion size, intensity, in rising speed, in deaccentuation and in phrasing. Table 98 (see H. A. Moussa, 2019, pp. 157) demonstrated that Jeddah Arabic differed from other Arabic dialects in marking focus.

training approaches (HVPT, Logan, Lively, & Pisoni, 1991), using highly variable natural speech material, also known as high-variability (HV) training (P. Iverson, Pinet, & Evans, 2012), or high variability pronunciation training (HVPT, Thomson, 2019), Alghabban (In preparation) is currently looking into the role of HVPT training on improving learning vowel contrasts in English as an L2 by Saudi Arabic learners[28]. A total of 117 participants (all female subjects) were divided into three groups:

- G1 (38 beginner learners) trained on to L1 SSBE English
- G2 (41 beginner learners) trained on variable L1 English, with standard American English (AmE), standard Australian English (AusE), and SSBE
- G3 (38 beginner learners) trained on variable L1 English, with standard American English (AmE), standard Australian English (AusE), and SSBE, in addition to Saudi Arabic-accented English (SA-E)

All learners followed a sixteen 35- to 55-minute training sessions, over an eight-week period (two sessions per week). They were trained on multiple vowel stimuli produced by 7, 9 or 10 native speakers (for groups G1, G2 and G3, respectively) from each of the varieties listed above. After describing the training activities, and the various tasks, they received a pre-test (week 4), a mid-test (week 9), a post-test (week 14), a generalisation-test 1 (between weeks 15 and 16) and generalisation-test 2 (after 1 to 2 months after training). The preliminary results reported in this thesis are promising as they show an improvement in G3 over G2 or G1 and highlight the importance of the HVPT method combining multiple inputs, including a Saudi Arabic-accented English to improving identification and discrimination of English vowels by beginner learners in a foreign context.

Lastly, the recent work of a current M2 student working under my supervision looked specifically at the acoustic correlates that can predict identification of ironic speech by Mandarin learners of French as a foreign language. Zhou (2023) first examined the cross-linguistic differences between French and Mandarin in marking ironic criticism (IC) and literal praise (LP) as produced by 20 speakers (10 speakers in French and 10 in Mandarin; balanced by age with 5 males and 5 females in each language). Using a predictive modelling approach and employing Random Forests, Zhou (2023, Figure 8, pp. 31) identified clear language-specific patterns, whereby French participants marked IC and LP differently by using voice quality (jitter and shimmer) in addition to $f0$ span, whereas Mandarin participants signalled the difference via speech rate, and mean $f0$ and intensity, among other measures. The second part of the study looked at how learners of French as a foreign language identified IC and LP based on a perception experiment. 42 listeners (21 high proficiency level; 13 females and 21 low to no proficiency level; 7 females) participated in this study. By employing a similar statistical approach to that used in J. Al-Tamimi (2017b); Baumann and Winter (2018), whereby listeners responses were modelled as a function of each of the acoustic parameters identified in the production task (after controlling for speaker, listener and items as random effects), which was then followed by a Random Forests analysis of the Ironic Score (quantified as the rate of ironic judgments). The results pointed to a difference in judgments of the low vs high proficiency groups, whereby the former found it difficult to judge ironic speech from French solely based on prosodic cues (without fully understanding the segmentals), whereas the latter identified ironic speech in both Mandarin and French sentences to

[28]Part of a continued supervision agreement with Newcastle University, UK.

a similar rate. In addition, an overall decrease in mean $f0$ and span, and mean intensity and span where the only predictors emerging as statistically significantly contributing to the perceptual judgments. Exploring the data further, Zhou (2023, Figure 14, pp. 46) showed that the variables identified by the Random Forests analysis to predict ironic judgments were a mixture of cues used in French and Mandarin, with $f0$ span, jitter, speech rate and intensity span as the top predictors to explain the results, likely showing an intermediate usage of correlates to explain the responses. Interestingly, level of proficiency did not prove to be a significant factor to explain the patterns. This research will continue to explore the individual strategies employed by Mandarin learners of English as an L2 and French as an L3 as a PhD project (Zhou, In preparation).

4.8.4.3 Conclusion

The results of the various studies conducted by the M2 and PhD students who worked with me highlighted an important aspect: an L2 learner can achieve a production close to a target-like and a native-like depending on the level of detail in the analysis and exposure to the language examined. Indeed, target-likeness is variable as in most cases, L2 learners try to assimilate their L1 categories to their learnt L2 categories. In doing so, the majority of learners have an intermediate category that can be identified as a "new" sound, evidence of which emerged for vowels (Almurashi et al., 2023, In preparation), fricatives (Kitikanan & Al-Tamimi, 2012; Kitikanan et al., 2015), stops (Ehbara et al., 2021) and prosody (Y. Chen et al., 2019). It is clear from these results that L2 learners are not simply transferring their L1 knowledge to the learnt categories; they are transferring knowledge of the L1 and the target L2 and produce "new" categories of sounds that are intermediate between the two languages. At a later stage of the learning process, these learners can achieve a target-like production and produce an L2 category almost similar to that of the target L2. In terms on native-likeness, this is more complex to achieve as it depends on various aspects. Motivations, exposure, knowledge of the L2, etc., can all impact on how native-like a learner can sound. Looking at phonetic detail allows to disentangle the differences that emerge when comparing the first language users of a language and the learners' variety of this language. As highlighted above, training experiments are crucial as can be seen in the training experiments implemented for prosody and intonation (Y. Chen et al., 2019), consonants and vowels (Ehbara et al., 2021) and an ongoing study on vowels (Alghabban, In preparation). Learners can achieve a native-like performance in particular situations when guided by knowledge of the theoretical underpinning of a system, but also using training techniques (e.g., High Variability training, visual and auditory cues, etc.).

All in all, these various studies put emphasis on the use of FPD to guide identification of the learner's variety and how systematic learners can be in learning the target language.

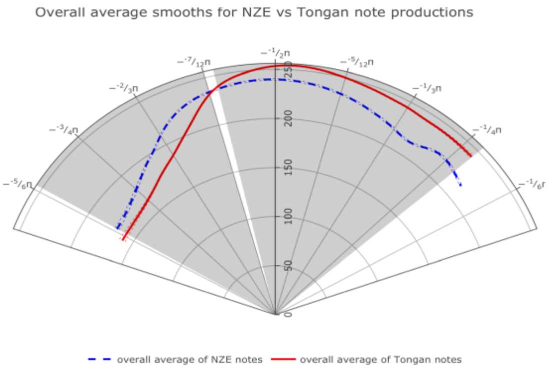
4.8.5 Impact of native language on "preferred" tongue postures

4.8.5.1 Preferred tongue contours

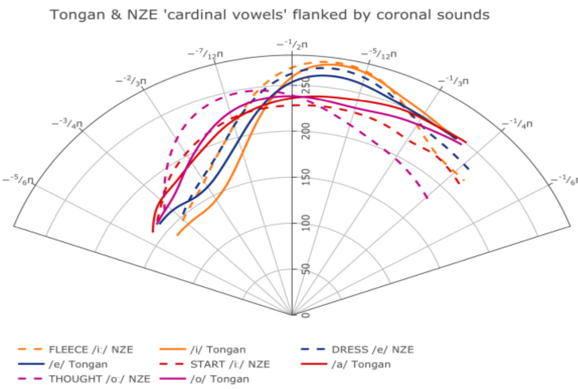
As highlighted in Section 4.5.3.2 (pp. 106), I have employed GAMMs (Generalised Additive Mixed-effects Models) on UTI data to quantify the degree of retraction and overall tongue contour changes associated with back consonants in Arabic. These were explained in terms of the Laryngeal Articulator Model (Esling, 2005; Esling et al., 2019)'s predictions of how the whole tongue contour will be influenced, when the epilarynx is constricted. The preliminary results presented in J. Al-Tamimi (2018a, 16th LabPhon satellite workshop: New developments in speech sensing and imaging) and as part of an open source R notebook J. Al-Tamimi (2022j, first version in 2018; see https://jalalal-tamimi.github.io/R-GAMM-LabPhon18/) led to a collaborative work with Matthias Heyne, University of Pittsburg, USA (and from September 2023 at the State University of New York at New Paltz) and Donald Derrick, University of Canterbury, New Zealand. This work tried to establish if there exists a relationship between the acquired tongue configurations in the production of vowels and the learnt tongue configurations while playing a musical instrument and how the former impacts the latter (work inspired and based on the PhD thesis published in Heyne, 2016). We quantified the full tongue contours obtained from the production of vowels in both New Zealand English (NZE) and from Tongan, in addition to musical notes (varying notes and intensities) from Trombone in both languages.

Heyne, Derrick, and Al-Tamimi (2019, the full R code is published in J. Al-Tamimi, Heyne, & Derrick, 2019 accessible from https://jalalal-tamimi.github.io/GAMM-Trombone-2019/) used GAMMs on the full tongue contours quantified via UTI. By proposing to use a maximal specification model (Barr, 2013; Barr, Levy, Scheepers, & Tily, 2013), which accounts for within and between speaker variation, we evaluated how the tongue configurations from vowel productions and from musical notes played via Trombones differed in the two languages. Figure 4.47 shows some results obtained from the study, specifically on the Notes (Figure 4.47a), the tongue contours obtained for the comparable cardinal vowels in both languages (Figure 4.47b) and the Variability within vowel production in both languages (Figure 4.47c).

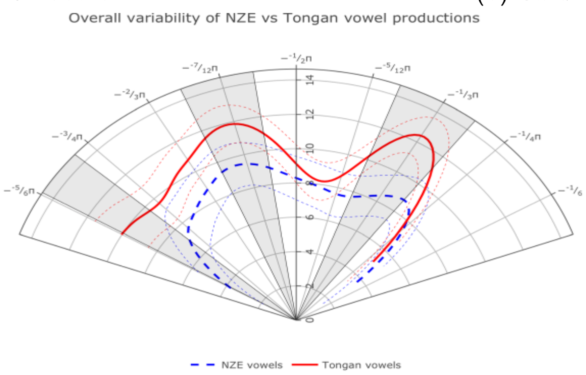
As highlighted in Figure 4.47a, tongue contours in NZE (solid red curves) are more tilted with a more lowered front tongue, and with a more retracted tongue back and root than in Tongan (dashed blue curves). This result takes into account the averaged smooths across all participants, notes, and intensities (for specific contours, see supplementary material for specific hypotheses in J. Al-Tamimi et al., 2019, accessible from https://jalalal-tamimi.github.io/GAMM-Trombone-2019/). Interestingly, Figure 4.47b shows the averaged tongue contours per comparable cardinal vowels across both languages (/iː e ɐː oː/ in NZE; /i e a o/ in Tongan). In each case, the NZE vowel articulations feature a more retracted tongue contour than the one used by the Tongan participants, in agreement with the overall differences observed at the back of the tongue during note productions reported in Figure 4.47a. Finally, we made the hypothesis that Tongan speakers will be more variable in the production of their vowels due to the influence of vowel density reported in previous research, including my own research (see, Section 4.2.2, pp. 54 J. Al-Tamimi & Ferragne, 2005). The expectation is that if a language has more vowels in its inventory, then the vowels will be produced as more extreme with less variability to retain distinctivity between them, opposite to languages with reduced vowel spaces and density. Figure 4.47c shows that overall averaged variance of tongue contours is variable across both languages at the front, middle, and back of the tongue, indicating that the Tongan participants' vowel productions were more variable than those produced by the NZE speakers (similar patterns were reported for the notes variability, see supplementary material for specific hypotheses in J. Al-Tamimi et al., 2019, accessible from https://jalalal-tamimi.github.io/GAMM-Trombone-2019/). During note playing, it seemed that NZE participants favoured an open-like and a back-like tongue position, similar to that of the vowel LOT /ɒ/, though the actual realisation in NZE is more raised and fronted; Tongan speakers seemed to favour a more centralised tongue posture, with a more back production similar to that of the vowel /u/ or /o/.



(a) Overall Smooths



(b) Cardinal Vowels



(c) Overall Vowel Variability

Figure 4.47: GAMMs Smooths in New Zealand English (NZE) and Tongan: (a) Trombone notes averaged smooths in NZE (dashed blue lines) and Tongan (solid red lines); (b) Smooths of Cardinal Vowels in NZE (dashed lines) and Tongan (solid lines); (c) Cardinal vowels variability averaged smooths in NZE (dashed blue lines) and Tongan (solid red lines); grey shadings signal statistically significant differences (reproduced from Heyne, Derrick, & Al-Tamimi, 2019, pp. 13, 16 and 18)

The results of this study provided articulatory evidence that structural differences and vowel density influenced the favoured tongue positions used in speech and Trombone note playing in both languages. The particular vocal tract configuration needed for producing the notes seems to relate the predictions of the Laryngeal Articulator Model (Esling, 2005; Esling et al., 2019). The favoured position in Tongan is similar to that of the back and high vowels, which are more "raised" following LAM, whereas those in NZE seem to be similar to the centralised non-final and final /ə/, which are located at the boundary between central, raised and retracted qualities, following LAM. Our predictions are that these favoured positions are the aerodynamically most optimal positions for Trombone playing, hence correlating these favoured positions to those of the native language.

4.8.5.2 SSANOVAs vs GAMMs on UTI data

The work published in Heyne, Derrick, and Al-Tamimi (2019) was the first step towards systematically using GAMMs over SSANOVAs on UTI data. Our approach relied heavily on specifying a maximal model following Barr (2013) and Barr et al. (2013) which includes random effects for speaker and items, as well as by-speaker and by-item random slopes for the fixed effects of interest and/or of controlling effects (e.g., sex of the speaker). The maximal specification model accounts for within and between-speaker variation, including differing vocal tract lengths and sizes (from UTI perspective), in addition to influences on the acoustics (including pitch, formants, etc.). This approach was also used in J. Al-Tamimi and Palo (2020, 12th ISSP) and J. Al-Tamimi and Palo (2023, 20th ICPhS).

As highlighted in Heyne, Al-Tamimi, and Derrick (2019, Boston Speech Motor Control Symposium); J. Al-Tamimi, Heyne, and Derrick (2020a, BAAP); J. Al-Tamimi, Heyne, and Derrick (2020b, 12th ISSP) and Derrick, Al-Tamimi, and Heyne (In preparation, Laboratory Phonology), GAMMs provide a streamlined and easy to implement approach to quantifying within and between-speaker differences when comparing tongue contours quantified via UTI. In fact, the majority of current research (for a summary, see Derrick et al., In preparation, Laboratory Phonology), still uses SSANOVAs to quantify tongue contours, by using by-speaker averaged contours, followed by comparing patterns across participants (for more details, see Davidson, 2006). While this approach has been extensively used in the literature on UTI, it suffers from Type I error given that random effects, which allow for within and between subject variability to be accounted for are not used[29]. Hence, our claim has been that using GAMMs over SSANOVAs reduces Type I error and provides the analyst with a straightforward way to quantify between-subject differences, after accounting for within and between-subject variations.

To evaluate this further and to complement the results obtained in Heyne, Derrick, and Al-Tamimi (2019), who adjusted the x and y coordinates of the full tongue contour in polar coordinates as a function of the coordinates of the tongue contour of the vowel /i/ obtained from the speaker having the smallest vocal tract (see Heyne, Derrick, & Al-Tamimi, 2019, pp. 9 and Figure 3 for more details on the normalisation procedure). In our case here, we wanted to evaluate which normalisation method allows to obtain a more robust evaluation of the differences, in addition to using raw data (i.e., without normalisation). As such, we normalised the tongue contours as a function of the speaker with the smallest vocal tract producing one of the three vowels /i a o/ in NZE, compared with raw data (i.e., with no normalisation). We then used the resulting x and y coordinates in our subsequent analyses. First, SSANOVAs were obtained for the four configurations, and AR-GAMMs with full maximal specifications.

The results presented in Figure 4.48 show the outcomes of this comparison. This example shows the tongue contours across NZE and Tongan productions of the note Bb2 (116.54Hz) in *forte*, with SSANOVA curves on the top row and the GAMMs on the bottom row, across the four cases (raw or one of the IPA /i a o/ normalisations), comparing the performance across NZE and Tongan. Our hypothesis here is that NZE will show a lowered tongue front and a more back tongue dorsum/root following the findings reported in Heyne, Derrick, and Al-Tamimi (2019). Indeed, this is the overall pattern when one looks at the smooths in Figure 4.48. However, we present a direct comparison of the performance of both SSANOVA and GAMMs. Starting with the SSANOVA curves on raw and normalised data (Figure 4.48, top row, a, b, c and d), the results show similarities in regions of significant difference (shaded areas; 95% CI) at front (right edge) and back of the tongue. There is clearly over-confidence in regions of significance, likely due to inflated Type I error given that random effects are omitted. All normalised datasets perform similarly to raw data, although contours look least "noisy" for /i/ normalisation; contours fit on /a o/-normalised data show additional inflections at edges, especially at the back of the tongue.

[29]It should be noted that the `ssanova` function from the `gss` package (Gu, 2014) allows the inclusion of random effects, via the option `random` although it is not as straightforward as within GAMMs. This is likely to be the reason behin not systematically using it within the community.

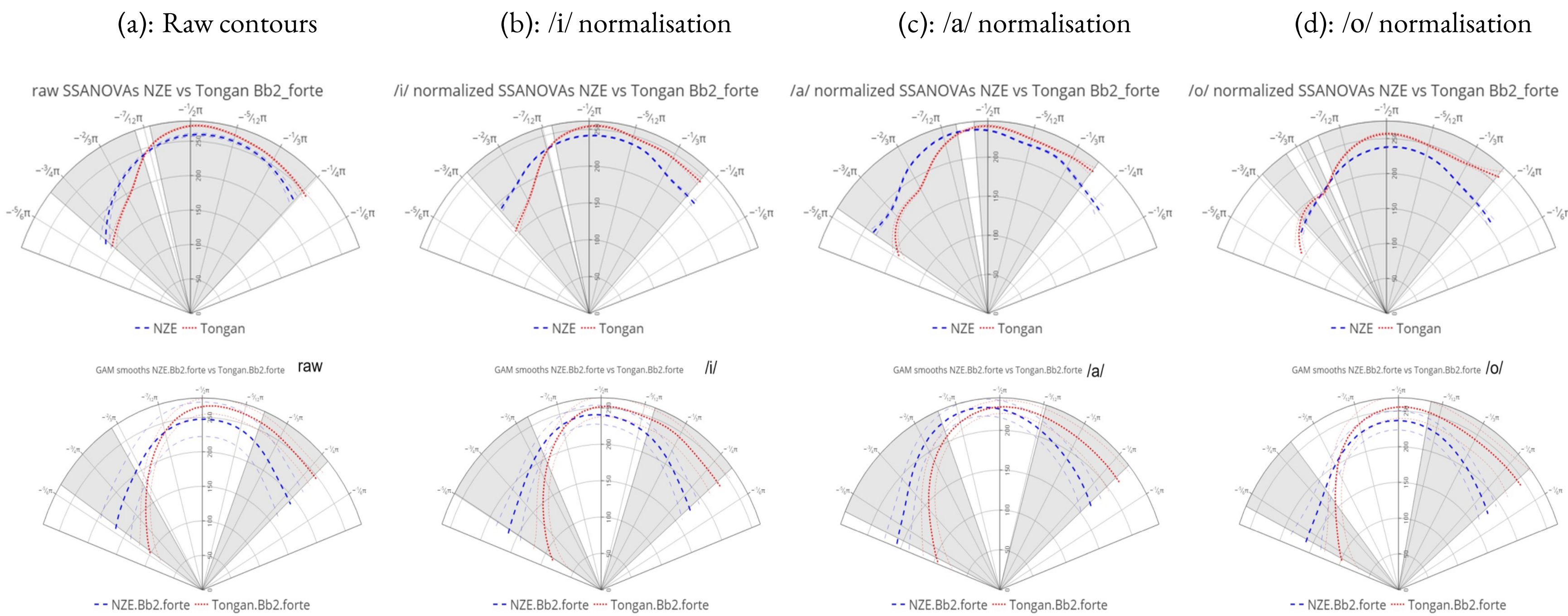


Figure 4.48: SSANOVAs (top row) vs GAMMs (bottom row) smooths for the musical note Bb2 (116.54 Hz) produced at *forte* intensity in both NZE (blue dashed curves) and Tongan (red dotted curves), with significance regions highlighted in grey: (a) raw contours without normalisation; (b) the same contours normalised based on /i/ vowels; (c) the same contours normalised based on /a/ vowels; (d) the same contours normalised based on /o/ vowels (reproduced from J. Al-Tamimi et al., 2020b)

Moving to GAMMs (Figure 4.48, bottom row, a, b, c and d), GAMM smooths display smaller regions of significant difference, likely due to the inclusion of random effects. When comparing the patterns emerging for raw and /i a o/ normalised datasets, the results show that across all notes and intensities, there are clear differences between SSANOVAs and GAMMs, with most variation obtained at the front area; the differences observed at the back area are more consistent across the two method. However, clearly in the case of SSANOVAs, the CIs are much smaller and the regions of statistically significant differences are wider. This shows an over-confidence in the results using SSANOVAs, with increased CIs and reduced significant regions using GAMMs. The results of this study show differences in regions of significant difference for both raw and normalised datasets. Interestingly, raw data findings seem to be more conservative, given that there is a clear reduction in regions of significance, when compared to any normalised dataset. This is likely due to the fact that when random effects are included, adjustments to the coefficients allow for statistical *normalisation* of the data given how random effects work; they allow for individual subjects' overall and specific variations to be taken into account, which is already the case when using raw data. Normalisation using /i/ is less variable than that with /o/ or /a/, with more variations in the former, due likely to different realisation within each language.

GAMMs outperform SSANOVA due to the ease in modelling random effects and accounting for within and between-speaker differences, which allow the emergence of generalisations with confidence. Clearly, differences in vocal tract sizes are an important factor impacting on generalisations of results; the patterns reported here show clearly how random effects structure and importantly the GAMMs framework can help in quantifying between-subject differences in UTI data with confidence. Normalisation techniques need to be adapted to the dataset, which preferably can be done based on an /i/-normalisation or using a bite-plate allowing to quantify the position of the hard palate relatively easily (though this can still be problematic to obtain, and when obtained, can be difficult to use with confidence). If normalisation is not possible, then raw data can be used with confidence, by including a maximally-specified random effects structure adapted to the data and by using GAMMs; the overall patterns will show similarities, albeit with slightly more variation and smaller regions of significant differences indicating a more conservative estimation (for more details, see Derrick et al., In preparation, Laboratory Phonology).

4.8.6 The "Many Speech Analyses" Project

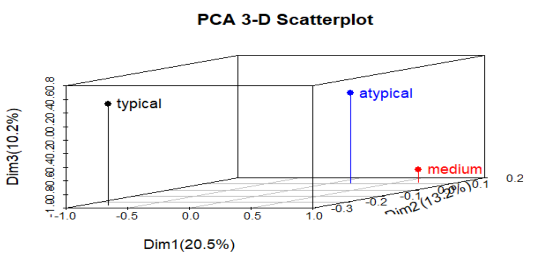
The "Many Speech Analyses" Project (Coretta et al., 2023) was a pre-registered report, which aimed at quantifying how variable researchers are in analysing speech data. A total of 157 researchers from 46 teams participated in this effort to quantify researchers degrees of freedom and how analytical choices influence the interpretation of the results. All teams used the same exact dataset, and were asked to answer an exact same research question. Given the nature of this project, of course, its main aims were not known in advance by research teams; rather they were asked to use any type of acoustic analyses on a specific dataset. The common research question that the teams aimed to answer was: *Do speakers acoustically modify utterances to signal atypical word combinations?* The data came from a production experiment whereby 30 German native speakers produced various types of utterances in three conditions NF (Noun Focus condition), AF (Adjective Focus condition) and ANF (Adjective/Noun Focus condition); the NF condition constituted the experimentally relevant condition, while the AF and ANF conditions acted as fillers. The Adjective and Noun had various combinations, with the NF condition being manipulated with respect to their typicality. The combinations were either typical (e.g. *orange mandarin*), medium typical (e.g. *green tomato*), or atypical (e.g. *yellow cherry*). Each subject produced 15 critical trials (NF condition). Each trial was repeated twice, yielding a total of 30 trials per participant and a grand-total of 900 ($15 \times 2 \times 30$ participants) spoken utterances. The teams had 4 months to complete their analyses, submit their report and all material used to generate the results. Then, they all engaged in a period of peer review, where each team peer-reviewed four other teams' results. They all then participated in peer reviewing the written manuscript and are all co-authors in Coretta et al. (2023).

I coordinated the activities of the team linckia_nattereri (J. Al-Tamimi, Caillol, Ferragne, & Pélissier, 2022, full code available at: https://osf.io/8r6x7/). In our team, we have decided to explore prosodic and voice quality changes associated with the NF condition, after having performed forced-alignment on the data. We started by a working hypothesis: *the more atypical the adjective-noun combination is, the more impact on a predictor there will be in the adjective in comparison with the noun. Hence we expect most of our metrics to display values reflecting more phonetic salience in the atypical context, in comparison with the medium or typical contexts*. To test this hypothesis, we evaluated the impact of 1) segmental (reflected by formant frequencies), 2) suprasegmental (reflected by prosodic measures of intensity and f0) and 3) voice quality measures (energy components in the spectrum) to the phonetic salience of the contexts. Based on our working hypothesis, we predict that if there is indeed an impact of the atypical context that multiple types of changes will be observed on any of the segmental, suprasegmental and voice quality measures. A total of 44 acoustic measures were used for subsequent statistical analyses (J. Al-Tamimi, Caillol, et al., 2022, see details within the submission available at: https://osf.io/8r6x7/). We started by exploring the data using correlation plots, and PCA, followed by Random Forests to reduce complexity in our dataset, then used a Bayesian Logit Mixed-effects Models to evaluate direction and significance of the effect in question.

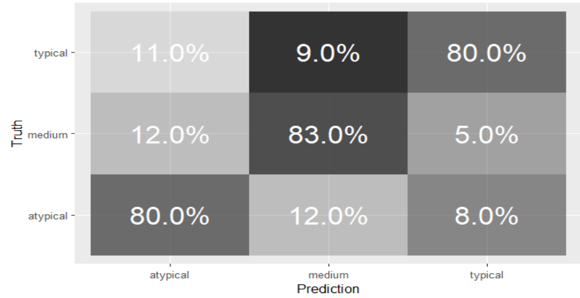
Figure 4.49 shows the results of our contribution to the research question. First, Figure 4.49a shows a 3D PCA plot for the three conditions, after taking into account the multiple acoustic correlates examined here. This shows proximity of the atypical and medium typicality on the first and second dimensions, with closeness between typical and atypical on the third dimension. The intensity mean (and amplitude-based measures) correlated positively well with the first dimension, whereas intensity_SD, F2_SD and HNR_0-3500Hz correlated positively with the second. Most of the HNR measures correlated positively with the third dimension (see supplementary material from J. Al-Tamimi, Caillol, et al., 2022).

Next, we used Random Forests to separate the three conditions. After removing highly collinear predictors, we retained 33 out of the 44 initial measures, and our classification results on the training set were 82%, with an AUC of 0.937 with relatively high sensitivity (0.82), specificity (0.91), and precision (0.82). On the testing set, our model achieved 81% accurate classification score, and an AUC of 0.94. This suggests that the subsequently reported patterns are robust to signal the differences between the three levels of typicality. Figure 4.49b shows the confusions matrice based on the testing set. All three classes were highly accurately classified, with rates close or above 80%; the medium and atypical classes were relatively highly confused where 12% of atypical cases were confused as medium and vice-versa.

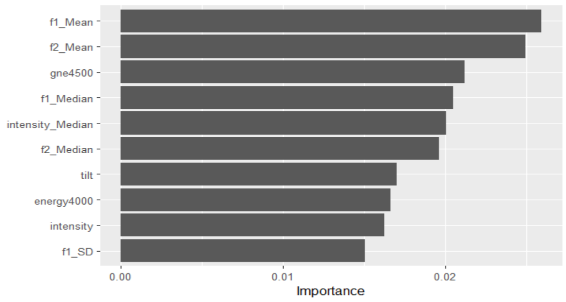
Using Variable Importance Scores presented in Figure 4.49c, the best 10 predictors used by



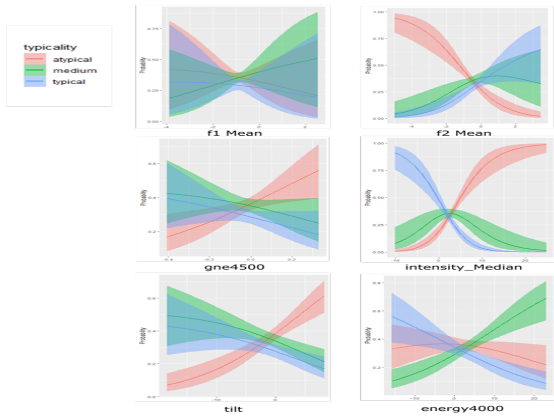
(a) 3D PCA



(b) Confusion Matrix



(c) Variable Importance Scores



(d) Conditional effects plots

Figure 4.49: Results of the "Many Speech Analyses": Team linckia_nattereri: (a) 3D PCA; (b) Confusion Matrix of classification results of the three levels of typicality; (c) Variable Importance Scores of top 10 predictors; (d) Conditional effects plots for the Bayesian Logit Mixed-effects Model on the top 6 predictors (reproduced from J. Al-Tamimi, Caillol, et al., 2022, part of the MSA project; Coretta et al., 2023)

the algorithm were a combination of formant-based, and energy/amplitude-based metrics, with F1 and F2 means being the top two predictors. Looking at the AUC curves per typicality (see supplementary material from J. Al-Tamimi, Caillol, et al., 2022), the typical context showed a perfect model fit, whereas the medium and atypical showed more variation: this highlights stability in how participants produced the typical adjective-nouns combinations, while they were more variable at producing medium and atypical adjective-nouns combinations.

Finally, we used a Bayesian Logit Mixed-effects Model on the top 10 predictors identified by our Random Forests. Figure 4.49d shows the conditional effects plots for 6 of the predictors. The results show F2 (mean and median) to be the best predictors of typicality, with large effect size, and intensity (mean and median), with a small effect size. Tilt, glottal excitation (gne4500Hz) and energy (4000Hz) also contribute to the distinction between typical categories. F1 mean, which is the best predictor identified in our Random Forests, does not seem like a good predictor in this analysis. However, f1 Median distinguishes well between medium and typical categories, which f2 Mean (the second best predictor in the random forests analysis) does not do, and therefore complements the predictive power of f2 and intensity. All in all, our results show participants to be stable when producing typical combinations of adjectives-nouns, while they are more variable in the medium and atypical combinations. The latter shows more variations overall.

An important result of the overall project is that giving the same dataset and the same research question to 46 teams showed that "*[t]he submitted analyses exhibited at least 52 unique ways of operationalizing the acoustic signal alongside 55 unique ways of constructing the statistical model. By multiplying the numbers of acoustic and model specifications, there are in principle 2860 possible unique combinations. Note that this is a conservative estimate of the number of possible analytic choices for our research question, ignoring many other degrees of freedom like e.g. acoustic parameter extraction, outlier treatment, and transformations, all of which might have an impact on the final results*" (Coretta et al., 2023, pp. 39). Researchers were variable in the type of acoustic measures used (temporal, formant, *f*0, intensity, voice quality, etc.), the portion where the measures were obtained (vowel, syllable, word, utterance), the type of statistical designs used (frequentist, Bayesian, LLMMs, GLMMs, GAMMS, Machine learning, etc.).

This study put emphasis on the fact that all research teams will be variable in their analytical choices of the same dataset and the same research question. We need clarity in our analytical methods and in the choices we make; transparency in research is a crucial element to allow for replicability of the findings. Without this, the replicability crisis will continue to emerge in various disciplines and impact on the findings we report!

4.9 Open Access

Since 2018, I have been engaged in transparency and replicability of research and have been instrumental in making my Praat scripts, and R notebooks available to the research community on my GitHub repository and in some instances on my OSF. Some of the material in the list below is supplementary material to published research, or training material for R courses I delivered previously and are now part of my taught courses at Université Paris Cité (at BA and MA levels). The list of items below shows how engaged I have been in this type of activity, which in many instances, led to collaborations with other researchers (see e.g., Section 4.8.5 on using GAMMs on UTI data). Future work is planned with a few researchers on using some of my scripts on Voice Quality estimation for detecting Schizophrenia in French and Italian (using the script published in J. Al-Tamimi, 2022h) or the use of the spectral estimations in fricative consonants on the velarised vs palatalised fricatives in Russian (using the script published in J. Al-Tamimi, 2022e). In terms of increasing usage of the software R, I was one of the founding members of the "Adventures in R" meetup group at Newcastle University, UK, which was financially supported by the R foundation between 2018-2020, where we delivered regular meetings and training events. Then I became engaged globally by co-founding the Global R User Group, with support from the R foundation, which organises regular meetings and online training events. In the future, I am planning a similar local group to enhance usage of the software R in addition to sharing advances in statistical designs. Related to that, I recently joined the "Statistics for Linguistics" network, which aims to bring together researchers and students in linguistics with an interest in developing/disseminating statistical methods and open research/scholarship practices. Hopefully, this involvement will allow researchers in linguistics in various places to share best practices in conducting research activities, transparency and be supported in their statistical choices.

1. J. Al-Tamimi, Caillol, et al. (2022): **Al-Tamimi, J.**, Caillol, C., Ferragne, E., & Pélissier, M. (2022). Many Speech Analyses: Team linckia_nattereri [Full analyses of the Many Speech Analyses project by the team members; contains processing files, Praat scripts to extract acoustic measures, exploratory and confirmatory statistical analyses]. Many Speech Analyses: Team Linckia_nattereri; OSF. `https://osf.io/8r6x7/`

2. J. Al-Tamimi (2022b): **Al-Tamimi, J.** (2022a). Praat-fo-Accurate-Estimation [Open source online material - Praat scripts to accurately estimate fo curves, using the two-pass method and then smoothing (original version 2018)]. Praat-Fo-Accurate-Estimation - Praat Script. https://jalalal-tamimi.github.io/Praat-f0-Accurate-Estimation/
3. J. Al-Tamimi (2022c): **Al-Tamimi, J.** (2022b). Praat-Measurement-points [Open source online material - Praat script to allow for an accurate measurement point (original version 2018)]. Praat-Measurement-Points - Praat Script. https://jalalal-tamimi.github.io/Praat-Measurement-points/
4. J. Al-Tamimi (2022d): **Al-Tamimi, J.** (2022c). Praat-Silence-Detection [Open source online material - Praat script for Silence/Speech detection using the two-pass method for fo estimation and intensity adapted to each speaker]. Praat-Silence-Detection - Praat Script. https://jalalal-tamimi.github.io/Praat-Silence-Detection/
5. J. Al-Tamimi (2022e): **Al-Tamimi, J.** (2022d). Praat-Spectral-Analyses-fricatives-stops [Open source online material - Praat script for computing spectral measurements for fricatives and stops]. Praat-Spectral-Analyses-Fricatives-Stops. https://jalalal-tamimi.github.io/Praat-Spectral-Analyses-fricatives-stops/
6. J. Al-Tamimi (2022f): **Al-Tamimi, J.** (2022e). Praat-Spectral-Tilt [Open source online material - Praat script for computing spectral tilt measurements for Voice Quality]. Praat-Spectral-Tilt - Praat Script. https://jalalal-tamimi.github.io/Praat-Spectral-Tilt/
7. J. Al-Tamimi (2022g): **Al-Tamimi, J.** (2022f). Praat-Voicing-detection [Open source online material - Praat script for Voicing detection using the two-pass method (original version 2018)]. Praat-Voicing-Detection - Praat Script. https://jalalal-tamimi.github.io/Praat-Voicing-detection/
8. J. Al-Tamimi (2022h): **Al-Tamimi, J.** (2022g). Praat-VQ-Measurements [Open source online material - Praat script for computing many acoustic measurements for Voice Quality]. Praat-VQ-Measurements - Praat Script. https://jalalal-tamimi.github.io/Praat-VQ-Measurements/
9. J. Al-Tamimi (2022k): **Al-Tamimi, J.** (2022h). R training [Open source online material aimed as a self-introduction to using R: Introduction, tidyverse, visualisation, inferential statistics (LM, GLM, CLM, LMM), PCA, decision trees, random forests]. R Training - Notebook. https://jalalal-tamimi.github.io/R-Training/
10. J. Al-Tamimi (2022i): **Al-Tamimi, J.** (2022i). R-Estimating Number Of Trees RF [Open source online material - This R script allows to estimate the optimal number of trees required to grow Random Forests via the Conditional Inference Trees framework (initial version in 2018)]. Estimating Number Of Trees RF - R Script. https://jalalal-tamimi.github.io/R-Estimating-Number-Of-Trees-RF/
11. J. Al-Tamimi (2022j): **Al-Tamimi, J.** (2022j). R-GAMM-LabPhon18 [Open source online material - Supplementary Material as an introduction to GAMMs (Generalized Additive Mixed-effects Modelling) on UTI (Ultrasound Tongue Imaging) data. First release in 2018; updated in 2022]. R-GAMM-LabPhon18 - Notebook. https://jalalal-tamimi.github.io/R-GAMM-LabPhon18/
12. Schiel and Al-Tamimi (2021): Schiel, F., & **Al-Tamimi, J.** (2021). WebMAUS and WebMINNI Arabic: Online resources for forced-alignment and speech-to-text transcription [Development of an Arabic variety-independent romanization system integrated within an Arabic variety-independent WebMAUS Forced Alignment and WebMINNI services (integrated with WebMAUS version 3.8)]. WebMAUS and WebMINNI Arabic: https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface
13. J. Al-Tamimi (2021b): **Al-Tamimi, J.** (2021a). Introduction to Random Forests [Open source online material - Introduction to Decision trees and Random Forests]. Introduction to Random Forests - Notebook. https://jalalal-tamimi.github.io/Intro-Random-Forests/
14. J. Al-Tamimi (2021c): **Al-Tamimi, J.** (2021b). Techniques in data analyses [Open source online material - advanced analyses using R: Linear, logistic, cumulative and mixed-effects models]. Techniques in Data Analyses - Notebook. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062021.nb.html
15. J. Al-Tamimi (2020b): **Al-Tamimi, J.** (2020). Techniques in data analyses [Open source online material - advanced analyses using R: Linear, logistic, cumulative and mixed-effects models]. Techniques in Data

Analyses - Notebook. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062020.nb.html

16. J. Al-Tamimi et al. (2019): **Al-Tamimi, J.**, Heyne, M., & Derrick, D. (2019a). GAMM-Trombone-2019 [Open source online material - Supplementary Material as an introduction to GAMMs (Generalized Additive Mixed-effects Modelling) on UTI (Ultrasound Tongue Imaging) data on trombone vs vowel production]. GAMM-Trombone-2019 - Notebook. https://jalalal-tamimi.github.io/GAMM-Trombone-2019/
17. J. Al-Tamimi (2019a): **Al-Tamimi, J.** (2019a). Introduction to R [Open source online material as an introduction R: Introduction, visualisation]. Introduction to R - Notebook. https://jalalal-tamimi.github.io/R-Introduction-to-R/
18. J. Al-Tamimi (2019d): **Al-Tamimi, J.** (2019b). Techniques in data analyses [Open source online material - advanced analyses using R: Linear, logistic, cumulative and mixed-effects models]. Techniques in Data Analyses - Notebook. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062019.nb.html
19. J. Al-Tamimi (2018d): **Al-Tamimi, J.** (2018a). R-Rating-data [Open source online material - Supplementary material as an Introduction to Cumulative Logit Regression]. R-Rating-Data - Notebook. https://jalalal-tamimi.github.io/R-Rating-data/
20. J. Al-Tamimi (2018e): **Al-Tamimi, J.** (2018b). R-Voicing-Gemination-VOT [Open source online material - Supplementary Material for the statistical analyses of mixed-effects models and random forests]. R-Voicing-Gemination-VOT - Notebook. https://jalalal-tamimi.github.io/R-Voicing-Gemination-VOT/
21. J. Al-Tamimi (2018f): **Al-Tamimi, J.** (2018c). Techniques in data analyses [Open source online material - advanced analyses using R: Linear, logistic, cumulative and mixed-effects models]. Techniques in Data Analyses - Notebook. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062018.nb.html

4.10 Conclusion

This chapter presented a summary of the research activity I have engaged in during the last 21 years of my career, with most of the major activities emerging in the last 10 years, first by exploring the dynamic nature of vowels, next by looking at gemination and the role of the feature [+Tense] in adult and child speech, then by exploring the acoustic correlates of the Voicing contrast in Arabic, then by looking at the articulatory and acoustic correlates of a gradient epilaryngeal constriction in guttural consonants. We then looked at the role of automated methods in speech processing in supporting diagnosis of cognitive disorders, the development of a forced-alignment system for Arabic and then concluding by a presentation of the various additional collaborative work I have engaged in.

My research activities put emphasis on the role of FPD by highlighting how variable and systematic speakers, listeners and learners are in their production, identification and learning journeys. This variability and systematicity are at the heart of the phonetic-phonology relationship because they allow for an accurate representation of the gradient nature of speech in general. Some aspects of my research tried to provide a gradient view of phonology as this better fits with the predictions of exemplar-based approaches. By focusing on the gradient nature of phonology, one can link speech production and perception to each other in an easier and more direct way, as the claim advanced here is that speakers use a rich production system to mark their individuality, which is then picked up by listeners to allow for an accurate identification of various patterns. Even perception is gradient, as highlighted in many recent research activities, which shows that speech perception is better represented as gradient rather than categorical (McMurray, 2022).

Chapter 5

Significant Activities as Principal Investigator (PI)

5.1 Gutturals and automatic methods in Arabic

As highlighted within Section 4.5, since I started working as an independent researcher in 2013, I led on this major activity of my research by examining the secondary correlates of Voice Quality changes first in pharyngealised consonants in Jordanian and Moroccan Arabic and then by examining the multiple articulatory and acoustic correlates of the gradient epilaryngeal constriction in gutturals consonants in Levantine Arabic. As highlighted previously, these results constitute a direct empirical evidence for the legitimacy of the gutturals as a natural class by employing non-invasive and indirect measures. These are the most significant publications related to this activity:

1. **Al-Tamimi, J.**, & Palo, P. (2023). Dynamics of the tongue contour in the production of guttural consonants in Levantine Arabic. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*. Prague, Czech Republic (7-11 August 2023).
2. **Al-Tamimi, J.** (2021). Acoustic correlates of epilaryngeal constriction in Levantine Arabic guttural consonants. *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).
3. **Al-Tamimi, J.** (2020). The role of the epilarynx in the production of guttural consonants in Levantine Arabic. *Speech science forum, hearing and phonetic sciences – University College London, UK (November 2020)*
4. **Al-Tamimi, J.**, & Ferragne, E. (2020). The phonetic basis of the guttural natural class in Levantine Arabic: Evidence from coarticulation and energy components using Deep Learning and Random Forests. *Proceedings of the 17th LabPhon (Laboratory Phonology) Conference*. Vancouver, Canada (5-9 July 2020).
5. **Al-Tamimi, J.** (2019). The role of the epilarynx in male/female and dialect discrimination in dialectal Arabic/Le rôle de l'épilarynx dans la discrimination entre locuteurs hommes/femmes et entre dialectes arabes. *Séminaire de recherche: Traitement du langage parlé (TLP), laboratoire d'Informatique pour la mécanique et les sciences de l'Ingénieur (LIMSI) - Université Paris-Sud, France (May 2019).*
6. **Al-Tamimi, J.** (2018). A Generalised Additive Modelling approach to ultrasound tongue surface: Quantifying retraction in Levantine Arabic back consonants. *Proceedings of the 16th LabPhon (Laboratory Phonology) Satellite Workshop: New Developments in Speech Sensing and Imaging.* Lisbon, Portugal (23 June 2018).
7. Khattab, G., **Al-Tamimi, J.**, & Alsiraih, W. (2018). Nasalisation in the production of Iraqi Arabic pharyngeals. *Phonetica, 75(4)*, 310–348. https://doi.org/10.1159/000487806
8. **Al-Tamimi, J.** (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8(1)*, 1–40. https://doi.org/10.5334/labphon.19
9. **Al-Tamimi, J.** (2015). Spectral tilt as an acoustic correlate to pharyngealisation in Jordanian and Moroccan Arabic (Article: 0436). *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015).

5.2 The feature [Tense] in Arabic

My work within this activity allowed to unravel new empirical evidence that the feature [Tense] is active in Arabic phonology. The majority of previous research in Arabic phonology identified the duration of the medial consonant as the major (and in most cases, only) relevant correlate for the gemination contrast. As highlighted in Section 4.3, the research I led and collaborated on during my post-doctoral activities highlighted the importance of the secondary correlates in the production of gemination in adult speech in addition to child speech. These are the most significant publications related to this activity:

1. **Al-Tamimi, J.**, & Khattab, G. (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese

Arabic: The case of fricative consonants. *The Journal of the Acoustical Society of America, 138(1)*, 344–360. https://doi.org/10.1121/1.4922514

2. Khattab, G., & **Al-Tamimi, J.** (2015). The acquisition of gemination in Lebanese Arabic children (article: 0870). *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*. University of Glasgow., Glasgow, UK (10-14 August 2015).
3. Khattab, G., & **Al-Tamimi, J.** (2014). Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 5(2)*, 231–269. https://doi.org/10.1515/lp-2014-0009
4. Khattab, G., & **Al-Tamimi, J.** (2013). Influence of geminate structure on early Arabic templatic patterns. *In M. Vihman & T. Keren-Portnoy (Eds.), The emergence of phonology: Whole-word approaches and cross-linguistic Evidence* (pp. 374–414). Cambridge University Press.
5. **Al-Tamimi, J.**, & Khattab, G. (2011). Multiple cues for the singleton-geminate contrast in Lebanese Arabic: Acoustic investigation of stops and fricatives. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, 212–215.
6. Khattab, G., & **Al-Tamimi, J.** (2008). Durational cues for gemination in Lebanese Arabic. *Language and Linguistics, 22*, 39–55.

5.3 The Voicing profile in Arabic

My work on describing the Voicing profile in Arabic in the light of the gradient privative features account advocated by J. Beckman et al. (2013) is a second major development in my activities as it highlights the links between FPD and gradiency in Phonology in a systematic way. Although the first part of this activity is related to gemination, the work below emerged after becoming an independent researcher in 2013 and continued as a PhD supervision activity. These activities put emphasis that Arabic, a Voicing language, is not simply specified with the feature [±Voice]; it requires additional active features, which are dialect-specific (e.g., [Tense], or [SG]). These are the most significant publications related to this activity:

1. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (In preparation). Voicing Contrast and Voicing Assimilation in Najdi Arabic. *Laboratory Phonology: Journal of the Association for Laboratory Phonology.*
2. Dallak, A., Khattab, G., & **Al-Tamimi, J.** (2023). Obstruent Voicing and Laryngeal Feature. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS).* Prague, Czech Republic (7-11 August 2023)
3. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (2022). Voicing contrast in Najdi Arabic stops: Implications for Laryngeal realism. *Proceedings of the 18th LabPhon (Laboratory Phonology) Conference.* (23-25 June 2022).
4. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (2019). The acoustic properties of laryngeal contrast in Najdi Arabic initial stops. *In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS) (pp. 2051–2055).* Canberra, Australia: Australasian Speech Science and Technology Association Inc.
5. Al-Gamdi, N., **Al-Tamimi, J.**, & Khattab, G. (2019). The acoustic properties of laryngeal contrast in Najdi Arabic initial stops. *In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS) (pp. 2051–2055).* Canberra, Australia: Australasian Speech Science and Technology Association Inc.
6. **Al-Tamimi, J.**, & Khattab, G. (2018a). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics, Invited Manuscript for the Special Issue of Journal of Phonetics, "Marking 50 Years of Research on Voice Onset Time and the Voicing Contrast in the World's Languages" (Eds., T. Cho, G. Docherty & D. Whalen), 71*, 306–325. https://doi.org/10.1016/j.wocn.2018.09.010
7. **Al-Tamimi, J.**, & Khattab, G. (2018b). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate plosives. *Proceedings of BAAP (The British Association of Academic Phoneticians Colloquium)*, Canterbury University, Kent, UK (12-14 April 2018).

8. Khattab, G., & **Al-Tamimi, J.** (2018). When lenis and fortis properties collide: The case of voiced geminates in Lebanese Arabic. *Proceedings of the 16th LabPhon (Laboratory Phonology).* Lisbon, Portugal (20-22 June 2018).

5.4 Forced Alignment system for Arabic

As highlighted in Section 4.7, the development of the Forced-Alignment system for Arabic was another major activity I led on since 2015. Due to the lack of an easy to use and open access system to provide automated alignment of the signal and pre-transcribed data for Arabic, I work in collaboration with various researchers to develop both a romanisation system: `Al-Tamimi Romanisation System; ATR` and a new macro, dialect-independent acoustic model for Arabic dialect. This led to integration of an Arabic WebMAUS and WebMINNI services. The two services are available as open-access online resources. In total, this activity led to two major publications, with subsequently planned activities:

1. **Al-Tamimi, J.**, Schiel, F., Khattab, G., Sokhey, N., Amazouz, D., Dallak, A., & Moussa, H. (2022). A Romanization System and WebMAUS Aligner for Arabic Varieties. *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), © European Language Resources Association (ELRA),* Licensed under CC-BY-NC-4.0, 7269–7276. http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.789.pdf
2. Schiel, F., & **Al-Tamimi, J.** (2021). WebMAUS and WebMINNI Arabic: Online resources for forced-alignment and speech-to-text transcription [Development of an Arabic variety-independent romanization system integrated within an Arabic variety-independent WebMAUS Forced Alignment and WebMINNI services (integrated with WebMAUS version 3.8)]. WebMAUS and WebMINNI Arabic: https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface

5.5 Speech as a bio-marker? Role of automated methods

As highlighted in Section 4.6, I have been interested in the use of automated methods in support of diagnosis of cognitive disorders, by examining speech as a potential bio-marker. During this activity, I have either directly worked on: 1) primary data (Schizophrenia) for both automatic prosodic profiles, statistical data analyses, writing up and presentation of work; 2) advising on automatic prosodic and VQ profiles and applying statistical techniques on results obtained from automatic method (on the various stages of Alzheimer and pre-AD), in addition to the use of Prosogram in a large-scale study on cognitively unimpaired older adults with and without $\alpha\beta$, and 3) by supervising an internship on differentiating Alzheimer from Controls using openSMILE in the ADReSS challenge dataset. In total, this activity led to four Articles in peer reviewed journals (one accepted, one under review and two in preparation); four conference proceedings, and three invited speaker talks, with one planned in 2025 as a thematic workshop:

1. **Al-Tamimi, J.** (2025). Classical and automatic methods in speech processing guiding diagnosis of cognitive disorders. *Invited Speaker on a 3-week thematic workshop at the "Institut Henri Poincaré", focusing on novel, rigorous statistical methodology for the analysis of speech and text signals; applications in the medical and political sciences..*
2. He, R., Chapin, K., **Al-Tamimi, J.**, Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (2023). Automated classification of cognitive decline and probable Alzheimer's dementia across multiple speech and language domains. *American Journal of Speech-Language Pathology..* https://doi.org/10.1044/2023_AJSLP-22-00403
3. He, R., **Al-Tamimi, J.**, Sánchez-Benavides, G., Montaña-Valverde, G., Gispert, J. D., Grau-Rivera, O., Suárez-Calvet, M., Minguillon, C., Fauria, K., Navarro, A., & Hinzen, W. (Under Review). Atypical cortical hierarchy

in Aβ-positive older adults in the context of atypical speech prosody. *Neurobiology of Aging*.

4. **Al-Tamimi, J.**, Lofgren, M., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (In preparation). Automated prosodic and voice quality profiles to distinguish between groups at elevated risk of Alzheimer's disease.
5. Lofgren, M., **Al-Tamimi, J.**, Chapin, K., García-Gutiérrez, F., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J., Sanabaria, A., Alegret, M., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (Under Review). An automated acoustic-prosodic analysis pipeline for identifying early cognitive decline and Alzheimer's disease. *Journal of Psychiatric Research*.
6. Jordan, E., & **Al-Tamimi, J.** (2023). Classification d'Alzheimer à partir de paramètres acoustiques et prosodique avec de l'apprentissage automatique. *Actes Des 9èmes Journées de Phonétique Clinique (JPC2023): 'Prendre La Mesure de La Parole'.* Toulouse 15-17 juin 2023.
7. **Al-Tamimi, J.** (2023). Différents stades d'Alzheimer: Rôle des méthodes automatisées de traitement de la prosodie et de la qualité de la voix. *Actes Des 9èmes Journées de Phonétique Clinique (JPC2023): 'Prendre La Mesure de La Parole'.* Toulouse 15-17 juin 2023.
8. Jones, W., **Al-Tamimi, J.**, & Hinzen, W. (2021). The role of automatic acoustic profiling of prosody in schizophrenia. *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).
9. Jones, W., & **Al-Tamimi, J.** (2020). Using Prosogram's automatic prosodic profiling in diagnosing schizophrenia. *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).

Chapter 6

Research Management and co-supervision

6.1 Research activities

As highlighted on my CV (Chapter 1) and throughout the summary of the main activities in which I was engaged (Chapter 4), I have had the chance to work with various researchers to reach my current research profile.

In terms of research activities, I was part of two research projects during my PhD between 2002 and 2008 (funded by the ANR and the "Ministère de la Recherche et de l'Enseignement Supérieur"), where I contributed to data design, collection and analyses for Arabic and French vowels assessing phonological complexity and the role of structural differences on the organisation of vowel systems.

When I became a post-doctoral researcher at Newcastle University in December 2007, I joined four major projects. The first two (between 2007 and 2010 and then 2009 and 2013), which were funded by the ESRC (Economic and Social Research Council, UK) were related to the work I engaged in on the feature [±Tense] in Arabic geminate consonants in both adults and children. On these two projects, I led on data transcription using both Praat and Phon, in addition to acoustic, auditory, phonological and statistical analyses. At the same time, and between 2011 and 2014, I was part of a European Union funded research network "NorPhlex" being part of the UK team (representing York and Newcastle sites). Finally, between 2012 and 2015, I joined an international research project funded by the Australian Research Council (ARC), Australia, where I led on the research activities for the York and Newcastle sites (the project involved various researchers from Australia, New Zealand, and the UK, from London, York and Newcastle). On both sites, I led on the recruitment of participants, data recordings, transcriptions, splicing and preparation of the stimuli for the subsequent perception experiments. I also managed the project's finances for the two sites.

Then, after starting as a lecturer at Newcastle University in September 2013, and then since joining the Université Paris Cité as an associate professor in October 2021, I engaged in various research activities, on my own and collaboratively, where I successfully secured various funding opportunities for either research, acquiring/upgrading equipment, and supporting MA and PhD students secure funding from various sources. The total of the secured funding at Newcastle University was of about £382k (£82k own funding; £300k supporting PhD applications) and since joining the Université Paris Cité is of about €233k (€96.9k own funding; €136k supporting MA and PhD applications).

Starting with research funding, in total, I led on eight research projects: five received external funding, and three with internal funding (total of £57k and €15.4k for projects + £5k and €4k for conferences) which were funded from various sources: UK British Academy/Leverhulme small grant; UK Leverhulme International Academic Fellowship; Faculty Research Funds from Newcastle University, UK; Personal Research account, Newcastle University, UK; AAP BRIO (Aide à la Recherche: Budget Restreint, Impact Optimal), Université Paris Cité, France and the Laboratoire de Linguistique Formelle, France.

During this time and as a recognition for my contributions, I was invited to join two major research projects (both with unsuccessful funding outcomes), one as a co-PI on an European funding (total €1,509k) and another as an ARC funding (total AUD $216k). I was also invited as a consultant on three projects: two unsuccessful (in the UK: £250k and in France: €428k) and one successful (total €10k). For all of these applications, I put forward my expertise in Arabic phonetics and phonology, in addition to experimental methods and statistical designs.

I also engaged in securing funding to upgrade and/or acquire new equipment either at Newcastle University or at the Université Paris Cité. I provided support and advisory role to upgrade the articulatory equipment at Newcastle University (new UTI machine, new portable EGG, microphones, software; totalling over £20k). Then and since joining the Laboratoire de Linguistique Formelle and the Université Paris Cité, France, I worked closely with various colleagues to secure funding in order to increase visibility of the phonetics and phonology research activities of our lab. For instance, I applied and secured funding to acquire specific articulatory equipment: €1.5k for a new portable EGG machine and software from the LLF; €11k to acquire a new portable UTI machine with software and new flexible helmet with integrated lip camera (from the settling-in funds, LLF and CLILLAC-ARP); and most recently, I secured funding of about €65k with various colleagues to acquire a new EMA (Electro Magnetic Articulatography) machine, via an AAP Petits & Moyens Equipement funding scheme, support from two labs (LLF and CLILLAC-ARP) and the Labex EFL.

Finally, I provided support to various Master and PhD applications to obtain funding from various funding bodies. This resulted in five PhD students at Newcastle University, UK, to secure funding via the NINEDTP funding scheme (£60k each; total of £300k); two Master students at the Université Paris Cité, France to obtain funding from the Labex EFL (€3k each; total of €6k) and two PhD at the Université Paris Cité, France to secure funding from the Chinese Scholarship Council (CSC) for 48-months each (€65k each; total of €130k).

In terms of leadership in research activities, and since September 2022, I became co-director of the Experimental Linguistics Research Strand, where I co-led on the writing up of the trajectory of the LLF's project from 2025, by organising various meetings and consulting with colleagues working within that strand. In addition, and since becoming a co-director of the Strand 1 "Complexity in Phonetics and Phonology" of the Labex EFL in November 2022, I have engaged in various activities with the other directors to prepare for our renewal project, which will be for the next 5-10 years, in the order of about €1,000k per year across research, training, outreach and various other research activities. We are currently working on consulting with various researchers and on the writing up of the future project, which should start in January 2025.

The summary above shows my continued engagement with research activities via funding applications and funding for Master and PhD students, in addition to my administrative and teaching duties presented in my CV.

6.2 RAs on various projects

During the various research activities summarised in the summary of my previous research activities (Chapter 4) and above (see as well my Curriculum Vitae in Chapter 1), I have had the chance to work with various Research Assistants/Associates and PhD students. For the development of the Forced Alignment system for Arabic, I have had the chance to collaborate with various researchers who shared their primary data with me to develop the Arabic WebMAUS (see full authorship in J. Al-Tamimi, Schiel, et al., 2022). For this, all participants in this effort had to follow the pre-ATR convention in transcribing their data and after a first time alignment of the data, manual verifications of the segmentations were done, which were then shared back with me for subsequent processing. For this part, I have also had the chance to work with five PhD students and a Research Associate on an hourly paid contract for data transcription, with three others for verifying the semi-automated segmentation and with an intern for three months trying to develop an automatic transcription system for Arabic, based of the ATR transliteration system (Section 4.7).

For my British Academy/Leverhulme Small Grants project investigating the role of the epilarynx in the production of guttural consonants in Arabic, I had the chance to work with a Research Associate for 5 months on automatic tracing and manual verifications of over 13k tongue contours obtained via UTI (Section 4.5.3.2). In addition, one PhD student is currently using the same dataset to investigate the phonetic and phonological patterning in guttural consonants, using a complementary approach with specific landmarks, VC to CV coarticulatory patterns and using classification techniques.

On the role of Automatic methods in supporting diagnosis of pathological voices, I have had the chance to work with two PhD students in Barcelona, advising and guiding them on using automatic methods via Prosogram to segment and obtain prosodic and voice quality profiles, in addition to supporting a PhD student at the Université Paris Cité working on a comparative study of schizophrenia in French and Italian, who is exploring the role of automatic methods via Prosogram. In addition to this, I have been working with an Master student doing their internship in Computational Linguistics on the role of automatic methods in differentiating between AD and non-AD and comparing the performance of Prosogram vs openSMILE (ComParE16 and e-GeMAPS).

6.3 Student supervision

Finally, I have had the chance to work with various students from 2002 and officially since 2007 (see Curriculum Vitae Chapter 1). Starting with BA/BSc students, and between 2002 and 2007, I informally advised five BA students in the Speech and Language Therapy department at the Université Lyon 1 on the analysis of pathological voices in comparison to neurotypical healthy controls on using objective acoustic measures in assessing the degree of dysphonia and breathiness due to various pathologies. This was then followed in 2007-2008, by a sole supervision of a BA thesis of two students in the same department on their final year project by looking at the impact of context and the index of breathiness on the productions of pathological voices. Following on from that, I worked with three other BSc students at the Speech and Language Therapy department at Newcastle University on their final year projects, on various topics ranging from typical productions to comparing neurotypical healthy controls to schizophrenia.

Since joining Newcastle University, and starting in 2009, I have been heavily involved with supervision of students working towards their PhDs, either as full PhD or as part of the IPhD (Integrated-PhD composed of a 1 year taught modules followed by 3 years research phase) in Phonetics and Phonology at Newcastle University. Starting with the IPhD, I have been heavily involved with four students completing their summer project to validate their taught year; two of whom I continued supervising as their primary PhD supervisor (see below). Since joining the Université Paris Cité, I advised two Master students in Phonetics and Phonology, who continue working with me on their PhD topics (see below).

In terms of students working towards their research PhD and/or the research phase of their IPhD, I have had the chance to work with 16 PhD students (6 as first/sole supervisor and 10 as second[1]). In total, 1£ have already completed and defended their PhDs and are all either lecturers and/or researchers in various places (1 in Iraq, 4 in Saudi Arabia, 2 in Thailand, 1 in China, 3 in the UK, 1 in Kuwait). Since starting at Université Paris Cite, I have a formal arrangement with Newcastle University to continue supervision of three PhD students, with one who had his pass list in July 2023 (becoming lecturer in Saudi Arabia), one who already submitted his PhD and having his viva in October 2023 and one who will be submitting her manuscript by November 2023. At Université Paris Cité, I am currently the sole supervisor of one PhD and another starting in October 2023 as second supervisor.

All 16 PhD students worked on topics directly related to my primary or secondary research activities: Four on vowels, specifically dynamics in production and/or perception (McCarthy, Warburton, Almurashi, and C. Li); two on the gutturals as a natural class and/or pharyngeals in Arabic (Alsharif and AlSiraih); two on the Voicing contrast in Arabic dialects (Al-Gamdi and Dallak); six working on L2 phonetics and phonology (Y. Li, Kitikanan, Chen, Ehbara, Kamphikul, and Alghabban), with four specific to the role of training method (Y. Li, Chen, Ehbara and Alghabban); five working on prosodic and rhythmic changes (Chen, Moussa, Ghadanfari, Kamphikul and Zhou).

I always encouraged the students working with me to publish their work. This is evidenced in the description of their work within the summary of my previous research (Chapter 4, but also within my publication list (Chapter 2). In total, there were 37 publications (12 articles, 8 conference proceedings and 8 presentations, and 1 book chapter), which I co-authored with the PhD (and MA) students working under my supervision (see Table 6.1). Some of these publications were based on their MA projects (2 journal articles, JASA and ARECLS; 3 conference proceedings, ICPhS; 5 conference presentation: 1 Journée de Phonétique Clinique, 3 BAAP, 1 Second Language Acquisition of Phonology conference), and the remaining on their PhD projects.

Table 6.1: Number of outputs with PhD and MA students

Articles	Published	5
	Under Revisions	1
	Under Review	3
	In preparation	4
Book chapter	Published	1
Conferences	5-pages articles	8
	Presentations	16
Total		38

I have also been part of the PhD panels for 11 PhD students at Newcastle University and in France. I was also part of 10 PhD examination committees: four as an external examiner with pre-viva report, examination and post-viva report (1 in the UK, 1 in India, 1 in Pakistan and 1 in France); two as an external examiner with examination and post-viva report (2 in France) and four as an internal examiner at Newcastle University with pre-viva report, examination and post-viva report.

[1]It is to be noted that at Newcastle University, all PhD students are required to have at least two PhD supervisors.

Chapter 7

Research Projects

I present here a snapshot of my future research plans, which will build on the various research activities summarised in Chapter 4. As highlighted in various places within that Chapter, and on the list of publications (see Chapter 2), a few publications are in preparation and will be submitted very soon. The descriptions below provide directions for future research I will be engaging in.

As highlighted previously, my research looks for an empirical evidence of the role of FPD in defining various categories. It is inspired by research conducted by other researchers and follows the approach described in Hawkins (2003, 2010). This is achieved by employing variable techniques, including: signal processing algorithms; non-invasive articulatory methods (UTI, EGG, EMA, etc.); perception experiment; informed by the "Signal Detection Theory" (Green & Swets, 1966; Huang & Ferreira, 2020; Macmillan & Creelman, 2005); computational approaches to speech and language, and relying on advanced statistical approaches (e.g., explainable machine leaning, mixed effects modelling).

My future research will continue exploring phonetic and phonological variability from a cross-linguistic perspective, and will be looking at variable participants' population, spanning adults, children, learners of an L2/L3, and with pathological voices. My future research will look at both segmental and suprasegmental aspects going from phonemes, to syllables to words, to utterances, to larger chunks of speech. It is clear that there is a connection between segmental and suprasegmental aspects of speech as they mutually influence each other. Hence, exploring the correlates of suprasegmentals will be important.

A brief summary of these future research activities is provided below.

7.1 Dynamic Specification of Vowels

When examining the role of dynamic specification of vowels presented in Section 4.2 (pp. 49), my research showed that using both intrinsic (the vowels' own trajectories) and extrinsic (coarticulation from surrounding consonants) approaches yielded a better description and characterisation of the produced vowels and allowed listeners to better identify the prototypical vowels. These results were evidenced mainly in Arabic, but also in French and English. In fact, extrinsic dynamic specification of vowels allowed to identify the syllable as the minimal unit of speech production and perception, while the intrinsic dynamic approach showed that [+High] vowels identified with the feature [−Tense] showed a different slope and direction to the same [+High] vowels with the feature [+Tense]. This result is expected for English vowels (Almurashi, 2022; Almurashi et al., Under Review; Slifka, 2003), but the results highlighted in (Almurashi, 2022; Almurashi et al., Under Revisions) showed a similar pattern for Arabic short vowels. This result was highlighted previously in J. Al-Tamimi (2007a, 2007b), which put emphasis on both qualitative and quantitative differences between the short and long vowels in Arabic.

Given that spectral information identified in this body of work highlight a potential for a universal pattern seen across languages, especially those showing a vocalic length contrast (though this can also apply for languages with a consonantal length contrast, see below). The changes observed for the short vowels series is similar to what one can find in English and in Quebec French, which seems to have developed a vowel contrast based on quantity and quality, akin that seen in English (see for instance Martin, 2002). The changes observed within the [±Tense] vowels seem to be related to differences in the degree of advancement/retraction of the root of the tongue, which is similar to that observed for the cases of the [±ATR] phonological contrasts: the [+ATR]/[+Tense] vowels seem to show an advancement of the root of the tongue, whereas those with a [−ATR]/[−Tense], show a possible retraction of the root of the tongue[1]. We ask the question whether these differences are systematic and can be found regardless of the language? or whether there are language-specific patterns that allow to differentiate languages based on the type of ambient phonological contrasts?

From a methodological point of view, the two perception experiments used in J. Al-Tamimi (2007a) were a prototype identification task and a categorisation task. We claim that the two showed different patterns simply due to what they led listeners to identify. The prototype task allowed for coarticulatory patterns to be taken into account by listeners as these dynamic prototypes contained rich and detailed information: this played a major role in facilitating a *phonetic* identification of the prototypes. The results of the experiment without transitions showed a tendency for the prototypes to not to be heavily impacted by coarticulatory patterns. Then, the categorisation experiment showed an almost categorical decision by the listeners to choose their categories; hence this can be seen as a *phonological* task. Of course, even the categorisation task with dynamic stimuli showed preference towards different categories. Exploring the relationship between *phonetic* and *phonological* modes in speech perception is essential as this will lend weight to the claims that speech perception favours gradience and not categoricity, and will depend on the task (see Gerrits, 2001; Gerrits & Schouten, 2004, for an excellent review on the topic, see McMurray, 2022). Exploring advanced methods, such as EEG will allow to identify whether these fine phonetic correlates are important for the listener. In fact,

[1]Although, Vaux (1999) and Gick, Wilson, Koch, and Cook (2004) explain that [−ATR] vowels are closer to the neutral position of the tongue root to differentiate them from [+RTR] vowels.

our claim is that these FPDs are important to allow listeners to identify their *phonological* categories, and hence we expect to identify an increasing ERP activity around the auditory N100, the phonological mismatch negativity (PMN), located between 150-250 ms after onset of the stimulus, in addition to a semantic N400. These changes will signal auditory detection, phonological changes and crucially, a change related to detecting lexico-semantic preferences (see Heidlmayr, Ferragne, & Isel, 2021, in the context of L2 acquisition). These dynamic cues will be predicted to show strong preference for identifying phonological categories in either Arabic, English, languages with a [±ATR] contrast, as opposed to metropolitan French, which does not show a length/quality contrast[2]. Examining these impacts on L2 acquisition is important to further evaluate any possible L1-L2 interferences and assimilations in an L2 context, akin to the work we have done on the dynamics of vowel production of Arabic learners of English as an L2 (Almurashi et al., 2023, In preparation). Exploring the perceptual relevance of these patterns in an L2 context will be important to showcase potential categorical differences between the learner's variety and that of their own L1 or the L1 variety they are targetting.

7.2 Gemination and Voicing - a cross linguistic perspective

The results of my work on gemination (see Section 4.3, on pp. 68) highlighted a major role for the secondary correlates in addition to the length contrast in signalling that geminate consonants in Arabic can be described in terms of longer productions yielding secondary correlates, which correlate with articulatory strength. This double evaluation of geminates is not new as there has been other research on other languages, which highlighted the role of both temporal and non-temporal correlates. The originality of our work lies in the fact that we systematically evaluated whether the gemination contrast is a length contrast with subsequent secondary correlates related to articulatory strength, or whether it is an articulatory strength contrast with subsequent length differences. The results of the classification algorithms allowed us to confirm the primacy of the length contrast over those related to the articulatory strength contrast. The second originality of our work was in evaluating how variable and systematic children were in acquiring gemination. Our results highlighted in Section 4.3.4 (on pp. 75) showed that children started off with a tertiary contrast: short consonants, short-strong consonants and long consonants. The first is correlated with singletons in adults; the last with geminates. The middle category is correlated with what we termed as an articulatory strength contrast, but without systematic lengthening. Here, children acquiring gemination in Arabic were able to identify the importance of the feature [+Tense] in Arabic and were systematically producing it instead of simple, complex and/or geminate consonants. At the end of the 25-words period, children displayed more of a categorical distinction, with productions in the range of the singletons and geminates in adults speech (identified via ADS).

The results here are similar to those seen for the short-long vowels described previously. We ask the question whether Arabic is specific in how it uses the feature [+Tense] for the length (both vocalic and consonantal) contrast? Are the acoustic correlates for [+Tense] similar across both consonants and vowels? Do we expect [+Tense] consonants (i.e., geminates) to show an Advanced Tongue Root, akin to that of the [+Tense] vowels? An area to explore further from an articulatory and acoustic point of view. We also, explored the relationship between gemination and voicing in Arabic (Section 4.4) allowing us to show that voiced geminates behave differently

[2]Although it is not clear if even in French, one can observe systematic tongue root changes for the so-called "median" vowels, i.e., /e ø o/ vs /ɛ œ ɔ/.

from voiced singletons, and that when evaluating the voicing contrast in this category, one can see that Arabic seems to borrow from both Voicing and Aspirating languages. Hence, here again, the feature [+Tense] was used as it better described the patterns observed in our data. This was also confirmed when looking at the Voicing contrast in singleton consonants in another Arabic dialect, where the feature [+Tense] explained the observed patterns (Section 4.4.3.2, but see Section 4.4.3.1 for differences in the implementation of the Voicing contrast in another Arabic dialect).

What is interesting in exploring gemination on its own or in interaction with the Voicing contrast, is that Arabic uses both temporal and non-temporal correlates borrowed from both Voicing and Aspirating languages. We suspect this is the case in other languages showing a consonantal (and possibly a vocalic) length contrast. It would be interesting to explore this further by looking at both acoustic and articulatory bases (explored via UTI, EPG, EMA, EGG, etc.) of the consonantal length contrast in a cross-linguistic perspective, by exploring the potential important role for non-temporal acoustic correlates and changes within the tongue tip, body, root, and laryngeal changes during the production of geminate consonants. Exploring the potential similarities and differences between languages will be important to inform theories of linguistic diversities and universals. We predict to find systematic differences between singletons and geminates by implementing articulatory and acoustic changes, which will be similar across languages, but with different scales. Here, gradiency in feature specification will become important to explain how languages exploit the feature [+Tense], accounting for language-specific patterns. For instance, previous work on Berber (Ouakrim, 2003) identified the length contrast to be primarily based on articulatory strength. If this is confirmed in our subsequent studies, then one can assign this language an [8Tense] (or [9Tense]). The results we obtained for Lebanese Arabic seem to point to a [4Tense] (or [3Tense]). This allows to explain that 1) the feature [Tense] is active in this specific language (in addition to the feature [Long]), 2) that it is variable in how it is implemented due to gradation in scaling and 3) that it has a different implementation across languages. We do not claim that [+Tense] is identical across Berber and Lebanese Arabic; the gradation in feature specification allows for language-specific patterns to emerge.

Looking at the interaction between the Gemination contrast and the Voicing contrast will explain how Geminating languages differ from each other. For instance, our own results presented in J. Al-Tamimi and Khattab (2018b) showed voiced geminates to display variable voicing and in some instances, specific female speakers tended to devoice more than others. This seems to be similar to the case of Japanese presented in Hussain and Shinohara (2019). Hence we ask the question whether devoicing in geminates is a universal tendency or language-specific, which can be better accounted for when using gradient privative feature. Extensions to this work looking at the Voicing contrast on its own will be important to identify how Voicing is implemented in singletons, which then can serve to explain how gemination impacts on Voicing.

Of course, the majority of work I have engaged in with respect to gemination and voicing was based on production and from an acoustic point of view. We already saw above how articulatory techniques can support evaluating the minute changes associated with the articulation of gemination and/or voicing. In addition, the next step will be exploring the perceptual relevance of the feature [+Tense] in perception. Running an identification or categorisation task by preserving non-temporal information and equalising the durations of the two categories (i.e., increasing the duration of singletons and reducing that of geminates), will lead to the acoustic correlates of the [+Tense] to be preserved and hence this will be a direct way to compare its primacy over the temporal domain. If languages do indeed use articulatory strength as a secondary mechanism (or primary depending on the language), then this simple test will allow to identify how variable and systematic listeners are. Exploring more complex designs via an EEG or Eye-tracker can also be used to explore the relationship between manipulated signals and their influence on identifying words; hence leading to an increase in N400 due to lexico-semantic incongruency, in addition to an increase in reading time from an Eye-Tracker perspective in identifying read and heard stimuli (for a similar approach, see Steffman, 2023).

Given that the feature [+Tense] is active in Arabic gemination, we ask the question on what happens in the cases of degemination (loss of the gemination contrast) described in some Arabic dialects (K. S. Jones, 2016). It is possible that both degemination and depharyngealisation (see below) play a major role from a sociolinguistic viewpoint, distinguishing between Urban and Bedouin dialects. Hence, we postulate that the feature [+Tense] will be important when one loses the length contrast. Is it likely that those who degeminate, produce "shorter" productions, which are stronger? Are these cases more emphasised to "mark" the phoneme or the word in which they occur so that any confusions are removed?

Looking at the role of secondary correlates to the length and voicing contrasts is crucial to identify language-specific patterns to evaluate cases of partial and complete neutralisation of voicing or length contrasts. Various points which will be explored in planned collaborative activities.

7.3 Guttural consonants in Arabic - A cross dialectal perspective

A major research activity that I have engaged in during the last few years was related to understanding the role of the epilarynx in the articulation of guttural consonants in Arabic (Section 4.5). My work showed that pharyngealised consonants are produced via a combined supra-laryngeal and laryngeal gestures (J. Al-Tamimi, 2017b). This same combined gesture is identified as the major articulatory and acoustic gesture for all three guttural contexts: uvular, pharyngealised and pharyngeal (J. Al-Tamimi & Ferragne, 2020b; J. Al-Tamimi & Palo, 2023).

The results pointed to a partial epilaryngeal constriction inducing both laryngeal and supra-laryngeal changes. In terms of the former, it is clear that the larynx is constricted and in some cases (e.g., uvular and pharyngeal), it is also raised. Raising of the larynx was also seen in the pharyngealised context, when comparing the EGG signals throughout the VCV sequence. Indeed, while Larynx raising was evident for the pharyngeal context throughout the VCV sequence and especially within the consonants, it was only evidenced for the pharyngealised context at the consonant release. Investigating this in a less controlled environment is crucial to generalise the results to other speech tasks. Acoustic results again pointed to a constricted glottis as a secondary feature in the production of gutturals; correlating with the feature [CG].

In terms of the supra-laryngeal changes, indeed, bark-distance measures outperformed absolute measures and they can be seen as a direct psychoacoustic measure to correlate production to perception. Exploring additional correlates of the noise components and of formant changes are essential to quantify the differences between gutturals and non-guttural consonants. In terms of the UTI results, these pointed to systematic articulatory changes towards the tongue front, body, dorsum and root that seem to combine all gutturals as a natural class. Exploring this articulatory similarity is important to understand how the whole tongue changes can explain the similarities between this natural class. With the UTI data, it was impossible to quantify tongue tip changes with precision. While it was evidenced that there was a small systematic difference between pharyngealised and plain contexts towards the tongue tip, this was not precise. Alternative methods, such as EMA or EPG are required to assess whether tongue tip changes are important to signal a potential dental articulation for the pharyngealised set[3]. What is the relationship between tongue tip, body, root and larynx in the production of all gutturals? How can we quantify the gestural coordination between all these measures using non-invasive techniques? I have explored using UTI and EGG together: can we expand this to EMA, UTI and EGG? What are the challenges from research point of view, and from an ecologic point of view? Combining multiple tools can be informative as to how each can predict specific changes, however, the interpretation of the findings may be hindered by the fact that speakers will be more and more constrained in their productions. Can MRI serve as a potential tool, even if any bony structure are invisible and imaging in supine position is different from standing up (A. Wrench, Cleland, & Scobbie, 2011)?

Crucially, we ask the question to know how robust these findings are from a perception point of view? If as we claim, that the feature [CG] is active in Arabic, what influence it has in the case of depharyngealisation or in acquisition? As with geminates above, manipulating the signal to only include the acoustic correlates of [CG] on samples of plain contexts will allow to evaluate how robust the findings are and inform phonological theory.

My own results showed that the epilarynx is partially constricted in Jordanian and in Levantine Arabic (grouping Jordanian Arabic, to Syrian, Lebanese and Palestinian Arabic dialects), in comparison to Moroccan Arabic, which showed a lowered type of constriction for the pharyngealised set. We ask the question then on how variable the dialects are in implementing the primary and secondary correlates for pharyngealisation and for gutturals in general? Can the degree of epilaryngeal constriction be used as an index to quantify dialectal proximity? Are North African Arabic dialects more prone to extreme epilaryngeal constrictions in comparison to Eastern Arabic dialects? As pointed out in Section 4.5.3.6.1, there is a tendency for some dialects to sound more guttural than others (Bellem, 2007), with no clear description of what does that mean. We suspect that this epilaryngeal constriction in gutturals can be used as both a geolinguistic and a sociolinguistic index of dialects, in a similar fashion as how /q/, /θ ð ðˤ/, /ʒ/, /x ɣ/ or vowels have been used as reflexes to distinguish Eastern from Western Arabic dialects (Embarki, 2008, 2013).

Finally, the acoustic results showed similarities in the laryngeal changes observed here and those in African languages with the [±ATR] contrast and with the singing voice. For the former, we ask then the question to know how similar these patterns will be when one looks at a cross language comparison between Arabic dialects and other languages with a [±ATR] contrast? We know that laryngeal changes are usually identified as secondary correlates to the [±ATR] contrast, which seems to be the case here for Arabic. Expanding on this further will be important to identify the role of the laryngeal activity in Arabic. As before, in the cases of depharyngealised sounds that lost their pharyngeal component, do speakers retain the laryngeal changes to remove

[3]Special thanks to John Esling for point out this to me while listening to some productions.

any potential confusions with plain pairs, or do they retain a different place of articulation for the depharyngealised sounds? How would the two following words be distinguished? The word /seːf/ (sword) and /sˤeːf/ (summer) show a perfect minimal pair. If the latter word loses its pharyngeal component and becomes depharyngealised, how will it sound? Would it be transcribed as: [seˀːf], [s̤e̤ːf], [s̰ḛːf] or [s̺e̞ːf]? How perceptible these forms are and how would they allow to distinguish participants who depharyngealise? Clearly, if the depharyngealised version of /sˤeːf/ became [seːf], then there is clear confusion between the two words, where the context becomes important to avoid confusions.

In terms of the signing voice, and when looking at the predictions advanced by Story (2019) in addition to the literature review on the topic presented in J. Al-Tamimi (2017b), it is clear that constricting the epilarynx maximally is used as a technique to obtain an "enhanced" voice by increasing the energy in the high frequencies, leading to a narrowing around F3, F4, and F5 to signal the singer formant Story (2019). We used these predictions in J. Al-Tamimi (2017b) by exploring the two metrics A^*1-A^*3 and A^*2-A^*3. Our results showed a reverse pattern whereby a decreased value of the difference was obtained in the pharyngealised context, due to a decrease in energy around F3, likely due to an abrupt closure of the glottis (Hanson et al., 2001) and the fact that we identified this epilaryngeal constriction to be partial. Clearly, the singing voice is different from normal speech, however, when singing or speaking, we use the same components in our vocal tract. We use the larynx to vibrate and produce voicing, we use the supralaryngeal cavities to shape our vocal tract for an optimal production and we use the lips in a specific way to modulate the productions. Hence, we ask the question to know how different is our vocal tract in singing and in speech? Do we use the exact same components, but to deferring degrees of constriction during the two acts? Are we expecting the use of the epilarynx to be universal across languages when singing or speaking? In the case of an extreme epilaryngeal constriction in speech, are we expecting to observe a similar increase in energy in higher frequencies as that observed for the signing voice? All interesting question that require additional experimentations.

7.4 Speaker-specific variation and dialect identification

One of the main findings from my research is that when using FPD to describe categories, we take into account speaker-specific details. Indeed, FPD refer to the detailed acoustic (and articulatory) correlates used by a speaker to produce speech sounds. These specific details are inherent to the speaker. The approach taken in my research is that by accounting for speaker-specific detail in our analyses, we try to show how systematic speakers are and how close they are to each other in the characterisation of a specific phonological category. By doing so, we provide generalisation as to how a specific contrast is produced, e.g., gemination, voicing, pharyngealisation, etc. However, we know well that speakers are variable. There are speaker-dependent variation that needs to be accounted for in order to be able to describe how the speaker-hearer are interacting with each other. Following Foulkes and Docherty (2006) and more recently Kendall, Pharao, Stuart-Smith, and Vaughn (2023), speakers actively produce specific features to signal their identity, sociolinguistic and geolinguistic background. Exploring the sociophonetic features in the productions of participants from various origins (Urban vs Bedouin); physiological sex; sociolinguistic gender; age, socioeconomic background, etc. is important in the context of Arabic and French research as there is a real need to explore the sociophonetic variation in the speech of participants to understand how they shape their productions in such a way to signal their identify. In some cases, and depending on the context, some of these topics can be seen unattainable simply because we cannot ask specific questions to participants related to their origin and/or sociolinguistic gender and/or economic status. However, a first step towards enriching this area is by exploring speaker-specific variation and use predictions from the literature to inform the theory.

I see speaker-specific variation as an important aspect to signal one's identity. We speak to communicate but also to signal who we are. This is clearly evidenced in nearly all of my research, either on L1 or L2, on neurotypical controls vs pathological voices, on male vs female differences, on dialectal differences between Jordanian Arabic or Moroccan Arabic, etc. In my subsequent research, speaker-specific variations will be explored further to highlight individual strategies in producing speech. Even in perception, we know very well that listeners do vary between each other, and as highlighted in the supplementary material of Khattab et al. (2018, for Voice Quality, see the Figures at: https://jalalal-tamimi.github.io/R-Rating-data/Rating-VQ-Nas-Phonetica.nb.html#32412_figures; for Nasalisation, see the Figures at: https://jalalal-tamimi.github.io/R-Rating-data/Rating-VQ-Nas-Phonetica.nb.html#32422_figures), raters identified some producers as being more or less nasal or with a tense voice quality than others. The same is true for the items and for the raters themselves. Exploring random effects in a mixed effects regression analysis in a similar fashion to that in Drager and Hay (2012) or in Baumann and Winter (2018) will be an important step forward to evaluate how speakers-hearers interact with each other.

Of course, some of my research attempted to explore how individual variations can be used to identify dialectal differences. This will be explored further in a similar fashion to that presented in Section 4.5.3.6.1. When looking at how specific acoustic correlates in specific contexts can be used to inform of dialectal differences, this was used to inform subsequent research that for instance, Moroccan Arabic is different from Jordanian Arabic, not only in terms of the number of vowels, the size of the vowel space, the jerkier rhythmic structure, but also in terms of the type of epilaryngeal constriction used (being closer to a maximally constricted), in comparison to that of Jordanian Arabic speakers. This correlates well with the few descriptions differentiating the two dialects on other features than the ones usually explored. It would be important to explore this in a more systematic way, ideally, on pre-existing large datasets, spanning multiple Arabic dialects, speakers, genders, and from various geographic areas. This will allow for an in-depth exploration of the acoustico-perceptual realisations of the pharyngealised and gutturals, which will be used first to describe how the dialects can be differentiated from each other in terms of how these sounds are produced, but also, how speakers vary with respect to the various sociolinguistic and geolinguistic features they use to signal their identity.

The combination of fine acoustic measurements as explored in my own research, but also exploring more advanced signal processing algorithms, e.g., openSMILE or wav2vec, in addition to ASR algorithms will guide research to explore speaker-specific details in dialect specification. Exploring random effects structures and using machine learning will guide research to further evaluate proximity and differences between various dialects from the same geographical area. From a perception point of view, it would be important to understand which correlates listeners pay attention to when identifying the dialectal origin of a speaker. Mimicking the methodology of Zellou (2017a, 2017b), we could evaluate the degree of gutturalness in the production of speakers from variable regions and explore the specificity of the laryngeal and supralaryngeal correlates used. These are a few ideas to expand on to allow enriching the analyses and descriptions of gutturals in Arabic dialects. Of course, these analyses will complement the formal accounts employed so far. As was already proposed for geminates and the feature [±Tense], here, we plan to explore the gradiency in the specification of the feature [±cet] and [CG]. We claim that all Arabic dialects have the two features active, but will vary with respect to the degree of constriction. Western and Bedouin dialects are likely to show an [8cet] and a [8CG] for the pharyngeal and pharyngealised, whereas Easter and Urban dialects will show a [7cet] and a [5CG] for the pharyngeal and a [5cet] and a [7CG]. Of course, these are predictions and will require additional examination of the data to have a clearer pattern.

7.5 L2 Variety

The work I engaged in within L1 and L2 will continue. I am planning a new direction of research on using bio-feedback as a tool to enhance learning. These can use UTI as a bio-feedback to learning sound categories, using multiple sources of variation (whether male vs female voices, young vs old, multiple accents, etc..) will allow the learner to enrich their inventory of the sounds they are learning as it will strengthen the exemplars they are constructing of a particular sound category. Combining knowledge of the L1 (whether target language or mother tongue) into learning an L2 is crucial to allow the examination of the learners varieties. For instance, we know well that learners will always deviate from their mother tongue to achieve a target-like and a native-like production in their targetted L2. How close are the productions of learners from their targetted L2? To be able to answer this, we of course need to rely on empirical examination of the targetted L2 as produced by native speakers to have a model to compare to. We can then assess where the learner's productions are with respect to the two extremes: the learner's L1 and the targetted language's L1 productions. Once this is quantified, we can then evaluate the closeness of the learner's productions to one or the other extreme. We can then be able to evaluate the role of extralinguistic factors on the learning strategies. As highlighted previously, most of the research I was engaged in looked at English as an L2. Of course, expanding these studies to cases of learners of French as a foreign/second language will strengthen the cross-linguistic comparison of the observed patterns. For instance, spectral dynamic cues are claimed to be universal and not language-specific and it would be interesting to observe how learners of French produce these dynamic cues and whether they are important to signal native-likeness and target-likeness.

In terms of acquisition of the L1, there is a scarcity of research putting emphasis on the role of FPD of phonological contrasts. I'd like to expand on the work I have done on Arabic, by exploring how children of languages with the consonantal length contrast acquire it. This study will be combined with that on adults cross-linguistic investigation to identify the role of the feature [+Tense]. If this feature is active in the adult variety, what are its phonetic correlates and how are these acquired and produced by children? In addition, I'd like to be engaged in research on developing a large scale database of Arabic acquisition of various contrasts. I heavily looked at the acquisition of gemination; what happens for pharyngealisation (or gutturals in general)? Are children expected to display a gradient change in their acquisition journey, where they will show gradient use of the feature [+cet] and/or [CG]? Are we expecting them to display specific voice quality changes similar to those observed for adults? Or are they going to deviate from these, especially depending on their dialectal origin? All questions that require more fine grained analysis of their speech outputs to appropriately correlate their productions to the expected categories in the adult speech.

7.6 WebMAUS and romanisation of Arabic scripts

The work on the development of the Arabic WebMAUS service continues. I have access to additional data from around 100 Arabic speakers from various regions (Morocco, Algeria, Egypt, Jordan, Saudi Arabia, Lebanon, Iraqi), and the work to transcribe their data has started. Given that most of the automatic transcription systems for Arabic usign Arabic scripts and do not add diacritics for short vowels, this causes issues for any algorithms to accurately segment the speech chunks. I am planning two exploratory directions of research within this area. The first is planned with the support of internships funded by the BRIO AAP, to identify solutions for automated transcription of the short vowels. This will be inspired by the Arabic WebMINNI Speech-to-Text. The second will be exploring Large Language Models and wav2vec, which will be used to provide a phonetically informed transcription, akin that done for low resourced languages. Once this hurdle is surpassed, the Arabic WebMAUS service will be enriched with this additional data, with the aim to strengthen its performance (version 2 contains over 500k sounds; version 3 will have over 1 Million sounds!). A formal comparison between manual segmentation and that obtained from WebMAUS services is planned to assess its performance. The next steps and depending on content of the material is to provide two additional tools:

- A region-specific Arabic aligner for Easter vs Western Arabic
- A possible extension to obtain WebMAUS aligners for each of the 5 geo-linguistic areas (Maghreb, Egyptian, Arabian Peninsula, Mesopotamia and Syro-Lebanese)

Further developments are planned in e.g., development of a Pronunciation Model that identifies a Target word with its variants. This can be used further to obtain frequencies of usage (phoneme or word-based), enriching the aligner with variants. To achieve all of these tasks, additional data will be required; likely by exploring the Mozilla Common Voice dataset.

7.7 Speech as a biomarker? Role of automated methods

The current research I have engaged in on using automatic methods of signal processing in supporting diagnosis of cognitive disorders will continue. Currently, we only explored one set of automated features that were based on a fully automated pipeline not relying on any prior transcriptions. What would happen if one explored full ASR algorithms to assessing the impact of the diseases on speech outputs? Of course, the expectations are that we will gain more in detection accuracy, albeit with less interpretability. From a clinical context, it is important to understand what the results are telling us and how they can inform diagnosis. For instance, if the increase in pausing is accompanied with a decrease in amount of spoken output and with an increase in breathiness, then these can be correlated to slowness in articulator movements likely due ageing but also to the cognitive disorder. Continuing with other speech processing algorithms, e.g., wav2vec, will allow us to explore the use of advanced methodologies to identify the role of speech as a potential biomarker. However, this will be least interpretable as what matters will mostly be performance. Of course, I'll keep working on interpretable algorithms to allow for an easy interpretation of the findings.

The current findings from my research have shown similarities in performances of automated speech methods in schizophrenia and dementia, even if the causes of these cognitive disorders are different. Can we think then of speech as a true biomarker carrying specific features to signal our state? This is a topic I'd like to explore further in both clinical and non-clinical contexts (see above).

From the clinical context, we know that the various neurodegenerative pathologies seem to impact on speech in a similar fashion. Parkinson's disease seems to impact on speech in a similar fashion, with an increase in pausing, increase in breathiness, etc. Are there expectations that any cognitive and/or degenerative pathology will have similar impacts on the speech outputs, like due to slowness in articulators and in motor planning? If this were the case, then this will need further exploration with the support of clinicians to identify how automated speech measures can be used as potential biomarkers of pathologies and how they can be used to support diagnosis. Importantly, it is not clear what the impact of normal ageing is in may instances in my research. As seen in Section 4.6.2 (pp. 129), the neurotypical controls were of an old age and we did not have any younger populations to compare to. Given the difficulties in some cases to dissociate normal ageing from pathological ageing (Ivanova et al., 2023), this will be an important field of investigation in my future research activities.

In addition, currently, Large Language Models have shown an excellent performance on unseen data and in cases of low resourced languages. I'd like to explore them further by looking at how they perform on pre-transcribed data obtained from the Dementia Bank. The problem here is that by relying on pre-transcribed data, we are unable to make fair and reasonable comparisons to automated speech measures. It would be important then to explore automatically transcribed speech outputs obtained by transformers and Large Language Models from clinical settings to evaluate how they perform on their own and in combination with automated speech measures.

Ultimately, I'd like to expand the research on a cross-language perspective. Looking at the DementiaBank, English is the most represented language, with other corpora available for other languages. Exploring whether our current findings from English (the ADReSS Challenge) and from Spanish/Catalan (the ACE foundation) can be found for other languages, will be a first step. If these speech measures turn up to be universal, this will be a major contribution to the field. However, we expect that some of the automatic measures will be language-specific and task specific. This is why exploring this further in a cross-language and task setting will be essential to allow identifying of universal and language-specific measures to quantify the impacts of pathologies on speech outputs. Finally, expanding this research on Arabic will be important to document the changes on other languages and to inform future researchers and clinicians on the specific signs associated with these pathologies.

Chapter 8

Bibliography

Abramson, A. S. (1986). The perception of word-initial consonant length: Pattani Malay. *Journal of the International Phonetic Association*, *16*(1), 8–16. doi: 10.1017/S0025100300003054

Abramson, A. S. (1999). Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. In *Proceedings of the 14th ICPhS* (pp. 591–594).

Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, *63*, 75–86. doi: 10.1016/j.wocn.2017.05.002

Al-Ani, S. (1970). *Arabic Phonology: An Acoustical and Physiological Investigation*. The Hague Mouton. doi: 10.1515/9783110878769

Al-Gamdi, N. (2022). *Voicing Contrast in Najdi Arabic Stops: Implications for Laryngeal Realism* (Unpublished doctoral dissertation). Newcastle University, UK.

Al-Gamdi, N., Al-Tamimi, J., & Khattab, G. (2019). The acoustic properties of laryngeal contrast in Najdi Arabic initial stops. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 2051–2055). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Al-Gamdi, N., Al-Tamimi, J., & Khattab, G. (2022a). Voicing contrast in Najdi Arabic stops: Implications for Laryngeal realism. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, UK (4-8 April 2022).

Al-Gamdi, N., Al-Tamimi, J., & Khattab, G. (2022b). Voicing contrast in Najdi Arabic stops: Implications for Laryngeal realism. In *Proceedings of the 18th LabPhon (Laboratory Phonology) Conferene.* Online conference (23-25 June 2022).

Al-Gamdi, N., Al-Tamimi, J., & Khattab, G. (Under Review). Voicing Contrast and Voicing Assimilation in Najdi Arabic. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.

Al-Gamdi, N., Khattab, G., & Al-Tamimi, J. (2018). Voicing contrast in Najdi Arabic initial stops. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).

Al-Tamimi, F., & Heselwood, B. (2011). Nasoendoscopic, videofluoroscopic and acoustic study of plain and emphatic coronals in Jordanian Arabic. In B. Heselwood & Z. Hassan (Eds.), *Instrumental studies in Arabic phonetics* (pp. 165–191). John Benjamins. doi: 10.1075/cilt.319

Al-Tamimi, J. (2001). *Description linguistique de 6 parlers jordaniens : Étude des interférences en production entre l'arabe dialectal jordanien et l'arabe classique ; une approche phonético-phonologique et acoustique* (Master Thesis - Maîtrise in Language Sciences - Phonetics and Phonology). Université Lyon 2, Lyon.

Al-Tamimi, J. (2002). *Variabilité phonétique en production et en perception de la parole : Le cas de l'arabe jordano-palestinien* (Master Thesis - DEA in Language Sciences - Phonetics and Phonology). Université Lyon 2, Lyon.

Al-Tamimi, J. (2004). L'équation du locus comme mesure de la coarticulation VC et CV: Étude préliminaire en Arabe Dialectal Jordanien. In *Actes des 25èmes journées d'Études sur la Parole (JEP)* (pp. 9–12). Fès, Morocco (19-22 April 2004).

Al-Tamimi, J. (2006). Analyse dynamique de la réduction vocalique en contexte CV à partir des pentes formantiques en arabe dialectal et en français. In *Actes des 26èmes journées d'Études sur la Parole (JEP)* (pp. 357–360). Dinard, France (12-16 June 2006).

Al-Tamimi, J. (2007a). *Indices dynamiques et perception des voyelles: Étude translinguistique en arabe dialectal et en français* (PhD Thesis in Language Sciences - Phonetics and Phonology). Université Lyon 2.

Al-Tamimi, J. (2007b). Static and dynamic cues in vowel production: A cross dialectal study in Jordanian and Moroccan Arabic. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)* (pp. 541–544). Saarbrücken, Germany (6-10 August 2007).

Al-Tamimi, J. (2008). L'espace vocalique perceptif dépend de la densité des systèmes vocaliques: Étude translinguistique en arabe marocain, en arabe jordanien et en français. In *Actes des 27èmes journées d'Études sur la Parole (JEP)* (pp. 1–4). Avignon, France (9-13 June 2008).

Al-Tamimi, J. (2009). Effect of pharyngealisation on vowels revisited: Static and Dynamic analyses of vowels in Moroccan and Jordanian Arabic. In *Proceedings of the International Workshop on Pharyngeals & Pharyngealisation.* Newcastle University, UK (26-27 Mars 2009).

Al-Tamimi, J. (2011). La palatalisation en coréen comme exemple de coarticulation : Données ciné-IRM stroboscopiques et acoustiques des mouvements graduels de la langue, de " (translation to French from: ", Hyunsoon KIM, Korean Palatalization as

Coarticulation: Stroboscopic cine-MRI and acoustic data on gradual tongue movements). In M. Embarki & C. Dodane (Eds.), *La coarticulation. Des indices à la représentation* (pp. 253–270). L'Harmattan.

Al-Tamimi, J. (2013). Articulatory Phonetics Edited by Bryan Gick, Ian Wilson Donald Derrick (Malden, MA: Wiley Blackwell, 2013) [Pp. 272] ISBN 978-1-4051-9321-4. £60.00 (hbk), ISBN: 978-1-4051-9320-7. £24.99 (pbk). *International Journal of Language & Communication Disorders*, *48*(5), 597–598. doi: 10.1111/1460-6984.12029

Al-Tamimi, J. (2014). An acoustic investigation of pharyngealisation effects on vowels: A cross-dialectal investigation in Jordanian and Moroccan Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* Oxford University, Oxford, UK (7-9 April 2014).

Al-Tamimi, J. (2015). Spectral tilt as an acoustic correlate to pharyngealisation in Jordanian and Moroccan Arabic (Article: 0436). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015.

Al-Tamimi, J. (2016). Is pharyngealization associated with an epilaryngeal constriction? An acoustic investigation in Jordanian and Moroccan Arabic. In *Proceedings of the Arabic linguistic forum (ALiF).* York University, UK (12-14 December 2016).

Al-Tamimi, J. (2017a). Acoustics of constricted epilarynx: The case of voiced coronal pharyngealized consonants in Jordanian and Moroccan Arabic. In *Proceedings of the 2nd PaPE (Phonetics and Phonology in Europe).* Cologne, Germany (12-14 June 2017).

Al-Tamimi, J. (2017b). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1), 1–40. doi: 10.5334/labphon.19

Al-Tamimi, J. (2018a). A Generalised Additive Modelling approach to ultrasound tongue surface: Quantifying retraction in Levantine Arabic back consonants. In *Proceedings of the 16th LabPhon (laboratory phonology) satellite workshop: New developments in speech sensing and imaging.* Lisbon, Portugal (23 June 2018).

Al-Tamimi, J. (2018b). Gradient epilaryngeal constriction in Levantine Arabic "gutturals": A Generalised Additive Modelling approach to ultrasound tongue surface. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).

Al-Tamimi, J. (2018c). *Gradient epilaryngeal constriction in Levantine Arabic "gutturals": A Generalised Additive Modelling approach to ultrasound tongue surface* [Atelier « Temporalité et Séquentialité Dans Les Formes Sonores », Nantes, France (June 2018)].

Al-Tamimi, J. (2018d). *R-Rating-data* [Open Source Online Material - Supplementary Material as an Introduction to Cumulative Logit Regression]. https://jalalal-tamimi.github.io/R-Rating-data/.

Al-Tamimi, J. (2018e). *R-Voicing-Gemination-VOT* [Open Source Online Material - Supplementary Material for the Statistical Analyses of Mixed-Effects Models and Random Forests]. https://jalalal-tamimi.github.io/R-Voicing-Gemination-VOT/.

Al-Tamimi, J. (2018f). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062018.nb.html.

Al-Tamimi, J. (2019a). *Introduction to R* [Open Source Online Material as an Introduction R: Introduction, Visualisation]. https://jalalal-tamimi.github.io/R-Introduction-to-R/.

Al-Tamimi, J. (2019b). *The role of the epilarynx in male/female and dialect discrimination in dialectal Arabic/Le rôle de l'épilarynx dans la discrimination entre locuteurs hommes/femmes et entre dialectes arabes* [Séminaire de Recherche : Traitement Du Langage Parlé (TLP), Laboratoire d'Informatique Pour La Mécanique et Les Sciences de l'Ingénieur (LIMSI) - Université Paris-Sud, France (May 2019)].

Al-Tamimi, J. (2019c). *The role of the epilarynx in the production of gutturals in Levantine Arabic/Rôle de l'épilarynx dans la production des gutturales en Arabe Levantin* [Séminaire de Recherche En Phonétique et Phonologie, Laboratoire de Phonétique et Phonologie (SRPP-LPP) - Paris 3 (March 2019)].

Al-Tamimi, J. (2019d). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062019.nb.html.

Al-Tamimi, J. (2020a). *The role of the epilarynx in the production of guttural consonants in Levantine Arabic* [Speech Science Forum, Hearing and Phonetic Sciences – University College London, UK (November 2020)].

Al-Tamimi, J. (2020b). *Techniques in data analyses* [Open Source Online Material - Advanced Analyses Usings R: Linear, Logistic, Cumulative and Mixed-Effects Models]. https://jalalal-tamimi.github.io/R-Techniques-in-Data-Analyses/Session_4-AnalysingData062020.nb.html.

Al-Tamimi, J. (2021a). Acoustic correlates of epilaryngeal constriction in Levantine Arabic guttural consonants. In *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).

Al-Tamimi, J. (2021b). *Introduction to Random Forests* [Open Source Online Material - Introduction to Decision Trees and Random Forests]. https://jalalal-tamimi.github.io/Intro-Random-Forests/.

Al-Tamimi, J. (2021c). *Session 4 : Techniques in data analyses* [Workshop Adventure in R: Training in Using R Delivered to PhD Students in Humanities and Sociale Sciences, Newcastle University (June 2021)].

Al-Tamimi, J. (2022a). *On the contribution of automatic methods at quantifying the impact of pathologies on speech* [Symposium Labex EFL (Empirical Foundations of Linguistics)]. Paris, France (16 June 2022).

Al-Tamimi, J. (2022b). *Praat-fo-Accurate-Estimation* [Open Source Online Material - Praat Scripts to Accurately Estimate Fo Curves, Using the Two-Pass Method and Then Smoothing (Original Version 2018)]. https://jalalal-tamimi.github.io/Praat-fo-Accurate-Estimation/.

Al-Tamimi, J. (2022c). *Praat-Measurement-points* [Open Source Online Material - Praat Script to Allow for an Accurate Measurement Point (Original Version 2018)]. https://jalalal-tamimi.github.io/Praat-Measurement-points/.

Al-Tamimi, J. (2022d). *Praat-Silence-Detection* [Open Source Online Material - Praat Script for Silence/Speech Detection Using the

Two-Pass Method for Fo Estimation and Intensity Adapted to Each Speaker]. https://jalalal-tamimi.github.io/Praat-Silence-Detection/.

Al-Tamimi, J. (2022e). *Praat-Spectral-Analyses-fricatives-stops* [Open Source Online Material - Praat Script for Computing Spectral Measurements for Fricatives and Stops]. https://jalalal-tamimi.github.io/Praat-Spectral-Analyses-fricatives-stops/.

Al-Tamimi, J. (2022f). *Praat-Spectral-Tilt* [Open Source Online Material - Praat Script for Computing Spectral Tilt Measurements for Voice Quality]. https://jalalal-tamimi.github.io/Praat-Spectral-Tilt/.

Al-Tamimi, J. (2022g). *Praat-Voicing-detection* [Open Source Online Material - Praat Script for Voicing Detection Using the Two-Pass Method (Original Version 2018)]. https://jalalal-tamimi.github.io/Praat-Voicing-detection/.

Al-Tamimi, J. (2022h). *Praat-VQ-Measurements* [Open Source Online Material - Praat Script for Computing Many Acoustic Measurements for Voice Quality]. https://jalalal-tamimi.github.io/Praat-VQ-Measurements/.

Al-Tamimi, J. (2022i). *R-Estimating Number Of Trees RF* [Open Source Online Material - This R Script Allows to Estimate the Optimal Number of Trees Required to Frow Random Forests via the Conditional Inference Trees Framework (Initial Version in 2018)]. https://jalalal-tamimi.github.io/R-Estimating-Number-Of-Trees-RF/.

Al-Tamimi, J. (2022j). *R-GAMM-LabPhon18* [Open Source Online Material - Supplementary Material as an Introduction to GAMMs (Generalized Additive Mixed-effects Modelling) on UTI (Ultrasound Tongue Imaging) Data. First Release in 2018; Updated in 2022]. https://jalalal-tamimi.github.io/R-GAMM-LabPhon18/.

Al-Tamimi, J. (2022k). *R training* [Open Source Online Material Aimed as a Self-Introduction to Using R: Introduction, Tidyverse, Visualisation, Inferential Statistics (LM, GLM, CLM, LMM), PCA, Decision Trees, Random Forests]. https://jalalal-tamimi.github.io/R-Training/.

Al-Tamimi, J. (2022l). *A whole tongue approach to gutturals in Levantine Arabic using Generalized Additive Mixed Modelling of Tongue surfaces* [Séminaire de Recherche En Phonétique et Phonologie, Laboratoire de Phonétique et Phonologie (SRPP-LPP) - Paris 3 (October 2022)].

Al-Tamimi, J. (2023). Différents stades d'Alzheimer : Rôle des méthodes automatisées de traitement de la prosodie et de la qualité de la voix. In *Actes des 9èmes Journées de Phonétique Clinique (JPC2023): "Prendre la mesure de la Parole".* Toulouse 15-17 juin 2023.

Al-Tamimi, J. (In preparation). Acoustic properties of the laryngeal and supra-laryngeal effects of epilaryngeal constrictions in Levantine Arabic guttural consonants. *The Journal of the Acoustical Society of America*.

Al-Tamimi, J., & Barkat-Defradas, M. (2002). Inter-dialectal and inter-individual variability in production and perception: A preliminary study in Jordanian and Moroccan Arabic. In I. Ferrando & J. J. S. Sanchez (Eds.), *Actes des 5èmes rencontres de l'Association internationale de dialectologie arabe (AIDA), cadiz, espagne* (pp. 171–186).

Al-Tamimi, J., Caillol, C., Ferragne, E., & Pélissier, M. (2022). *Many Speech Analyses: Team linckia_nattereri* [Full Analyses of the Many Speech Analyses Project by the Team Members; Contains Processing Files, Praat Scripts to Extract Acoustic Measures, Exploratory and Confirmatory Statistical Analyses]. https://osf.io/8r6x7/. OSF.

Al-Tamimi, J., & Ferragne, E. (2005). Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. In *Proceedings of the 9th Eurospeech (European Conference on Speech Communication and Technology)* (pp. 2465–2468). Lisbon, Portugal (4-8 September 2005).

Al-Tamimi, J., & Ferragne, E. (2020a). Deep Learning and Random Forests show acoustic-phonetic evidence supporting the guttural natural class in Levantine Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).

Al-Tamimi, J., & Ferragne, E. (2020b). The phonetic basis of the guttural natural class in Levantine Arabic: Evidence from coarticulation and energy components using Deep Learning and Random Forests. In *Proceedings of the 17th LabPhon (Laboratory Phonology) Conference.* Vancouver, Canada (5-9 July 2020).

Al-Tamimi, J., Girard, M., & Marsico, E. (2002). Variabilité inter-langue et inter-individuelle en production et en perception: Étude préliminaire en arabe dialectal et en français. In *Actes des 24èmes Journées d'Études sur la Parole (JEP)* (pp. 169–172). Nancy, France (June 2002).

Al-Tamimi, J., Heyne, M., & Derrick, D. (2019). *GAMM-Trombone-2019* [Open Source Online Material - Supplementary Material as an Introduction to GAMMs (Generalized Additive Mixed-effects Modelling) on UTI (Ultrasound Tongue Imaging) Data on Trombone vs Vowel Production]. https://jalalal-tamimi.github.io/GAMM-Trombone-2019/.

Al-Tamimi, J., Heyne, M., & Derrick, D. (2020a). Accounting for within and between-subject variation using Generalized Additive Mixed Models: Application on Ultrasound Tongue contours with and without normalization. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).

Al-Tamimi, J., Heyne, M., & Derrick, D. (2020b). From SS-ANOVA to GAMMs: Accounting for within and between-subject variation using generalized additive mixed models on ultrasound tongue contours. In *Proceedings of the 12th International Seminar on Speech Production (ISSP).* Online Virtual Conference (14-18 December 2020).

Al-Tamimi, J., Jones, W., & Hinzen, W. (In preparation). Automatic acoustic profiling of prosody in schizophrenia. *The Journal of the Acoustical Society of America, Express Letters*.

Al-Tamimi, J., & Khattab, G. (2010). Multiple cues for the singleton-geminate contrast in Lebanese Arabic: The role of non-temporal characteristics. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* London, UK (29-31 March 2010).

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

Al-Tamimi, J., & Khattab, G. (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. *The Journal of the Acoustical Society of America*, *138*(1), 344–360. doi: 10.1121/1.4922514

Al-Tamimi, J., & Khattab, G. (2018a). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate plosives. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).

Al-Tamimi, J., & Khattab, G. (2018b). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics, Invited manuscript for the special issue of Journal of Phonetics, "Marking 50 Years of Research on Voice Onset Time and the Voicing Contrast in the World's Languages" (eds., T. Cho, G. Docherty & D. Whalen)*, *71*, 306–325. doi: 10.1016/j.wocn.2018.09.010

Al-Tamimi, J., Lofgren, M., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (In preparation). Automated prosodic and voice quality profiles to distinguish between groups at elevated risk of Alzheimer's disease. *Speech Communication*.

Al-Tamimi, J., & Palo, P. (2020). Quantifying gradience of epilaryngeal constriction in Levantine Arabic "gutturals": A Generalized Additive Modelling approach to ultrasound tongue contours. In *Proceedings of the 12th International Seminar on Speech Production (ISSP).* Online Virtual Conference (14-18 December 2020).

Al-Tamimi, J., & Palo, P. (2023). Dynamics of the tongue contour in the production of guttural consonants in Levantine Arabic. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 2095–2099). Prague, Czech Republic (7-11 August 2023): Guarant International.

Al-Tamimi, J., & Palo, P. (In preparation). Gradient epilaryngeal constriction of back consonants in Levantine Arabic: A generalised additive mixed modelling of Ultrasound Tongue Imaging. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.

Al-Tamimi, J., Schiel, F., Khattab, G., Sokhey, N., Amazouz, D., Dallak, A., & Moussa, H. (2022). A Romanization System and WebMAUS Aligner for Arabic Varieties. In *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), © European Language Resources Association (ELRA), licensed under CC-BY-NC-4.0* (pp. 7269–7276). Marseille, 20-25 June 2022.

Alghabban, S. (In preparation). *Improving the Perception and Production of English Vowels among Arabic EFL learners in Saudi Arabia* (Unpublished doctoral dissertation). Newcastle University, UK.

Almurashi, W. (2022). *Acoustic Cues in Production of English Vowels by Hijazi Arabic Learners* (Unpublished doctoral dissertation). Newcastle University, UK.

Almurashi, W., Al-Tamimi, J., & Khattab, G. (2019). Static and dynamic cues in vowel production in Hijazi Arabic. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 3468–3472). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Almurashi, W., Al-Tamimi, J., & Khattab, G. (2020a). Acoustic cues in production of english vowels by Hijazi Arabic learners. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).

Almurashi, W., Al-Tamimi, J., & Khattab, G. (2020b). Static and dynamic cues in vowel production in Hijazi Arabic. *The Journal of the Acoustical Society of America: Special Issue on the "Phonetics of Under-Documented Languages"*, *147*(4), 2917–2927. doi: 10.1121/10.0001004

Almurashi, W., Al-Tamimi, J., & Khattab, G. (2021). Acoustic analysis of the acquisition of english vowels by Hijazi Arabic learners. In *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).

Almurashi, W., Al-Tamimi, J., & Khattab, G. (2023). Static and dynamic features of English monophthongal vowels production by Hijazi Arabic L2 learners. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 2562–2566). Prague, Czech Republic (7-11 August 2023): Guarant International.

Almurashi, W., Al-Tamimi, J., & Khattab, G. (Under Review). Acoustic Cues in Production of Southern British English. *Speech Communication*.

Almurashi, W., Al-Tamimi, J., & Khattab, G. (Under Revisions). Dynamic specification of vowels in Hijazi Arabic. *Phonetica*.

Almurashi, W., Khattab, G., & Al-Tamimi, J. (2018). Static and dynamic cues in vowel production in Hijazi Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).

Almurashi, W., Khattab, G., & Al-Tamimi, J. (In preparation). Acoustic Analysis of English Vowel Production by Hijazi Arabic L2 Learners. *The Journal of the Acoustical Society of America*.

Alsharif, B. (In preparation). *Retraction in Postvelar consonants in Levantine Arabic* (Unpublished doctoral dissertation). Newcastle University, UK.

Alsiraih, W. (2013). *Voice quality features in the production of pharyngeal consonants by Iraqi Arabic speakers* (Unpublished doctoral dissertation). Newcastle University, UK.

Aralova, N., Grawunder, S., & Winter, B. (2011). The Acoustic Correlates of Tongue Root Vowel Harmony in Even (Tungusic). In *The 17th International Congress of Phonetic Sciences, Hong Kong* (pp. 240–243).

Arvaniti, A., & Tserdanelis, G. (2000). On the phonetics of geminates: Evidence from Cypriot Greek. In *Proceedings of 6th International Conference on Spoken Language Processing* (Vol. 2, pp. 559–562).

Association, A. P. (Ed.). (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed ed.). Washington, D.C: American Psychiatric Association.

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. Palgrave Macmillan.

Barkat-Defradas, M., Al-Tamimi, J., & Benkirane, T. (2003). Phonetic variation in production and perception of speech: A comparative study of two Arabic dialects. In *Prroceedings of the 15th International Congress of Phonetic Sciences (ICPhS).* Barcelona, Spain (3-9 August 2003).

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Quantitative Psychology and Measurement*, *4*, 1–2. doi: 10.3389/fpsyg.2013.00328

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68(3)*, 255–278. doi:10.1016/j.jml.2012.11.001

Barrett, S. E. (1998). *Prototypes in Speech Perception* (Unpublished doctoral dissertation). University of Cambridge.

Barsties v. Latoszek, B., Maryn, Y., Gerrits, E., & De Bodt, M. (2017). The Acoustic Breathiness Index (ABI): A Multivariate Acoustic Model for Breathiness. *Journal of Voice*, *31*(4), 511.e11-511.e27. doi:10.1016/j.jvoice.2016.11.017

Baumann, S., & Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *Journal of Phonetics*, *70*, 20–38. doi:10.1016/j.wocn.2018.05.004

Becker, J. T., Boiler, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The Natural History of Alzheimer's Disease: Description of Study Cohort and Accuracy of Diagnosis. *Archives of Neurology*, *51*(6), 585–594. doi:10.1001/archneur.1994.00540180063015

Beckman, J., Jessen, M., & Ringen, C. (2009). German Fricatives: Coda Devoicing or Positional Faithfulness? *Phonology*, *26*(2), 231–268.

Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics*, *49*(2), 259–284. doi:10.1017/S0022226712000424

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309. doi:10.1017/S095267570000066X

Bellem, A. (2007). *Towards a Comparative Typology of Emphatics: Across Semitic and Into Arabic Dialect Phonology.* (Unpublished doctoral dissertation). School of Oriental and African Studies (University of London).

Best, C. T. (1993). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In B. Boysson-Bardies, S. Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 167–224). Dordrecht: Springer Netherlands. doi:10.1007/978-94-015-8234-6

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

Best, C. T., Shaw, J., Docherty, G., Evans, B., Foulkes, P., Hay, J., Al-Tamimi, J., Mair, K., Mulak, K., & Wood, S. (2015). From newcastle MOUTH to Aussie ears: Australians' perceptual assimilation and adaptation for Newcastle UK vowels. In *Proceedings of the Annual Conference of the International Speech Communication Association (ISCA), INTERSPEECH* (pp. 1932–1936). Dresden, Germany (6-10 September 2015).

Best, C. T., Shaw, J., Mulak, K., Docherty, G., Evans, B., Foulkes, P., Hay, J., Al-Tamimi, J., Mair, K., & Wood, S. (2015). Perceiving and adapting to regional accent differences among vowel subsystems (Article 0964). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015.

Best, C. T., & Tyler, M. D. (2007). Commonalities and complementarities: Nonnative and second-language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins Publishing Company. doi:10.1075/lllt.17.07bes

Bin-Muqbil, M. S. (2006). *Phonetic and phonological aspects of Arabic emphatics and gutturals* (Unpublished doctoral dissertation). University of Wisconsin-Madison.

Boersma, P., & Weenink, D. (2013). *Praat software*. University of Amsterdam.

Broselow, E., Chen, S.-I., & Huffman, M. (1997). Syllable weight: Convergence of phonology and phonetics. *Phonology*, *14*(1), 47–82.

Cantineau, J. (1960). *Etudes de linguistique arabe*. Paris: Klincksieck.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., Frahm, J., & Harrington, J. (2020). Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time {MRI}. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1). doi:10.5334/labphon.214

Carré, R., & Mrayati, M. (1992). Distinctive regions in acoustic tubes. Speech production modeling. *Journal d'Acoustique*, *5*, 141–159.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh University Press.

Chapin, K., Clarke, N., Garrard, P., & Hinzen, W. (2022). A finer-grained linguistic profile of Alzheimer's disease and Mild Cognitive Impairment. *Journal of Neurolinguistics*, *63*, 101069. doi:10.1016/j.jneuroling.2022.101069

Chen, M. Y. (1995). Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *The Journal of the Acoustical Society of America*, *98*(5), 2443–2453. doi:10.1121/1.414399

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*(4), 2360–2370. doi:10.1121/1.419620

Chen, Y. (2020). *The effect of explicit instruction and auditory/audio-visual training on Chinese EFL learner's perception of intonation* (Unpublished doctoral dissertation). Newcastle University, UK.

Chen, Y., Khattab, G., & Al-Tamimi, J. (2016). Auditory or audio-visual feedback in EFL intonation training. In *Proceedings of the New Sounds 2016: 8th international conference on second-language speech. Aarhus university, denmark.* Aarhus University, Denmark (10-12 June 2016).

Chen, Y., Khattab, G., & Al-Tamimi, J. (2019). The effect of explicit instruction and auditory/audio-visual training on Chinese learners' acquisition of English intonation. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 2286–2290). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Chiba, T., & Kajiyama, M. (1941). *The Vowel: Its Nature and Structure*. Tokyo-Kaiseikan Pub. Co., Ltd., Tokyo.

Chistovich, L. A., & Lublinskaya, V. V. (1979). The 'center of gravity' effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. *Hearing Research*, *1(3)*, 185–195. doi:10.1016/0378-5955(79)90012-1

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Cohen, A. S., Mitchell, K. R., & Elvevåg, B. (2014). What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. *Schizophrenia Research*, *159*(2-3), 533–538. doi: 10.1016/j.schres.2014.09.013

Cohn, A. C. (2003). Phonological Structure and Phonetic Duration: The Role of the Mora. *Working Papers of the Cornell Phonetics Laboratory*, *15*, 69–100.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second Language Speech: Theory and Practice*. Cambridge University Press. doi: 10.1017/CBO9781139087636

Coretta, S., Casillas, J. V., Roessig, S., Franke, M., Ahn, B., Al-Hoorie, A. H., Al-Tamimi, J., Alotaibi, N. E., AlShakhori, M. K., Altmiller, R. M., Arantes, P., Athanasopoulou, A., Baese-Berk, M. M., Bailey, G., Sangma, C. B. A., Beier, E. J., Benavides, G. M., Benker, N., BensonMeyer, E. P., Benway, N. R., Berry, G. M., Bing, L., Bjorndahl, C., Bolyanatz, M., Braver, A., Brown, V. A., Brown, A. M., Brugos, A., Buchanan, E. M., Butlin, T., Buxó-Lugo, A., Caillol, C., Cangemi, F., Carignan, C., Carraturo, S., Caudrelier, T., Chodroff, E., Cohn, M., Cronenberg, J., Crouzet, O., Dagar, E. L., Dawson, C., Diantoro, C. A., Dokovova, M., Drake, S., Du, F., Dubuis, M., Duême, F., Durward, M., Egurtzegi, A., Elsherif, M. M., Esser, J., Ferragne, E., Ferreira, F., Fink, L. K., Finley, S., Foster, K., Foulkes, P., Franzke, R., Frazer-McKee, G., Fromont, R., García, C., Geller, J., Grasso, C. L., Greca, P., Grice, M., Grose-Hodge, M. S., Gully, A. J., Halfacre, C., Hauser, I., Hay, J., Haywood, R., Hellmuth, S., Hilger, A. I., Holliday, N., Hoogland, D., Huang, Y., Hughes, V., Icardo Isasa, A., Ilchovska, Z. G., Jeon, H.-S., Jones, J., Junges, M. N., Kaefer, S., Kaland, C., Kelley, M. C., Kelly, N. E., Kettig, T., Khattab, G., Koolen, R., Krahmer, E., Krajewska, D., Krug, A., Kumar, A. A., Lander, A., Lentz, T. O., Li, W., Li, Y., Lialiou, M., Lima, R. M., Lo, J. J. H., Lopez Otero, J. C., Mackay, B., MacLeod, B., Mallard, M., McConnellogue, C.-A. M., Moroz, G., Murali, M., Nalborczyk, L., Nenadić, F., Nieder, J., Nikolić, D., Nogueira, F. G. S., Offerman, H. M., Passoni, E., Pélissier, M., Perry, S. J., Pfiffner, A. M., Proctor, M., Rhodes, R., Rodríguez, N., Roepke, E., Röer, J. P., Sbacco, L., Scarborough, R., Schaeffler, F., Schleef, E., Schmitz, D., Shiryaev, A., Sóskuthy, M., Spaniol, M., Stanley, J. A., Strickler, A., Tavano, A., Tomaschek, F., Tucker, B. V., Turnbull, R., Ugwuanyi, K. O., Urrestarazu-Porta, I., Van De Vijver, R., Van Engen, K. J., Van Miltenburg, E., Wang, B. X., Warner, N., Wehrle, S., Westerbeek, H., Wiener, S., Winters, S., Wong, S. G.-J., Wood, A., Wottawa, J., Xu, C., Zárate-Sández, G., Zellou, G., Zhang, C., Zhu, J., & Roettger, T. B. (2023). Multidimensional Signals and Analytic Flexibility: Estimating Degrees of Freedom in Human-Speech Analyses. *Advances in Methods and Practices in Psychological Science*, *6*(3), 25152459231162567. doi: 10.1177/25152459231162567

Dallak, A. (2023). *Arabic Obstruents: Laryngeal Contrast and Representation* (Unpublished doctoral dissertation). Newcastle University, UK.

Dallak, A., Khattab, G., & Al-Tamimi, J. (2023). Obstruent Voicing and Laryngeal Features in Arabic. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 2154–2158). Prague, Czech Republic (7-11 August 2023): Guarant International.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America*, *120*(1), 407–415. doi: 10.1121/1.2205133

Davis, S. (2011). Geminates. In *The Blackwell Companion to Phonology* (pp. 1–25). John Wiley & Sons, Ltd. doi: 10.1002/9781444335262.wbctp0037

Derrick, D., Al-Tamimi, J., & Heyne, M. (In preparation). Accounting for within and between-subject variation using generalized additive mixed models on ultrasound tongue contours. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.

DiCanio, C. T. (2012). Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics*, *40*(1), 162–176. doi: 10.1016/j.wocn.2011.10.006

Dickey, C. C., Vu, M.-A. A. T., Voglmaier, M. M., Niznikiewicz, M. A., McCarley, R. W., & Panych, L. P. (2012). Prosodic abnormalities in schizotypal personality disorder. *Schizophrenia Research*, *142*(1-3), 20–30. doi: 10.1016/j.schres.2012.09.006

Dodane, C., & Al-Tamimi, J. (2007). An acoustic comparison of vowel systems in adult-directed-speech and child-directed-speech: Evidence from french, english & japanese. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)* (pp. 1573–1576). Saarbrücken, Germany (6-10 August 2007).

Drager, K., & Hay, J. (2012). Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, *24*(1), 59–78. doi: 10.1017/S0954394512000014

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology*, *23(2)*, 157–191. doi: 10.1017/S095267570600087X

Ehbara, H. (2021). *Impact of computer-assisted pronunciation training on child English language learners* (Unpublished doctoral dissertation). Newcastle University, UK.

Ehbara, H., & Al-Tamimi, J. (2018). The impact of computer assisted pronunciation training on Libyan child learners of english. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* Canterbury University, Kent, UK (12-14 April 2018).

Ehbara, H., Young-Scholten, M., & Al-Tamimi, J. (2021). The role of delayed output on Second/Foreign language pronunciation in children. In C. N. Giannikas (Ed.), *Teaching practices and equitable learning in children's language education* (pp. 23–44). IGI Global. doi: 10.4018/978-1-7998-6487-5.ch002

Embarki, M. (2008). Les dialectes arabes modernes : État et nouvelles perspectives pour la classification géo-sociologique. *Arabica*, *55*(5/6), 583–604. doi: 10.1163/157005808X364616

Embarki, M. (2013). Phonetics. In J. Owens (Ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 24–44). Oxford University Press. doi: 10.1093/oxfordhb/9780199764136.013.0002

Engstrand, O., & Krull, D. (1991). Effects of inventory size on the distribution of vowels in the formant space: Preliminary data from seven languages. *PERILUS (Phonetic Experimental Research at the Institute of Linguistics University of Stockholm)*, *13*, 15–18.

Ennaji, M. (1991). Aspects of multilingualism in the Maghreb. *ijsl*, *1991*(87), 7–26. doi: 10.1515/ijsl.1991.87.7

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization* (No. 113). Utrecht: LOT.

Esling, J. (1999). The IPA Categories "Pharyngeal" and "Epiglottal" Laryngoscopic Observations of Pharyngeal Articulations and Larynx Height. *Language and Speech*, *42(4)*, 349–372. doi:10.1177/00238309990420040101

Esling, J. (2005). There Are No Back Vowels: The Laryngeal Articulator Model. *The Canadian Journal of Linguistics / La revue canadienne de linguistique*, *50*(1), 13–44. doi:10.1353/cjl.2007.0007

Esling, J., Moisik, S., Benner, A., & Crevier-Buchman, L. (2019). *Voice Quality: The Laryngeal Articulator Model* (1st ed.). Cambridge University Press. doi:10.1017/9781108696555

Esposito, A., & DiBenedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America*, *106*(4, Pt. 1), 2051–2062. doi:10.1121/1.428056

Eyben, F. (2016). *Real-time Speech and Music Classification by Large Audio Feature Space Extraction*. Cham: Springer International Publishing. doi:10.1007/978-3-319-27299-3

Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., André, E., Busso, C., Devillers, L. Y., Epps, J., Laukka, P., Narayanan, S. S., & Truong, K. P. (2016). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transactions on Affective Computing*, *7*(2), 190–202. doi:10.1109/TAFFC.2015.2457417

Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile: The munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia* (pp. 1459–1462). Firenze Italy: ACM. doi:10.1145/1873951.1874246

Fahey, R. P., Diehl, R. L., & Traunmüller, H. (1996). Perception of back vowels: Effects of varying F1-F0 Bark distance. *Journal of the Acoustical Society of America*, *99*(4), 2350–2357. doi:10.1121/1.415422

Fant, G. (1960). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations, second edition*. Mouton.

Farinas, J., Rouas, J.-L., Pellegrino, F., & André-Obrecht, R. (2005). Extraction automatique de paramètres prosodiques pour l'identification automatique des langues (Automatic extraction of prosodic parameters for the use in automatic language identification). *Traitement du Signal*, *22(2)*, 81–97.

Ferguson, C. A. (1959). Diglossia. *Word*, *15*(2), 325–340. doi:10.1080/00437956.1959.11659702

Ferragne, E., Gendrot, C., & Pellegrini, T. (2019). Towards phonetic interpretability in deep learning applied to voice comparison. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 790–794). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Flege, J. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47–65. doi:10.1016/S0095-4470(19)30537-6

Flege, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–273). Timonium, MD: York Press.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409–438. doi:10.1016/j.wocn.2005.08.002

Fulop, S. A., Kari, E., & Ladefoged, P. (1998). An Acoustic Study of the Tongue Root Contrast in Degema Vowels. *Phonetica*, *55(1-2)*, 80–98. doi:10.1159/000028425

Garellek, M. (2012). The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics*, *40(1)*, 152–161. doi:10.1016/j.wocn.2011.10.003

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. F. Assmann (Eds.), *The Routledge Handbook of Phonetics* (1st ed., pp. 75–106). Abingdon, Oxon ; New York, NY : Routledge, 2019. | Series: Routledge handbooks in linguistics: Routledge. doi:10.4324/9780429056253-5

Garellek, M. (2020). Acoustic Discriminability of the Complex Phonation System in !Xóõ. *Phonetica*, *77*(2), 131–160. doi:10.1159/000494301

Garellek, M., Samlan, R., Gerratt, B. R., & Kreiman, J. (2016). Modeling the voice source in terms of spectral slopes. *The Journal of the Acoustical Society of America*, *139*(3), 1404–1410. doi:10.1121/1.4944474

Gendrot, C., Ferragne, E., & Pellegrini, T. (2019). Deep learning and voice comparison: Phonetically motivated vs. automatically learned features. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 1332–1336). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Gerrits, E. (2001). *The categorisation of speech sounds by adults and children: A study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues* (Unpublished doctoral dissertation). LOT, Utrecht, The Netherlands.

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *66*(3), 363–376. doi:10.3758/BF03194885

Ghazeli, S. (1977). *Back Consonants and Backing Coarticulation in Arabic* (Unpublished doctoral dissertation). University of Texas at Austin, USA.

Gick, B., Wilson, I., Koch, K., & Cook, C. (2004). Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica*, *61(4)*, 220–233. doi:10.1159/000084159

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.

Gu, C. (2014). Smoothing Spline ANOVA Models: R Package gss. *Journal of Statistical Software*, *58*(5), 1–25-1–25.

Guion, S. G., Post, M. W., & Payne, D. L. (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics*, *32(4)*, 517–542. doi:10.1016/j.wocn.2004.04.002

Haeri, N. (1991). *Sociolinguistic variation in Cairene Arabic: Palatalization and the ${$$$it qaf$$/$}$ in the speech of men and women* (Unpublished doctoral dissertation). University of Pennsylvania.

Haeri, N. (1992). How different are men and women: Palatalization in Cairo. *Perspectives on Arabic linguistics IV. Amsterdam and Philadelphia: John Benjamins*, 165–180.

Haeri, N. (1994). A linguistic innovation of women in Cairo. *Language variation and change*, *6*(1), 87–112.

Haider, F., Luz, S., Fromm, D., & MacWhinney, B. (Eds.). (2021). *Alzheimer's Dementia Recognition Through Spontaneous Speech*. Frontiers Media SA. doi:10.3389/978-2-88971-854-2

Halle, M., & Stevens, K. N. (1969). On the Feature 'Advanced Tongue Root'. In *Quarterly Progress Report No. 94, Research Laboratory of Electronics, M.I.T* (pp. 209–215).

Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. *MIT Quarterly Progress Report*, *101*, 198–212.

Ham, W. (2001). *Phonetic and Phonological Aspects of Geminate Timing*. Routledge.

Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America*, *101*(1), 466–481.

Hanson, H. M., & Chuang, E. S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America*, *106*(2), 1064–1077.

Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, *29*(4), 451–480. doi:10.1006/jpho.2001.0146

Harris, J. (1994). *English sound structure*. Oxford, UK ; Cambridge, Mass: Blackwell.

Haspelmath, M., & Sims, A. D. (2010). *Understanding morphology* (2nd ed ed.). London: Hodder Education.

Hassan, Z. M., & Esling, J. H. (2011). Investigating the emphatic feature in Iraqi Arabic: Acoustic and articulatory evidence of coarticulation. In B. Heselwood & Z. Hassan (Eds.), *Instrumental studies in Arabic phonetics* (pp. 218–234). John Benjamins. doi:10.1075/cilt.319

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The Elements of Statistical Learning* (Second Edition ed.). New York, NY: Springer New York. doi:10.1007/978-0-387-84858-7

Hawkins, S. (2003). Contribution of Fine Phonetic Detail to Speech Understanding. In *15th ICPhS* (p. 4).

Hawkins, S. (2010). Phonological features, auditory objects, and illusions. *Journal of Phonetics*, *38*(1), 60–89. doi:10.1016/j.wocn.2009.02.001

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, *20(2)*, 253–306.

He, R., Al-Tamimi, J., Sánchez-Benavides, G., Montaña-Valverde, G., Gispert, J. D., Grau-Rivera, O., Suárez-Calvet, M., Minguillon, C., Fauria, K., Navarro, A., & Hinzen, W. (Under Review). Atypical cortical hierarchy in Aβ-positive older adults in the context of atypical speech prosody. *Neurobiology of Aging*.

He, R., Chapin, K., Al-Tamimi, J., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (2023). Automated classification of cognitive decline and probable Alzheimer's dementia across multiple speech and language domains. *American Journal of Speech-Language Pathology*, *32*(5), 2075–2086. doi:10.1044/2023_AJSLP-22-00403

Heidlmayr, K., Ferragne, E., & Isel, F. (2021). Neuroplasticity in the phonological system: The PMN and the N400 as markers for the perception of non-native phonemic contrasts by late second language learners. *Neuropsychologia*, *156*, 107831. doi:10.1016/j.neuropsychologia.2021.107831

Hellmuth, S. (2013). Phonology. In J. Owens (Ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 46–70). Oxford University Press. doi:10.1093/oxfordhb/9780199764136.013.0003

Henton, C. (1989). Fact and fiction in the description of female and male pitch. *Language & Communication*, *9*(4), 299–311. doi:10.1016/0271-5309(89)90026-8

Henton, C. (1995a). Cross-language variation in the vowels of female and male speakers. In *Proceedings of the XIIIth international congress of phonetic sciences* (Vol. 4, pp. 420–423).

Henton, C. (1995b). Pitch dynamism in female and male speech. *Language & Communication*, *15*(1), 43–61. doi: 10.1016/0271-5309(94)00011-Z

Heselwood, B. (2007). The 'tight approximant' variant of the Arabic 'ayn. *Journal of the International Phonetic Association*, *37*(01), 1. doi:10.1017/S0025100306002787

Heselwood, B., & Al-Tamimi, F. (2011). A study of the laryngeal and pharyngeal consonants in Jordanian Arabic using nasoendoscopy, videofluoroscopy and spectrography. In B. Heselwood & Z. Hassan (Eds.), *Instrumental studies in Arabic phonetics* (pp. 101–128). John Benjamins. doi:10.1075/cilt.319

Hess, S. A. (1998). *Pharyngeal Articulations* (Unpublished doctoral dissertation). UCLA, Los Angeles, California, USA.

Heyne, M. (2016). *The influence of first language on playing brass instruments: An ultrasound study of Tongan and New Zealand trombonists.* (Unpublished doctoral dissertation). Christchurch, New Zealand: University of Canterbury.

Heyne, M., Al-Tamimi, J., & Derrick, D. (2019). Using Generalized Additive (Mixed) Models to uncover novel patterns of speech production and account for by-group random patterns of variation. In *Proceedings of the Boston Speech Motor Control Symposium.* Boston University, Boston, USA (21 June 2019).

Heyne, M., Derrick, D., & Al-Tamimi, J. (2019). Native language influence on brass instrument performance: An application of generalized additive mixed models (GAMMs) to midsagittal ultrasound images of the tongue. *Frontiers in Psychology*, *10*(2597), 1–26. doi:10.3389/fpsyg.2019.02597

Hillenbrand, J. M. (2013). Static and Dynamic Approaches to Vowel Perception. In G. S. Morrison & P. F. Assmann (Eds.), *Vowel Inherent Spectral Change* (pp. 9–30). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-14209-3_2

Hoekert, M., Kahn, R., Pijineborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research*, *96*(1-3), 135–145. doi:10.1016/j.schres.2007.07.023

Hoemeke, K. A., & Diehl, R. L. (1994). Perception of vowel height: The role of F1-F0 distance. *Journal of the Acoustical Society of America*, *96*(2), 661–674. doi:10.1121/1.410305

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female

speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, *84*(2), 511–529. doi: 10.1121/1.396829
Hombert, J.-M., & Puech, G. (1984). Espace vocalique et structuration perceptuelle : Application au Swahili. *PHOLIA (PHOnétique et LInguistique Africaine)*, *1*, 199–208.
Honeybone, P. (2005). Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications. In *Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications* (pp. 317–352). De Gruyter Mouton. doi: 10.1515/9783110890402.317
Huang, Y., & Ferreira, F. (2020). The Application of Signal Detection Theory to Acceptability Judgments. *Frontiers in Psychology*, *11*, 73. doi: 10.3389/fpsyg.2020.00073
Hussain, Q., & Shinohara, S. (2019). Partial devoicing of voiced geminate stops in Tokyo Japanese. *The Journal of the Acoustical Society of America*, *145*(1), 149–163. doi: 10.1121/1.5078605
Idemaru, K., & Guion, S. G. (2008). Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association*, *38*, 167–186. doi: 10.1017/S0025100308003459
Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America*, *121*(4), 2283–2295. doi: 10.1121/1.2697522
Ivanova, O., Martínez-Nicolás, I., & Meilán, J. J. G. (2023). Speech changes in old age: Methodological considerations for speech-based discrimination of healthy ageing and Alzheimer's disease. *International Journal of Language & Communication Disorders*, 1–25. doi: 10.1111/1460-6984.12888
Iverson, G. K., & Salmons, J. C. (1995). Aspiration and Laryngeal Representation in Germanic. *Phonology*, *12*(3), 369–396.
Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, *33*(1), 145–160. doi: 10.1017/S0142716411000300
Jaeger, J. (1983). The fortis/lenis question: Evidence from Zapotec and Jawoñ. *Journal of Phonetics*, *11*, 177–189.
Jakobson, R. (2002). Mufaxxama – the 'Emphatic' Phonemes in Arabic. In *Phonological Studies* (pp. 510–522). De Gruyter Mouton. doi: 10.1515/9783110892499.510
Jakobson, R., Fant, G., & Halle, M. (1952/1976). *Preliminaries to Speech Analysis: The distinctive features and their correlates*. Cambridge, Massachusetts: MIT Press.
Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch* (Unpublished doctoral dissertation). University Library Groningen.
Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German* (Vol. 44). John Benjamins Publishing.
Jessen, M. (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In T. A. Hall (Ed.), *Distinctive Feature Theory* (pp. 237–294). Berlin, Boston: De Gruyter Mouton. doi: 10.1515/9783110886672.237
Jessen, M., & Ringen, C. (2002). Laryngeal features in German. *Phonology*, *19*(2), 189–218.
Jesus, L. M., & Shadle, C. H. (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics*, *30*(3), 437–464. doi: 10.1006/jpho.2002.0169
Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego: Academic Press.
Johnson, K. (2000). Adaptive Dispersion in Vowel Perception. *Phonetica*, *57*(2-4), 181–188. doi: 10.1159/000028471
Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*(4), 485–499. doi: 10.1016/j.wocn.2005.08.004
Johnson, K., Flemming, E., & Wright, R. (1993). The Hyperspace Effect: Phonetic Targets Are Hyperarticulated. *Language*, *69*(3), 505. doi: 10.2307/416697
Johnson, K., Sherman, V. C., & Sherman, S. G. (2012). *Acoustic and Auditory Phonetics - 3rd edition* (3rd ed.).
Johnson, K., & Sjerps, M. J. (2021). Speaker Normalization in Speech Perception. In J. S. Pardo, L. C. Nygaard, R. E. Remez, & D. B. Pisoni (Eds.), *The Handbook of Speech Perception* (1st ed., pp. 145–176). Wiley. doi: 10.1002/9781119184096.ch6
Jones, K. S. (2016). The acoustics of word-final fake gemination in Egyptian Arabic. In *PROCEEDINGS of the 22nd International Congress on Acoustics* (pp. 1–10). Buenos Aires – 5 to 9 September, 2016 Acoustics for the 21st Century....
Jones, W., & Al-Tamimi, J. (2020). Using Prosogram's automatic prosodic profiling in diagnosing schizophrenia. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).
Jones, W., Al-Tamimi, J., & Hinzen, W. (2021). The role of automatic acoustic profiling of prosody in schizophrenia. In *Proceedings of the 4th PaPE (Phonetics and Phonology in Europe).* Barcelona, Spain (21-23 June 2021).
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. doi: 10.1016/j.wocn.2010.11.007
Joos, M. (1948). Acoustic Phonetics. *Language*, *24*(2), 5. doi: 10.2307/522229
Jordan, E., & Al-Tamimi, J. (2023). Classification d'Alzheimer à partir de paramètres acoustiques et prosodique avec de l'apprentissage automatique. In *Actes des 9èmes Journées de Phonétique Clinique (JPC2023): "Prendre la mesure de la Parole".* Toulouse 15-17 juin 2023.
Kamphikul, N. (2022). *Intonation of Thai Learners for Marking Narrow Focus* (Unpublished doctoral dissertation). Newcastle University, UK.
Kang, H., & Ko, S. (2012). In search of the acoustic correlates of tongue root retraction in three Altaic languages: Western Buriat, Tsongol Buriat, and Ewen. *Altai Hakpo*, *22*, 179–203.
Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 286–319. doi: 10.2307/413642
Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. In *The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK: The University of Glasgow. ISBN 978-0-85261-941-4. Paper number 0821.*

Kendall, T., Pharao, N., Stuart-Smith, J., & Vaughn, C. (2023). Advancements of phonetics in the 21st century: Theoretical issues in sociophonetics. *Journal of Phonetics*, *98*, 101226. doi:10.1016/j.wocn.2023.101226

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and Overlap in the Speech Chain. *Language*, *82(1)*, 33–63. doi:10.1353/lan.2006.0051

Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and Auditory Differences in the /t/-/ṭ/ Opposition in Male and Female Speakers of Jordanian Arabic. In S. Boudelaa (Ed.), *Current Issues in Linguistic Theory* (Vol. 266, pp. 131–160). Amsterdam: John Benjamins Publishing Company. doi:10.1075/cilt.266.09kha

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

Khattab, G., & Al-Tamimi, J. (2011a). Phonological patterns in early word production by Lebanese-Arabic speaking children. In *Proceedings of the Child Language Seminar.* Newcastle University, UK (13-14 June 2011).

Khattab, G., & Al-Tamimi, J. (2011b). Phonological templates in lebanese arabic. In *Proceedings of the 12th IASCL (International Association for the Study of Child Language).* Montreal, Canada (19-23 July 2011).

Khattab, G., & Al-Tamimi, J. (2011c). The role of geminates in shaping early word patterns by Lebanese-Arabic speaking children. In *Proceeings of the International Child Phonology Conference.* York University, UK (15-18 June 2011).

Khattab, G., & Al-Tamimi, J. (2012a). Learning to control phonetic timing for phonological purposes: The acquisition of gemination by Lebanese-speaking children. In *Proceedings of the International Child Phonology Conference (ICPC).* Minneapolis, USA (04-06 June 2012).

Khattab, G., & Al-Tamimi, J. (2012b). Learning to control phonetic timing for phonological purposes: The acquisition of gemination by Lebanese-speaking children. In *Proceedings of BAAP (The British Association of Academic Phoneticians Colloquium,.* Leeds University, Leeds, UK (26-28 March 2012).

Khattab, G., & Al-Tamimi, J. (2013a). Influence of geminate structure on early Arabic templatic patterns. In M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic Evidence* (pp. 374–414). Cambridge University Press.

Khattab, G., & Al-Tamimi, J. (2013b). The role of French loan word adaptation in the acquisition of consonant length in Lebanese Arabic. In *Proceedings of the International Child Phonology Conference (ICPC).* Radboud University, Nijmegen, Netherlands (10-12 June 2013).

Khattab, G., & Al-Tamimi, J. (2014a). Disentangling phonetic from phonological length: The effect of multilingual input on Lebanese-Arabic-learning children's production. In *Proceedings of BAAP (The British Association of Academic Phoneticians) Colloquium.* Oxford University, Oxford, UK (7-9 April 2014).

Khattab, G., & Al-Tamimi, J. (2014b). Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *5*(2), 231–269. doi:10.1515/lp-2014-0009

Khattab, G., & Al-Tamimi, J. (2015a). The acquisition of gemination in Lebanese Arabic children (article: 0870). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015.

Khattab, G., & Al-Tamimi, J. (2015b). Spectral properties of vowels surrounding singleton and geminate sonorants in Lebanese Arabic. In *In Proceedings of GemCon2015 - ICPhS 2015 satellite workshop on 'Geminate consonants across the world'.* Glasgow University, Glasgow, UK (12 August 2015).

Khattab, G., & Al-Tamimi, J. (2017). Contact-induced constraints on the acquisition of phonological length contrast by Lebanese Arabic-speaking children. In *Proceedings of the 14th IASCL (International Association for the Study of Child Language).* Lyon, France (17-21 July 2017).

Khattab, G., & Al-Tamimi, J. (2018). When lenis and fortis properties collide: The case of voiced geminates in Lebanese Arabic. In *Proceedings of the 16th LabPhon (laboratory phonology).* Lisbon, Portugal (20-22 June 2018).

Khattab, G., Al-Tamimi, J., & Alsiraih, W. (2018). Nasalisation in the production of Iraqi Arabic pharyngeals. *Phonetica*, *75*(4), 310–348. doi:10.1159/000487806

Khattab, G., Vihman, M., Al-Tamimi, J., Nakai, S., & Kunnari, S. (2014). A comparison of long consonant acquisition in Arabic, Finnish, Japanese and Welsh. In *Proceedings of the 14th Laboratory Phonology Conference (LabPhon14).* Tokyo, Japan (25-27 July 2014).

Kim, M. R., Beddor, P. S., & Horrocks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics*. doi:10.1006/jpho.2001.0152

Kingston, J., & Diehl, R. L. (1994). Phonetic Knowledge. *Language*, *70*(3), 419–454. doi:10.2307/416481

Kitikanan, P. (2016). *L2 English fricative production by Thai learners* (Unpublished doctoral dissertation). Newcastle University, UK.

Kitikanan, P., & Al-Tamimi, J. (2012). The earliest stage of voiceless fricative acquisition among Thai learners of Mandarin Chinese. *Annual Review of Education, Communication & Language Sciences*, *9*, 91–114.

Kitikanan, P., Al-Tamimi, J., & Khattab, G. (2015). An acoustic investigation of the production of English /s/ by L2 Thai learners (Article: 0655). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* University of Glasgow., Glasgow, UK (10-14 August 2015): The scottish consortium for ICPhS 2015.

Kitikanan, P., Al-Tamimi, J., & Young-Scholten, M. (2012). The earliest stage of fricative acquisition among Thai learners of Mandarin Chinese. In *Proceedings of the Second Language Acquisition of Phonology Conference.* York University, York, UK (6 July 2012).

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, *87*(2), 820–857. doi:10.1121/1.398894

Kohler, K. (1984). The phonetic explanation in phonology: The feature fortis/lenis. *Phonetica*, *41*, 150–174. doi:10.1159/000261721

Kreiman, J., Lee, Y., Garellek, M., Samlan, R., & Gerratt, B. R. (2021). Validating a psychoacoustic model of voice quality. *The Journal of the Acoustical Society of America*, *149*(1), 457–465. doi:10.1121/10.0003331

Kuang, J., & Keating, P. (2012). Glottal articulations of phonation contrasts and their acoustic and perceptual consequences. *UCLA Working Papers in Phonetics*, *111*, 123–161.

Kuang, J., & Keating, P. (2014). Vocal fold vibratory patterns in tense versus lax phonation contrasts. *Journal of the Acoustical Society of America*, *136*(5), 2784–2797. doi:10.1121/1.4896462

Kuhl, P. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93–107. doi:10.3758/BF03212211

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science*, *277*(5326), 684–686. doi:10.1126/science.277.5326.684

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer.

Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge University Press. doi:10.1017/CBO9780511808814

Ladd, D. R., & Schmid, S. (2018). Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics*, *71*, 229–248. doi:10.1016/j.wocn.2018.09.003

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

Laufer, A., & Baer, T. (1988). The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech*, *31*, 181–205. doi:10.1177/002383098803100205

Laver, J. (1980). *The phonetic description of voice quality*. Cambridge University Press.

Laver, J. (1994). *Principles of Phonetics*. Cambridge University Press.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, *105*(3), 1455–1468. doi:10.1121/1.426686

Lehn, W. (1963). Emphasis in Cairo Arabic. *Language*, *39(1)*, 29–39. doi:10.2307/410760

Li, C. (2022). *The role of factors influencing the realization of diphthongs in Standard Mandarin: Dynamic modeling with GAMMs* (Unpublished master's thesis). Université Paris Cité.

Li, C. (In preparation). *The role of factors influencing the realization of diphthongs in Mandarin: A sociophonetic and phonological perspective* (Unpublished doctoral dissertation). Université Paris Cité.

Li, C., & Al-Tamimi, J. (Under Review). The Impact of the Tonal Factor on Diphthong Realizations in Standard Mandarin: Explanations within the Articulatory Phonology Framework modelled via Generalized Additive Mixed Models. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.

Li, C., Al-Tamimi, J., & Wu, Y. (2023). Tone as a factor influencing the dynamics of diphthong realizations in Standard Mandarin. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 1876–1880). Prague, Czech Republic (7-11 August 2023): Guarant International.

Li, Y. (2015). *Audio-visual training effect on L2 perception and production of English /θ/-/s/ and /ð/-/z/ by Mandarin speakers* (Unpublished doctoral dissertation). Newcastle University, UK.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461. doi:10.1037/h0020279

Liljencrants, J., & Lindblom, B. (1972). Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast. *Language*, *48*(4), 839–862. doi:10.2307/411991

Lindblom, B. (1963). *On vowel reduction*. Quarterly Progress Status Report 29, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden.

Lindblom, B. (1986). On the origin and purpose of discreteness and invariance in sound patterns. In *Invariance and variability in speech processes* (pp. 493–510). Erlbaum Hillsdale, NJ.

Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403–439). Dordrecht: Springer Netherlands. doi:10.1007/978-94-009-2037-8_16

Lindblom, B. (1996). Role of Articulation in speech perception: Clues from production. *Journal of the Acoustical Society of America*, *99*(3), 1683–1692.

Lindblom, B., & Sussman, H. M. (2012). Dissecting coarticulation: How locus equations happen. *Journal of Phonetics*, *40*(1), 1–19. doi:10.1016/j.wocn.2011.09.005

Lindsay, H., Tröger, J., & König, A. (2021). Language Impairment in Alzheimer's Disease—Robust and Explainable Evidence for AD-Related Deterioration of Spontaneous Speech Through Multilingual Machine Learning. *Frontiers in Aging Neuroscience*, *13*.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422. doi:10.1080/00437956.1964.11659830

Livijn, P. (2000). Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion? In *In the Proceedings of the XIIIth Swedish Phonetics Conference (FONETIK 2000)* (pp. 93–96).

Local, J., & Simpson, A. (1999). Phonetic implementation of geminates in Malayalam nouns. In *Proceedings of the 14th ICPhS* (pp. 595–598).

Lofgren, M., Al-Tamimi, J., Chapin, K., García-Gutiérrez, F., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J., Sanabaria, A., Alegret, M., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (Under Review). An automated acoustic-prosodic analysis pipeline for identifying early cognitive decline and Alzheimer's disease. *Journal of Psychiatric Research*.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874–886. doi:10.1121/1.1894649

Lubbers, M., & Torreira, F. (2013/2018). *Praatalign: An interactive Praat plug-in for performing phonetic forced alignment: https://github.com/dopefishh/praatalign. Version 2.0a.*

Luz, S., Haider, F., de la Fuente, S., Fromm, D., & MacWhinney, B. (2020). Alzheimer's Dementia Recognition Through Spontaneous Speech: The ADReSS Challenge. In *Interspeech 2020* (pp. 2172–2176). ISCA. doi:10.21437/Interspeech.2020-2571

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed ed.). Mahwah, N.J: Lawrence Erlbaum Associates.

Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *Journal of the Acoustical Society of America*, *125*(6), 3962–3973.

Manuel, S. Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *The Journal of the Acoustical Society of America*, *88*(3), 1286–1298. doi: 10.1121/1.399705

Martin, P. (2002). Le système vocalique du français du Québec. De l'acoustique à la phonologie. *La linguistique*, *38*(2), 71–88. doi: 10.3917/ling.382.0071

Martínez-Nicolás, I., Llorente, T. E., Martínez-Sánchez, F., & Meilán, J. J. G. (2021). Ten Years of Research on Automatic Voice and Speech Analysis of People With Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article. *Frontiers in Psychology*, *12*.

Martínez-Sánchez, F., Muela-Martínez, J. A., Cortés-Soto, P., Meilán, J. J. G., Ferrándiz, J. A. V., Caparrós, A. E., & Valverde, I. M. P. (2015). Can the Acoustic Analysis of Expressive Prosody Discriminate Schizophrenia? *The Spanish Journal of Psychology*, *18*, E86. doi: 10.1017/sjp.2015.85

Mayo-Yáñez, M., & Cabo-Varela, I. (2020). Voice and Speech Disorders in Alzheimer's Disease: A State-of-the-Art Review. *OBM Neurobiology*, *4*(3), 1–9. doi: 10.21926/obm.neurobiol.2003072

McAleer, P., & Belin, P. (2018). The Perception of Personality Traits from Voices. In *The Oxford Handbook of Voice Perception* (pp. 584–606). Oxford University Press. doi: 10.1093/oxfordhb/9780198743187.013.26

McCarthy, D. (2019). *The acoustics of place of articulation in English plosives* (Unpublished doctoral dissertation). Newcastle University, UK.

McCarthy, D., & Al-Tamimi, J. (2019). F2R: A Technique for collapsing F2 onset and F2 mid into a single acoustic attribute. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 3777–3781). Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc.

McCarthy, J. (1979). *Formal Problems in Semitic Phonology and Morphology* (Unpublished doctoral dissertation). MIT.

McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. Keating (Ed.), *Phonological Structure and Phonetic Form* (pp. 191–233). Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511659461.012

McKay, G. (1980). Medial stop gemination in Rembarrnga: A spectrographic study. *Journal of Phonetics*, *8*, 343–352.

McMurray, B. (2022). The myth of categorical perception. *The Journal of the Acoustical Society of America*, *152*(6), 3819–3842. doi: 10.1121/10.0016614

Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., López, D. E., Millian-Morell, L., & Arana, J. M. (2014). Speech in Alzheimer's Disease: Can Temporal and Acoustic Parameters Discriminate Dementia? *Dementia and Geriatric Cognitive Disorders*, *37*(5-6), 327–334. doi: 10.1159/000356726

Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., Sánchez, J. A., & Pérez, E. (2012). Acoustic Markers Associated with Impairment in Language Processing in Alzheimer's Disease. *The Spanish journal of psychology*, *15*(2), 487–494. doi: 10.5209/rev_SJOP.2012.v15.n2.38859

Mertens, P. (2004). The prosogram: Semi-automatic transcription of prosody based on a tonal perception model. *Paper presented at the Speech Prosody, 2004*.

Meunier, C., Frenck-Mestre, C., Lelekov-Boissard, T., & Besnerais, M. L. (2003). Production and perception of vowels: Does the density of the system play a role? In *In Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)* (pp. 723–726).

Michaud, A. (2004). A measurement from electroglottography: DECPA, and its application in prosody. In *Speech Prosody 2004* (pp. pp–633). International Speech Communication Association (ISCA).

Miller, J. D. (1989). Auditory-perceptual interpretation of the vowel. *The Journal of the Acoustical Society of America*, *85*(5), 2114–2134. doi: 10.1121/1.397862

Mittal, V. K., Yegnanarayana, B., & Bhaskararao, P. (2014). Study of the effects of vocal tract constriction on glottal vibration. *The Journal of the Acoustical Society of America*, *136*(4), 1932–1941. doi: 10.1121/1.4894789

Moisik, S. (2013a). *The Epilarynx in Speech* (Thesis).

Moisik, S. (2013b). Harsh voice quality and its association with blackness in popular American media. *Phonetica*, *69(4)*, 193–215. doi: 10.1159/000351059

Moisik, S., Czaykowska-Higgins, E., & Esling, J. (2012). The epilaryngeal articulator: A new conceptual tool for understanding lingual-laryngeal contrasts. *McGill Working Papers in Linguistics*, *22(1)*.

Moisik, S., Czaykowska-Higgins, E., & Esling, J. (2019). Phonological potentials and the lower vocal tract. *Journal of the International Phonetic Association*, 1–35. doi: 10.1017/s0025100318000403

Moisik, S., & Esling, J. (2010). Examining the Acoustic Contributions of the Epilaryngeal Tube to the Voice Source and Vocal Tract Resonance. *Canadian Acoustics*, *38(3)*, 138–139.

Moisik, S., & Esling, J. (2011). The 'whole larynx' approach to laryngeal features. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS), Hong Kong* (pp. 1406–1409).

Moisik, S., Esling, J., & Crevier-Buchman, L. (2010). A high-speed laryngoscopic investigation of aryepiglottic trilling. *Journal of the Acoustical Society of America*, *127*(3), 1548–1558. doi: 10.1121/1.3299203

Mok, P. P. (2013). Does Vowel Inventory Density Affect Vowel-to-Vowel Coarticulation? *Language and Speech*, *56*(2), 191–209. doi: 10.1177/0023830912443948

Morrison, G. S., & Assmann, P. F. (Eds.). (2013). *Vowel inherent spectral change*. Heidelberg; New York: Springer.

Moussa, H., & Al-Tamimi, J. (2018). Focus and F0 patterns in Jeddah Arabic. In *Proceedings of BAAP (The British Association of Academic Phoneticians)*. Canterbury University, Kent, UK (12-14 April 2018).

Moussa, H., & Al-Tamimi, J. (In preparation). Marking Focus in Jeddah Arabic. *Laboratory Phonology: Journal of the Association for*

Laboratory Phonology.

Moussa, H. A. (2019). *Jeddah Arabic intonation : An autosegmental-metrical approach* (Unpublished doctoral dissertation). Newcastle University, UK.

Mrayati, M., Carré, R., & Guérin, B. (1988). Distinctive regions and modes: A new theory of speech production. *Speech Communication*, *7(3)*, 257–286. doi:10.1016/0167-6393(88)90073-8

Nasr, R. T. (1966). *Colloquial Arabic: An oral approach*. Beirut: Librarie du Liban.

Nellis, D. G., & Hollenbach, B. E. (1980). Fortis versus Lenis in Cajonos Zapotec Phonology. *International Journal of American Linguistics*. doi:10.1086/465639

Nguyen, N. (2001). Rôle de la coarticulation dans la reconnaissance des mots. *Année psychologique*, *101*, 125–154.

Nguyen, N. (2005). La perception de la parole. In N. Nguyen, S. Wauquier, & J. Durand (Eds.), *Phonologie et phonétique* (pp. 425–447). Hermès.

Nguyen, N., Wauquier, S., & Tuller, B. (2009). The dynamical approach to speech perception: From fine phonetic detail to abstract phonological categories. In F. Pellegrino, E. Marsico, I. Chitoran, & C. Coupé (Eds.), *Approaches to Phonological Complexity* (pp. 191–218). Walter de Gruyter. doi:10.1515/9783110223958.191

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

Ouakrim, O. (2003). A Perceptual Study of Tenseness. Some Acoustic Cues Identifying Tense Vs Non-Tense Contrast in Berber. In *XV Internation Congress on Phonetic Sciences* (pp. 795–798).

Overall, J. E., & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports*, *10*(3), 799–812. doi:10.2466/pr0.1962.10.3.799

Parola, A., Simonsen, A., Bliksted, V., & Fusaroli, R. (2019). Voice patterns in schizophrenia: A systematic review and Bayesian meta-analysis. *Schizophrenia Research*. doi:10.1016/j.schres.2019.11.031

Peterson, G. E., & Barney, H. L. (1952). Control Methods Used in a Study of the Vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. doi:10.1121/1.1906875

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (Thesis). Massachusetts Institute of Technology.

Pierrehumbert, J. B. (2000). The Phonetic Grounding of Phonology. *Bulletin de la Communication Parlee*, *5*, 7–23.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee & P. J. Hopper (Eds.), *Typological Studies in Language* (Vol. 45, p. 137). Amsterdam: John Benjamins Publishing Company. doi:10.1075/tsl.45.08pie

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics*, *34*(4), 516–530. doi:10.1016/j.wocn.2006.06.003

Pierrehumbert, J. B. (2016). Phonological Representation: Beyond Abstract Versus Episodic. *Annual Review of Linguistics*, *2*(1), 33–52. doi:10.1146/annurev-linguistics-030514-125050

Punnoose, R., Khattab, G., & Al-Tamimi, J. (2013). The contested fifth liquid in Malayalam: A window into the lateral-rhotic relationship in dravidian languages. *Phonetica*, *70*(4), 274–297. doi:10.1159/000356359

Reisberg, B., & Gauthier, S. (2008). Current evidence for subjective cognitive impairment (SCI) as the pre-mild cognitive impairment (MCI) stage of subsequently manifest Alzheimer's disease. *International Psychogeriatrics*, *20*(1), 1–16. doi:10.1017/S1041610207006412

Ridouane, R. (2003). Geminates vs. Singleton Stops in Berber: An Acoustic, Fiberscopic and Photoglottographic study. In *XIV International Congress on Phonetic Sciences* (pp. 1743–1746).

Ridouane, R. (2006). On the feature [spread glottis].

Ridouane, R. (2007). Gemination in Tashlhiyt Berber: An Acoustic and Articulatory Study. *Journal of the International Phonetic Association*, *37*(2), 119–142. doi:10.1017/S0025100307002903

Ridouane, R. (2010). Geminates at the junction of phonetics and phonology. In *Geminates at the junction of phonetics and phonology* (pp. 61–90). De Gruyter Mouton. doi:10.1515/9783110224917.1.61

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon: A Software Solution for the Study of Phonological Acquisition. *Proceedings of the ... Annual Boston University Conference on Language Development. Boston University Conference on Language Development*, *2006*, 489–500.

Royal, A.-M. (1985). Male-Female Pharyngealization Patterns in Cairo Arabic: A Sociolinguistic Study of Two Neighborhoods. In *Texas Linguistic Forum.*

Schiel, F. (2015). A statistical model for predicting pronunciation. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS).* The scottish consortium for ICPhS 2015.

Schiel, F. (2018). *The Munich Automatic Segmentation System (MAUS).*

Schiel, F., & Al-Tamimi, J. (2021). *WebMAUS and WebMINNI Arabic: Online resources for forced-alignment and speech-to-text transcription* [Development of an Arabic Variety-Independent Romanization System Integrated within an Arabic Variety-Independent WebMAUS Forced Alignment and WebMINNI Services (Integrated with WebMAUS Version 3.8)]. https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface.

Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J. K., Baird, A., Elkins, A., Zhang, Y., Coutinho, E., & Evanini, K. (2016). The INTERSPEECH 2016 Computational Paralinguistics Challenge: Deception, Sincerity & Native Language. In *Interspeech 2016* (pp. 2001–2005). ISCA. doi:10.21437/Interspeech.2016-129

Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics*, *25*(3), 255–286. doi:10.1006/jpho.1997.0043

Schwartz, S. R., Cohen, S. M., Dailey, S. H., Rosenfeld, R. M., Deutsch, E. S., Gillespie, M. B., Granieri, E., Hapner, E. R., Kimball, C. E., Krouse, H. J., McMurray, J. S., Medina, S., O'Brien, K., Ouellette, D. R., Messinger-Rapport, B. J., Stachler, R. J., Strode, S., Thompson, D. M., Stemple, J. C., Willging, J. P., Cowley, T., McCoy, S., Bernad, P. G., & Patel, M. M. (2009). Clinical Practice

Guideline: Hoarseness (Dysphonia). *Otolaryngology–Head and Neck Surgery*, *141*(S1), 1–31. doi:10.1016/j.otohns.2009.06.744
Scobbie, J. M., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (1996). Covert contrast as a stage in the acquisition of phonetics and phonology. *QMC Working Papers in Speech and Language Sciences*, *1*, 43–62.
Scobbie, J. M., Punnoose, R., & Khattab, G. (2013). Articulating five liquids: A single speaker ultrasound study of Malayalam. In L. Spreafico & A. Vietti (Eds.), *Rhotics. New Data and Perspectives* (pp. 99–124). Bozen-Bolzano University Press.
Shah, Z., Sawalha, J., Tasnim, M., Qi, S.-a., Stroulia, E., & Greiner, R. (2021). Learning Language and Acoustic Models for Identifying Alzheimer's Dementia From Speech. *Frontiers in Computer Science*, *3*.
Shaw, J. A., Best, C. T., Mulak, K. E., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., Al-Tamimi, J., Mair, K., & Peek, M. (2014). Effects of short-term exposure to unfamiliar regional accents: Australians' categorization of London and Yorkshire English consonants. In *Proceedings of the 15th Australasian International Speech Science and Technology conference.* (pp. 72–75). Christchurch, New Zealand (6-9 December 2014).
Shimizu, K. (1989). A Cross-Language Study of Voicing Contrasts of Stops. *STUDIA PHONOLOGICA XXIII*.
Simpson, A. P. (2012). The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics*, *40*(3), 477–490. doi:10.1016/j.wocn.2012.02.001
Slifka, J. (2003). Tense/lax vowel classification using spectral dynamic cues. In *Proc. of the 15th ICPhS* (pp. 921–924).
Steffman, J. A. (2023). Vowel-initial glottalization as a prominence cue in speech perception and online processing. *Laboratory Phonology*, *14*(1). doi:10.16995/labphon.8753
Stevens, K. N. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, *34(4)*, 264–279. doi:10.1159/000259885
Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*(1-2), 3–45. doi:10.1016/S0095-4470(19)31520-7
Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). MIT press.
Stevens, K. N., & Keyser, S. J. (1989). Primary Features and Their Enhancement in Consonants. *Language*, *65(1)*, 81–106. doi:10.2307/414843
Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, *38*(1), 10–19. doi:10.1016/j.wocn.2008.10.004
Story, B. (2019). The Vocal Tract in Singing. In *The Oxford Handbook of Singing* (pp. 144–166). Oxford University Press. doi:10.1093/oxfordhb/9780199660773.013.012
Strange, W. (1989a). Dynamic specification of coarticulated vowels spoken in sentence context. *The Journal of the Acoustical Society of America*, *85*(5), 2135–2153. doi:10.1121/1.397863
Strange, W. (1989b). Evolving theories of vowel perception. *The Journal of the Acoustical Society of America*, *85*(5), 2081–2087. doi:10.1121/1.397860
Strange, W., & Jenkins, J. J. (2013). Dynamic Specification of Coarticulated Vowels. In G. S. Morrison & P. F. Assmann (Eds.), *Vowel Inherent Spectral Change* (pp. 87–115). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-14209-3_5
Strange, W., Jenkins, J. J., & Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *The Journal of the Acoustical Society of America*, *74*(3), 695–705. doi:10.1121/1.389855
Strange, W., Verbrugge, R. R., Shankweiler, D. P., & Edman, T. R. (1976). Consonant environment specifies vowel identity. *The Journal of the Acoustical Society of America*, *60*(1), 213–224. doi:10.1121/1.381066
Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, *9*(1), 307. doi:10.1186/1471-2105-9-307
Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*(1), 25. doi:10.1186/1471-2105-8-25
Strobl, C., Hothorn, T., & Zeileis, A. (2009). Party On! a New, Conditional Variable-importance Measure for Random Forests Available in the Party Package. *The R Journal, Volume 1/2*, 14–17.
Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*(4), 323–348. doi:10.1037/a0016973
Suleiman, S. M. (1985). *Jordanian Arabic between Diglossia and Bilingualism: Linguistic analysis* (Vol. VI:8). Amsterdam: John Benjamins Publishing Company. doi:10.1075/pb.vi.8
Sundberg, J., & Askenfelt, A. (1981). Larynx height and voice source. A relationship? In *Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology* (Vol. 22(2-3), pp. 23–36).
Sussman, H. M., Hoemeke, K., & Ahmed, F. (1993). A cross-linguistic investigation of locus equations as a relationally invariant descriptor for place of articulation. *Journal of the Acoustical Society of America*, *94*(3), 1256–1268. doi:10.1121/1.408178
Sussman, H. M., Hoemeke, K. A., & McCaffrey, H. A. (1992). Locus Equations As an Index of Coarticulation for Place of Articulation Distinctions in Children. *Journal of Speech, Language, and Hearing Research*, *35*(4), 769–781. doi:10.1044/jshr.3504.769
Sussman, H. M., McCaffrey, H. A., & Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *The Journal of the Acoustical Society of America*, *90*(3), 1309–1325. doi:10.1121/1.401923
Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better Decisions through Science. *Scientific American*, *283*(4), 82–87. doi:10.1038/scientificamerican1000-82
Sylak, J. (2011). Pharyngealization in Chechen is gutturalization. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 37(2), pp. 81–95).
Sylak-Glassman, J. (2014a). *Deriving Natural Classes: The Phonology and Typology of Post-Velar Consonants* (Unpublished doctoral dissertation). University of California, Berkeley.
Sylak-Glassman, J. (2014b). An Emergent Approach to the Guttural Natural Class. *Proceedings of the Annual Meetings on Phonology*, 1–12. doi:10.3765/amp.v1i1.44
Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English

vowels. *Journal of the Acoustical Society of America*, *79*(4), 1086–1100. doi: 10.1121/1.393381
Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2020). Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. *PLoS ONE*, *15*(7), e0236009. doi: 10.1371/journal.pone.0236009
Thomson, R. I. (2019). High Variability [Pronunciation] Training (HVPT): A proven technique about which every language teacher and learner ought to know:. *Journal of Second Language Pronunciation*, *4*(2), 208–231. doi: 10.1075/jslp.17038.tho
Titze, I. R., & Story, B. H. (1997). Acoustic interactions of the voice source with the lower vocal tract. *Journal of the Acoustical Society of America*, *101*(4), 2234–2243. doi: 10.1121/1.418246
Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *The Journal of the Acoustical Society of America*, *69*(5), 1465–1475. doi: 10.1121/1.385780
Turton, D., Khattab, G., Alsharif, B., & Al-Tamimi, J. (2020). Syllabic and emphatic conditioning of /l/ in Levantine Arabic: An auditory, acoustic and articulatory analysis. In *Proceedings of BAAP (The British Association of Academic Phoneticians).* York University, York, UK (1-3 April 2020; postponed due to Covid).
Vaissière, J. (2011). On the acoustic and perceptual characterization of reference vowels in a cross-language perspective. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)* (p. 8). Hong Kong, China (17-21 August 2011).
van Dyck, C. H., Swanson, C. J., Aisen, P., Bateman, R. J., Chen, C., Gee, M., Kanekiyo, M., Li, D., Reyderman, L., Cohen, S., Froelich, L., Katayama, S., Sabbagh, M., Vellas, B., Watson, D., Dhadda, S., Irizarry, M., Kramer, L. D., & Iwatsubo, T. (2022). Lecanemab in Early Alzheimer's Disease. *New England Journal of Medicine*, NEJMoa2212948. doi: 10.1056/NEJMoa2212948
Vaux, B. (1999). A note on pharyngeal features. *Harvard Working Papers in Linguistics*, *7*, 39–63.
Versteegh, K. (2001). *The Arabic Language*. Edinburgh University Press.
Vihman, M. M., & McCune, L. (1994). When is a word a word? *Journal of Child Language*, *21*(3), 517–542. doi: 10.1017/S0305000900009442
Vijayan, K., & Murty, K. S. R. (2014). Epoch extraction from allpass residual of speech signals. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1493–1497). doi: 10.1109/ICASSP.2014.6853846
Watson, J. C. E. (2007). *The phonology and morphology of Arabic* (1. paperback ed ed.). Oxford: Oxford University Press.
Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *The Journal of the Acoustical Society of America*, *73*(4), 1322–1336. doi: 10.1121/1.389236
Woo, P., Casper, J., Colton, R., & Brewer, D. (1992). Dysphonia in the Aging: Physiology Versus Disease. *The Laryngoscope*, *102*(2), 139–144. doi: 10.1288/00005537-199202000-00007
Wood, S. (1986). The acoustical significance of tongue, lip, and larynx maneuvers in rounded palatal vowels. *The Journal of the Acoustical Society of America*, *80*(2), 391–401. doi: 10.1121/1.394090
Wrench, A., Cleland, J., & Scobbie, J. M. (2011). An ultrasound protocol for comparing tongue contours: Upright vs. supine. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 2161–2164). Hong Kong, China (17-21 August 2011).
Wrench, A. A. (2018). Articulate Assistant Advanced User Guide (Version 2.17). *Edinburgh: Articulate Instruments Ltd*.
Yanagihara, N. (1967). Hoarseness: Investigation of the physiological mechanisms. *The Annals of Otology, Rhinology, and Laryngology*, *76*(2), 472–488. doi: 10.1177/000348946707600215
Yeou, M. (1997). Locus Equations and the Degree of Coarticulation of Arabic Consonants. *Phonetica*, *54*(3-4), 187–202. doi: 10.1159/000262221
Zawaydeh, B. (1999). *The phonetics and phonology of gutturals in Arabic* (Unpublished doctoral dissertation). PhD dissertation, Bloomington, IN: Indiana University.
Zawaydeh, B., & de Jong, K. (2011). The phonetics of localising uvularization in Arabic. In B. Heselwood & Z. Hassan (Eds.), *Instrumental studies in Arabic phonetics* (pp. 257–276). John Benjamins. doi: 10.1075/cilt.319
Zellou, G. (2017a). Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics*, *61*, 13–29. doi: 10.1016/j.wocn.2016.12.002
Zellou, G. (2017b). Specificity for coarticulatory vowel nasality in lexical representations. *The Journal of the Acoustical Society of America*, *142*(4), EL375-EL380. doi: 10.1121/1.5006817
Zeroual, C. (2003). Aerodynamic Study of Moroccan Arabic Guttural Consonants. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1859–1862).
Zeroual, C., & Clements, G. N. (2015). The feature [pharyngeal]. In A. Rialland, R. Ridouane, & H. van der Hulst (Eds.), *Features in Phonology and Phonetics: Posthumous Writings by George N. Clements and Coauthors* (pp. 247–276). Mouton de Gruyter: Berlin.
Zeroual, C., Esling, J. H., & Hoole, P. (2011). EMA, endoscopic, ultrasound and acoustic study of two secondary articulations in Moroccan Arabic. In B. Heselwood & Z. Hassan (Eds.), *Instrumental studies in Arabic phonetics* (pp. 277–297). John Benjamins. doi: 10.1075/cilt.319
Zhou, Z. (2023). *What makes an utterance ironic? Perception of Irony by L2 Learners of French* (Unpublished master's thesis). Université Paris Cité.
Zhou, Z. (In preparation). *Phonological transfer in L3 acquisition of ironic tone of voice: An acoustic and perceptual study of Mandarin-English bilinguals learning French as a third language* (Unpublished doctoral dissertation). Université Paris Cité, Paris, France.