

Studies in Arabic Linguistics

Perspectives on Arabic Linguistics XXXII

Edited by Elly van Gelderen

9

John Benjamins Publishing Company

Perspectives on Arabic Linguistics XXXII

Studies in Arabic Linguistics

This book series aims to publish original research in all fields of Arabic linguistics, including – but not limited to – theoretical linguistics, historical linguistics, sociolinguistics, pragmatics, typology, and language acquisition. Submissions from all current theoretical frameworks are welcome. Studies may deal with one or more varieties of Arabic, or Arabic in relation to or compared with other languages. Both monographs and thematic collections of research papers will be considered.

The series includes monographs and thematically coherent collective volumes, in English.

For an overview of all books published in this series, please see
benjamins.com/catalog/sal

Editors

Elabbas Benmamoun
Duke University

Enam Al-Wer
University of Essex

Editorial Board

Mahasen Hasan Abu-Mansour
Umm Al-Qura University

Sami Boudelaa
United Arab Emirates University

Stuart Davis
Indiana University

Mushira Eid
University of Utah

Clive Holes
The Oriental Institute, Oxford

Jean Lowenstamm
CNRS-Université Paris 7

Mustafa A. Mughazy
Western Michigan University

Jamal Ouhalla
University College Dublin

Jonathan Owens
University of Bayreuth

Janet C.E. Watson
University of Leeds

Manfred Woidich
University of Amsterdam

Volume 9

Perspectives on Arabic Linguistics XXXII. Papers selected from the Annual Symposium on Arabic Linguistics, Tempe, Arizona, 2018

Edited by Elly van Gelderen

Perspectives on Arabic Linguistics XXXII

Papers selected from the Annual Symposium
on Arabic Linguistics, Tempe, Arizona, 2018

Edited by

Elly van Gelderen

Arizona State University

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/sal.9

Cataloging-in-Publication Data available from Library of Congress.

ISBN 978 90 272 0759 3 (HB)

ISBN 978 90 272 6070 3 (E-BOOK)

© 2020 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · www.benjamins.com

Table of contents

| | |
|--|-----|
| Introduction <i>Elly van Gelderen</i> | 1 |
| Part I. Phonetics and phonology | |
| Stress assignment and foot construction in a Southwestern Saudi Arabic dialect <i>Musa Alahmari</i> | 11 |
| What triggers <i>'imāla</i> : Focus on a Palestinian variety with phonological analysis <i>Noor Abo Mokh and Stuart Davis</i> | 35 |
| A study of the place of articulation of the Arabic voiceless dorsal fricative <i>Noor Abo Mokh, Steven M. Lulich, Abdullah Alfaifi, Sarah Robinson, Sherman Charles and Kenneth De Jong</i> | 55 |
| Part II. Syntax and semantics | |
| On the syntax of correlation: Evidence from Egyptian Arabic <i>Usama Soltan</i> | 71 |
| The complementizer layer in Standard Arabic revisited <i>Salem Albuhayri and Hamid Ouali</i> | 109 |
| Sluicing and sprouting in Jordanian Arabic <i>Juman Al Bukhari</i> | 135 |
| Part III. Clinical linguistics | |
| Clinical linguistic research in the study of Arabic diglossia <i>Reem Khamis-Dakwar</i> | 155 |
| Index | 173 |

Introduction

Elly van Gelderen

Arizona State University

Significant recent additions to Arabic linguistics and applied linguistics have been *The Oxford Handbook of Arabic Linguistics* by Owens in 2013, *The Routledge Handbook of Arabic Linguistics* by Benmamoun & Bassiouney in 2017, and *The Routledge Handbook of Arabic Second Language Acquisition* by Alhawary in 2018. Previous volumes in the Studies in Arabic Linguistics' series have also provided excellent introductions to the traditions of Arabic linguistics and to current trends, e.g. Broselow & Ouali (2011). The present volume adds to this knowledge.

This volume includes work presented at the Annual Symposium on Arabic Linguistics 32 (ASAL), which was held at Arizona State University from 23–25 February 2018. The goal of the symposium is to bring together linguists to share their recent work on Arabic linguistics and applied linguistics, to discuss current issues in the Arabic language and its many varieties, and to contribute to the field of linguistics. Three plenaries, four posters, and 23 papers were presented, of which seven contributions appear here. The two co-chairs of ASAL 32 were graduate students Mekhlid Alsaeedi and Mohammed AlMutlaq with Elly van Gelderen as the faculty coordinator. Other students and faculty involved were Ibrahim Alaswad, Haroon Alsager, Mansour Altamimi, Noor Alzubaidi, Jaber Harthi, William Kruger, Neimeh Mousa, Kathryn Pruitt, Umar Sulayman, and Johanna Wood.

In Section 1 of this introduction, a few trends are discussed that complement the background provided in earlier introductions. Section 2 reviews the contributions in the volume.

1. Contributions of Arabic to typology and theoretical linguistics

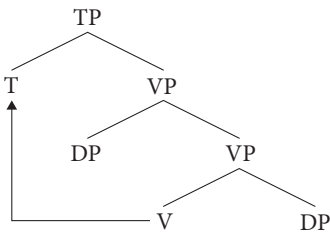
Standard Arabic and varieties of Arabic are fascinating both from an empirical and theoretical perspective. Their consonantal inventories are large (for instance, Standard Arabic has 28, cf. Embarki 2013: 27); they include a number of uvular, pharyngeal(ized), and glottal sounds and fricatives occur in seven different places of articulation. This rich inventory of consonants is in stark contrast with the vowels

where three basic phonemes occur in Standard Arabic, namely /i, a, u/, with their long counterparts /i:, a:, u:/.

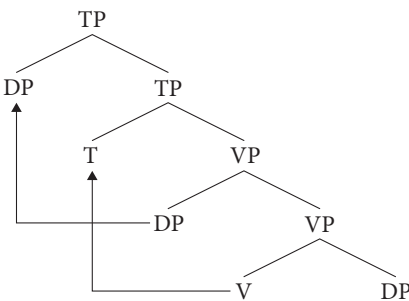
On another level, the consonantal root and vocalic patterns have challenged the representation of the word (e.g. Goldsmith 1990) and, initially, this pattern was used to account for voice alternations. However, since the 1990s, many morphologists have turned from the root and pattern analysis to a word-based morphology (e.g. Ratcliffe 1997; Ussishkin 2003). The traditional analysis resulted in a very complex representation of the verbal conjugation system and of plurals whereas a word-based model produces a simpler derivation of verbal aspect, voice, and nominal plurals.

The word order of Standard Arabic shows SVO and VSO alternations where the various dialects have SVO. The dialects have also lost the nominal case endings. Many debates have raged as to which order is basic, SVO or VSO. For both SVO and VSO, generative approaches assume that the V(erb) moves to a higher position but they are not completely clear which position this is. The subject in VSO structures is assumed to stay in the VP, as in (1a), but the subject in SVO structures either (a) moves to the Specifier of phrase that the V moves to, as in (1b), where I use T for the higher position but this is not crucial, or (b) is basegenerated in a higher position (Soltan 2007).

(1) a.



b.



Current cartographic frameworks have stimulated a more articulated left periphery for Arabic (Shlonsky 2000; Al Sager 2017; Albuhayri & Ouali, this volume) and that periphery is relevant for where the DP lands or is base generated. Unfortunately, Arabic lacks a lot of sentence-internal adverbs to check this cartography.

The two orders also result in different agreement patterns (Mohammad 2000) with the V in VSO orders agreeing in person and gender and in SVO orders in person, number, and gender. This fact has given rise to many analyses, from different checking mechanisms for Spec-Head structures and agree structures (Koopman and Sportiche 1991) to empty expletives in VSO orders (Mohammad 2000; van Gelderen 1997) to a basegenerated DP and pro (Soltan 2007: 68). Using current theories, the richer agreement pattern in the SVO order may provide evidence for phase transfer (Chomsky 2015) and special solutions for labelling paradoxes, the $\langle \phi, \phi \rangle$ labeling. However, in this Problems of Projections' approach, VSO remains a riddle because the subject in (1a) should not be allowed to stay in the VP because this results in {DP, VP} structure, which cannot be labeled.

Arabic varieties can be compared and then provide more typological and historical insights. For instance, negatives and verbal and pronominal copulas show significant variation across regions and these differences have provided evidence for the Jespersen and Croft Cycles (Lucas 2007; van Gelderen 2011; Alsaedi 2019; Alqassas 2019). These cycles show that negation is crucial in every language and can be reinforced by additional lexical material and that copulas are optional. Earlier ASAL Proceedings, e.g. Eid & Parkinson (1996), included work on variation in negation in dialects of Yemen, Egyptian and Moroccan Arabic. Negation in Arabic has also provided insight on the position of the NegP, e.g. in a relatively high position (Fassi Fehri 1993), and on the morphological status (phrase or head) of the various negatives: *ma*, *-sh*, *l-* (Benmamoun 1996). Copulas interact with negatives and vary widely across Arabic. They are still important in the debate if there is an empty verb (e.g. Jelinek 1981) or no verb (see Benmamoun 2000).

Another area where typology and historical linguistics aid synchronic analysis is on the distinction between arguments and agreement. What is the status of pronominal subjects and objects? In (2), from Egyptian Arabic, the object pronoun and an argument nominal are in complementary distribution showing the pronoun is not an agreement marker. In Standard Arabic, the marking in (3) is obligatory, indicating the marking on the verb is agreement.

- (2) a. *šuft-uh* Egyptian Arabic
 saw.1s-him
 'I saw him.'
- b. *šuft il-walad*
 saw.1s the-boy
 'I saw the boy.' (Jelinek 1989: 121)
- (3) a. *šuft-ik ?inti* Standard Arabic
 saw.1s-2FS you.2FS

- b. *šuft ?inti
 saw.1s you
 ‘I saw you.’

(Shlonsky 1997: 197)

Change of a pronoun to a clitic to an agreement marker is common so varieties of Arabic are expected to differ slightly. Other grammaticalizations similarly ‘explain’ inter-Arabic variation; see e.g. Esseesy (2010) for prepositions and conjunctions.

Arabic linguistics is thriving as befits a language that may be the fourth largest in the world (<https://www.ethnologue.com/statistics/size>). Owens (2013: 2) puts the number of L1 speakers as 300 million and that of L2 speakers as 60 million. Its phonological, morphological, syntactic, and semantic structure remains interesting to many theoretical frameworks and the sociolinguistic situation is evolving, with new pidgins/creoles emerging (Tosco & Manfredi 2013) and increased study of regional variants (Brustad 2000). A recent volume edited by Manfredi & Tosco (2018) adds to the discussion on Arabic in contact. From a pedagogical point of view, the study of L2 Arabic and heritage speakers is important and warrants further study. New (social) media have changed the methodologies used in studying (mainly) pragmatic uses. This volume brings together some of these broader themes.

2. The contributions in this volume

Part I starts the volume with three papers on phonetics and phonology. All three focus on regional varieties, namely on Saudi Arabian, on Palestinian Arabic, and on a selection of six varieties. These varieties differ from Modern Standard Arabic in stress assignment and directionality of foot construction, the presence of *‘imāla*, and the place of articulation of the dorsal fricative. The frameworks and approaches used are Metrical Stress Theory, Optimality Theory, Feature Geometry, and Feature Analysis.

Stress in Arabic is sensitive to syllable weight. **Alahmari** investigates how stress is assigned under the influence of geminates and long vowels in a Southwestern Saudi Arabian Arabic variety as well as how feet are constructed in this variety. Stress shows a moraic trochee binary system where ternary effects arise as a result of Weak Local Parsing. The paper also considers stress assignment influenced by geminates and long vowels to be prominence-based: stress goes to the heaviest syllable using a scalar quantity-sensitive system. The paper shows how other phonological aspects interact with stress, e.g. high vowel deletion and the quality of word-final vowels.

Abo Mokh & Davis examine *‘imāla*, a term used to describe the raising of the low vowel /a/ to [e] or [i], a process typical of Levantine varieties of Arabic. The paper discusses the phonological features of the consonants that block *‘imāla*,

the semantic classes of adjectives describing color ('black', 'green' etc) and mental/physical challenges ('toothless', 'clumsy') that interact with the phenomenon, and the absence of *'imāla* in loanwords in a rural Palestinian variety. Short nouns whose consonantal roots start with an unpronounced /w/ also fail to undergo *'imāla*, a novel observation.

Abo Mokh, Lulich, Alfaifi, Robinson, Charles, & De Jong study the place of articulation of the voiceless dorsal fricative in six varieties of Arabic (i.e. in Faifi, Algerian, Cairene, Syrian, Moroccan, and Palestinian). Experiments using ultrasound show that most speakers have a posterior articulation similar to uvulars and that no speaker has a more anterior, velar articulation. This has implications for their phonological representation.

Part II consists of three papers on Syntax and Semantics. Again, two of the three examine regional varieties, namely Egyptian and Jordanian. All papers use a generative approach with the first two employing a cartographic framework, in which the CP layer is expanded. The third paper examines ellipsis and assumes a rich structure that is deleted in sluicing and sprouting.

Soltan provides evidence for two types of correlatives in Egyptian Arabic. He argues that the syntax of correlatives can be understood in a compositional manner rather than as constructions. Their characteristics follow from general principles of grammar, such as variable binding, parallelism, and cartography, and from properties of the relevant heads, for instance, *kull* and *mā*.

Albuhayri & Ouali examine three aspects in the complementizer layer (CP) in Standard Arabic. First, they scrutinize the positional differences between *ʔinna* and *ʔanna* and argue that *ʔinna* is a verum operator heading a VerumP in the CP-layer. Second, they look at the agreement elements and argue they are pronominal. Finally, they show that the ban on extraction out of a subject in SVO structures indicates a higher position for the subject DP, in accordance with approaches such as Soltan (2007), mentioned above.

Al Bukhari discusses elliptical constructions in Jordanian Arabic, a variety not yet investigated in this regard. Sluicing involves the ellipsis of a clause after a *wh*-word that has moved from this clause and sprouting is a subtype thereof. The paper investigates a variety of constructions with the *wh*-word an argument or not and a PP or not in Jordanian and contributes to the theoretical debate on the best analysis of sluicing.

Part III contains one paper, on clinical linguistics. The paper concerns the issue of diglossia and includes a call to action.

Khamis-Dakwar provides an overview of the Speech Language Pathology profession in the Arabic-speaking world. This paper advocates for more work and programs dedicated to clinical linguistics in the Arabic speaking world. First, there are no comprehensive data on the availability of Arabic Language Pathologists.

Secondly, the data that are available are not adequate. This overview is followed by descriptions of preliminary studies focussing on diglossic situations. One study looks at aphasia recovery in diglossic speech communities and the other on diglossic codeswitching in Arabic-speaking individuals with Autism Spectrum Disorder.

3. Note on reviewing, the transcriptions, glosses, and examples

For help reviewing the chapters, the volume editor would like to thank Ali Abdel-Khalig, Ahmad Alqassas, Sara Al Tubuly, Haroon Al Sager, Elabbas Benmamoun, Robert Borsley, Naomi Danton, Aydin Durgunoglu, Rania Habib, Youssef Haddad, Lutfi Hussein, Tommy Leung, Narin Loa, Roy Major, Sam Rosenthal, Ur Shlonsky, Usama Soltan, Janet Watson, and Islam Youssef.

Because the papers represent different areas in linguistics, the examples are presented in different ways. The papers on phonetics, phonology, and prosody use IPA for the examples; the papers on syntax use appropriate morphological glossing, where relevant.

References

- Alqassas, Ahmad. 2019. *A Multi-locus Analysis of Arabic Negation*. Edinburgh: Edinburgh University Press. <https://doi.org/10.3366/edinburgh/9781474433143.001.0001>
- Al Sager, Haroon. 2017. Phasehood of *Wh*-Questions in Modern Standard Arabic. ASU PhD.
- Alsaeedi, Mekhlid. 2019. Negative Existential and Existential Copula cycles in Arabic. In Ljuba Veselinova & Arja Hamari (eds), *The Negative Existential Cycle from a historical-comparative perspective*. Language Science Press.
- Benmamoun, Elabbas. 1996. Negative Polarity and Presupposition in Arabic. In Mushira Eid (eds), *Perspectives on Arabic Linguistics VIII: Papers from the Third Annual Symposium on Arabic Linguistics*, 47–66. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.134.06ben>
- Benmamoun, Elabbas. 2000. *The feature structure of functional categories: a comparative study of Arabic Dialects*. Oxford: Oxford University Press.
- Broselow, Ellen & Hamid Ouali. 2011. *Perspectives on Arabic Linguistics: Papers from the Annual Symposia 2008 and 2009*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.317>
- Brustad, Kristen. 2000. *The Syntax of Spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington: Georgetown University Press.
- Chomsky, Noam. 2015. Problems of Projection: Extensions. In: *Structures, Strategies, and Beyond*, ed. Elisa Di Domenico et al., 3–16. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.223.01cho>
- Eid, Mushira and Dilworth Parkinson (eds.). 1996. *Perspectives on Arabic Linguistics IX*. Amsterdam: John Benjamins.
- Embarki, Mohamed. 2013. Phonetics. In: Jonathan Owens (ed), *The Oxford Handbook of Arabic Linguistics*, 23–44. Oxford: Oxford University Press.

- Esseesy, Mohssen. 2010. *Grammaticalization of Arabic Prepositions and Subordinators*. Leiden: Brill. <https://doi.org/10.1163/9789004187634>
- Fassi Fehri, Abdelkader. 1993. *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.
- Gelderen, Elly van. 1997. *Verbal Agreement and the Grammar of its Breakdown*. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110959994>
- Gelderen, Elly van. 2011. *The Linguistic Cycle*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199756056.001.0001>
- Goldsmith, John. 1990. *Autosegmental & Metrical Phonology*. Oxford: Blackwell.
- Jelinek, Eloise. 1981. *On defining categories*. PhD University of Arizona.
- Jelinek, Eloise. 1989. The Case Split in Choctaw. In Laszlo Maracz & Pieter Muysken (eds), *Configurationality*, 117–141. Dordrecht: Foris. <https://doi.org/10.1515/9783110884883-007>
- Koopman, Hilda & Dominique Sportiche. 1991. The Position of Subjects. *Lingua* 85.2/3: 211–258. [https://doi.org/10.1016/0024-3841\(91\)90022-W](https://doi.org/10.1016/0024-3841(91)90022-W)
- Lucas, Chris. 2007. Jespersen's Cycle in Arabic and Berber. *Transactions of the Philological Society* 105.3: 398–431. <https://doi.org/10.1111/j.1467-968X.2007.00189.x>
- Manfredi, Stefano & Mauro Tosco. 2018. *Arabic in contact*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sal.6>
- Mohammad, Mohammad. 2000. *Word Order, Agreement, and Pronominalization in Standard and Palestinian Arabic*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.181>
- Owens, Jonathan. 2013. *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199764136.001.0001>
- Ratcliffe, Robert. 1997. Prosodic Templates in a Word-Based Morphological Analysis of Arabic. In Mushira Eid & Robert Ratcliffe (eds.), *Perspectives on Arabic Linguistics X*, 147–171. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.153.10rat>
- Shlonsky, Ur. 1997. *Clause Structure and Word Order in Hebrew and Arabic*. Oxford: Oxford University Press.
- Shlonsky, Ur. 2000. Remarks on the complementizer layer of Standard Arabic. In Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.), *Research in Afroasiatic grammar: Papers from the third conference on Afroasiatic languages*, 325–343. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.202.15shl>
- Soltan, Usama. 2007. *On formal feature licensing in minimalism: Aspects of Standard Arabic morphosyntax*. University of Maryland PhD.
- Tosco, Mauro & Stefano Manfredi. 2013. Pidgins and Creoles. In: Jonathan Owens (ed), *The Oxford Handbook of Arabic Linguistics*, 495–519. Oxford: Oxford University Press.
- Ussishkin, Adam. 2003. Templatic effects as fixed prosody: the verbal system in Semitic. In Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky (eds), *Research in Afroasiatic Grammar III*, 511–530. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.241.23uss>

PART I

Phonetics and phonology

Stress assignment and foot construction in a Southwestern Saudi Arabic dialect

Musa Alahmari
King Khalid University

This paper presents and analyzes some aspects of the stress system of a Southwestern Saudi Arabic dialect (SSA). In an Optimality-Theoretic framework, the analysis focuses on issues related to how feet are constructed in this dialect as well as how stress is assigned in the language under the effect of geminates and long vowels. In the presented analysis, I argue that the stress pattern of SSA is a moraic trochee binary stress system in which ternary effects arise as a result of Weak Local Parsing that is responsible for the pre-antepenultimate stress in the language (Hayes, 1995). I also argue that the analysis of stress assignment under the effect of geminates and long vowels is prominence-based in such a way that stress is assigned to the heaviest syllable in the word in a scalar quantity-sensitivity stress system that gives priority to syllables that are inherently bimoraic. The analysis seeks a unified account that interacts with other aspects of the phonology in the investigated dialect such as high vowel deletion and the quality of word-final vowels.

Keywords: optimality theory, moraic theory, metrical theory, quantity, prominence

1. Introduction

The study of word stress patterns has played an important role in the developments of metrical phonology. Languages with various properties of word stress patterns have provided evidence for typological preferences in stress assignment and foot structure, of which Arabic dialects have been a focus of substantial metrical stress research (McCarthy, 1979; Kenstowicz, 1980; Kabrah, 2004; Watson, 2002, 2011; Bamakhramah, 2009; Kager, 2009). It has been shown that stress assignment in Arabic is sensitive to syllable weight, where stress is usually assigned to the heavy syllable in the word. Nonetheless, Arabic dialects do differ in other aspects of stress assignment and foot structure. One of the major differences pertains to how far

away from the right edge of the word the word stress can be realized. That is, what is the extent of the stress window from the right edge of the word? In three syllable window dialects, such as Cairene and Meccan Arabic dialects, stress does not retract beyond the antepenultimate syllable while in other dialects such as Palestinian and San'ani Arabic dialects, stress can retract to the pre-antepenultimate syllable (Hayes, 1995; Watson, 2011). In addition, dialects differ in the directionality of foot construction. In some dialects, feet are constructed from left to right while in others feet are constructed from right to left. Dialects also differ as to whether they allow or prohibit degenerate feet, with the prohibition of degenerate feet being strong in some dialects while weak in others (Hayes, 1995).

In this paper, I aim to present and analyze the stress pattern in a Southwestern Saudi Arabic dialect (SSA), an understudied Arabic dialect spoken in Saudi Arabia, in the area of Tihama (Northern Tihama), south of the Makkah region, east of the Hali valley (Alahmari, 2018). The paper specifically aims to explore interesting aspects of stress assignment and foot construction in this dialect such as pre-antepenultimate stress, geminate and long vowel effects on stress assignment, foot extrametricality, and the ban on degenerate feet. In addition, the paper deals with the issues of high vowel syncope and the quality of word-final vowels. Based on the principles of Metrical Stress Theory (Hayes, 1995), I argue that pre-antepenultimate stress in the language can be accounted for under Weak Local Parsing. Not only can it account for pre-antepenultimate stress in four light syllable words such as ['ra.ga.ba.ti] 'my neck', it can also account for words with a pre-antepenultimate stressed non-final closed CVC syllable such as ['mak.ta.ba.ti] 'my library'. In particular, I show that weak local parsing leaves behind unparsed light syllables in stress clash environments. Under this analysis, I also argue for the application of foot extrametricality and the strong ban on degenerate feet. For the effect of geminates and long vowels on stress assignment, I demonstrate that these effects follow from a scalar-quantity sensitivity stress system and that they can be accounted for in a prominence-based analysis (Davis, 1989; Hayes, 1995; Watson, 2002). It is shown that the proposed analysis cannot only account for the general stress pattern in the language, but it can also extend naturally to account for other related phenomena such as high vowel syncope and the quality of word-final vowels. To highlight the contributions of this paper to Arabic dialectology more generally, first, the paper documents a dialect where geminates and long vowels have priority for stress assignment over heavy CVC syllables. Second, it brings evidence for abstract foot construction in Arabic. Third, it provides an unprecedented analysis of the quality of word-final vowels.

This paper is structured as follows. In Section 2, the general stress pattern of the SSA dialect is presented with special reference to syllable quantity. In addition,

I discuss how stress is assigned based on syllable quantity, and how it shifts from one syllable to another. In Section 3, I present the proposed analysis of the general stress pattern in this dialect under the principles of Metrical Stress Theory (Hayes, 1995). Section 3 also deals with the quality of word-final vowels and the application of high vowel syncope in the dialect under investigation. Section 4 provides an optimality-theoretic analysis that accounts for the various aspects of the stress pattern in this dialect based on constraint interactions (Kager, 1999; Prince and Smolensky, 1993; McCarthy and Prince, 1993). Section 5 concludes this paper.

2. Stress pattern

The general stress pattern in SSA is predictable based on syllable quantity. The distribution of syllables in SSA ranges from light to heavy to superheavy syllables. In moraic theory (Hayes, 1989; Davis 2011a, 2011b), heavy syllables are bimoraic (two moras) while light syllables are monomoraic (one mora). Heavy syllables include CVV (syllables with a long vowel), CVG syllables (syllables with a geminate), and non-final CVC syllables. A non-final CVC syllable becomes heavy due to the constraint Weight-By-Position (Hayes, 1989). Light syllables include CV syllables or syllables with a short vowel and word-final CVC syllables. A word-final CVC syllable is monomoraic due to consonant extrametricality. Superheavy syllables such as CVVC and CVCC syllables are restricted to word-final position only. Thus, they are bimoraic syllables because of consonant extrametricality. The interactions between syllable quantity and stress assignment is discussed in more detail in Section (2.1).

2.1 Syllable quantity

In SSA, stress is always initial in words with only light syllables. In such forms, stress can retract to the pre-antepenultimate syllable as the following examples in (1) show.

- (1) Light syllables:
- a. [ʕa.fɑ] dinner
 - b. [ʕa.tab] he wrote
 - c. [ra.ga.ba] neck
 - d. [ra.ga.ba.ti] my neck

Heavy geminate and long vowel syllables (CVV/CVG) always attract stress. Such syllables attract stress regardless of their position in the word. Consider the following examples in (2).

- (2) Heavy CVV/CVG syllables:
- a. ['daw.wart] I searched
 - b. ['dar.ra.sa.tuh] she taught him
 - c. ['saa.fart] I travelled
 - d. ['saa.ʃa.da.tuh] she helped him

Non-final CVC syllables attract stress only in the absence of a geminate or a long vowel in the word. As such, non-final CVC syllables can attract stress up to the pre-antepenultimate. Consider the following examples in (3).

- (3) Non-final CVC syllables:
- a. ['mak.ta.ba] library
 - b. ['mak.ta.ba.ti] my library

Superheavy CVVC syllables always attract stress regardless of the presence of other heavy syllables in the word while superheavy CVCC syllables only attract stress from other non-final CVC syllables. Consider the following examples in (4).

- (4) Superheavy CVVC/CVCC syllables:
- a. [dʒii.'raan] neighbors
 - b. [kab.'kabt] I spilled

There is a crucial relation between syllable quantity and stress assignment in this dialect. On the basis of that several generalizations can be made. Heavy and super-heavy syllables seem to always attract stress. Of these heavy syllables, geminates and long vowels appear to attract stress regardless of their position in the word. For light syllables, stress retracts up to the pre-antepenultimate. Heavy non-final CVC syllables also attract stress up to the pre-antepenultimate syllable in the absence of a geminate or a long vowel.

2.2 Stress shift

The previous section makes it clear that a proper understanding of the general stress pattern in SSA can only become possible through reference to syllable quantity. Nonetheless, we need to consider other forms in which stress shifts from one syllable to another apart from syllable quantity. This occurs after morphological concatenation results in forms with more than four syllables. Consider now the following examples in which stress shifts from one syllable to another after morpheme concatenation.

- (5) Morpheme concatenation and stress shift:
- a. /mintabih-a/ ['min.ta.bi.ha] aware (fem.)
 - b. /mintabih-a-l-uh/ [min.ta.'bi.ha.luh] aware (fem.) of him
 - c. /maʕrifa-ti/ ['maʕ.rfa.ti] my knowledge
 - d. /maʕrifa-ti-l-uh/ [maʕ.'rfa.ti.luh] my knowledge of him

The forms in (5) are crucial in explaining stress shift in the dialect. In (5a), stress is assigned to the pre-antepenultimate with no stress shift after the feminine suffix marker /-a/ is attached to the stem /mintabih/ 'aware'. This means that suffixation does not necessarily result in stress shift. However, after morpheme concatenation in (5b), stress shifts to the antepenultimate. Similarly, stress is assigned to the pre-antepenultimate in (5c) (before high vowel syncope), while in (5d) stress shifts to the penultimate, after morpheme concatenation.

The stress pattern as observed in the presented forms is analyzed in the following section. Based on the principles of Metrical Stress Theory (Hayes, 1995), the analysis will take into account pre-antepenultimate stress in the language, stress shift in concatenated forms, and the effect of geminates and long vowels on stress assignment.

3. Analysis

The data presented in the previous section show that heavy and superheavy syllables attract stress predominantly. Heavy CVG/CVV syllables appear to always attract stress. Superheavy CVVC syllables always attract stress while CVCC syllables attract stress only if there is no geminate or a long vowel in the word. Heavy non-final CVC syllables are similar to superheavy CVCC syllables, suggesting a prosodic difference between types of syllables in the dialect. Light syllables attract stress either word initially in four-light syllable words or through stress shift after morpheme concatenation. Simply put, the generalization is that stress in SSA is attracted to geminates and long vowels. Otherwise, stress is attracted to the heavy syllable or the initial light syllable in the absence of other heavy syllables in the word up to the pre-antepenultimate. In what follows, I will argue for the following stress facts based on the observations drawn from the presented data. The general stress pattern in SSA clearly reflects a trochaic stress system with left to right direction of parsing. Feet are binary with a strong prohibition on degenerate feet. Stress is always non-final with End Rule Right (ERR).

Based on the principles of the metrical theory of stress, this stress system can be accounted for under Weak Local Parsing (WLP), the Peripherality Condition, the Priority Clause, and Persistent Footing (Hayes, 1995). WLP is part of the parsing

parameter proposed by Hayes (1995) in which he distinguishes between two types of parsing parameters: strong and weak.

- (6) Foot Parsing Locality Parameter (Hayes, 1995: 308):
- a. **Strong Local Parsing** When a foot has been constructed, align the window for further parsing at the next unfooted syllable. (unmarked value of the parameter).
 - b. **Weak Local Parsing** When a foot has been constructed, align the window for further parsing by skipping over /~/ [i.e. a light syllable], where possible. (marked value of the parameter).

Hayes (1995) points out that WLP is motivated by the tendency in languages to avoid stress clashes and stress lapses. He also notes that the major difference between the two parsing parameters is that the target of strong parsing is to achieve exhaustive parsing while the target of weak parsing is to keep feet apart minimally by skipping light syllables. Under WLP, a binary foot plus a light syllable will deem the light syllable unparsed. Parsing will resume then to parse other syllables into binary feet (Hayes, 1995). In the current analysis, I assume that foot parsing in SSA is weak local because it accounts for the general stress pattern in this dialect. It also reaches out to account for other related phenomenon in the language such as high vowel syncope and makes accurate predictions on which high vowels delete and which high vowels do not. As will be shown, WLP interacts with other principles of metrical theory.

An important principle to be implemented in this analysis pertains to foot extrametricality. It is invoked by the Peripherality Condition, which deems a peripheral element or constituent extrametrical (Hayes, 1995).

- (7) Peripherality Condition (Hayes, 1995: 57):
A constituent may be extrametrical only if it is at a designated edge (left or right) of its domain.

Evidence for foot extrametricality in this dialect comes from the fact that stress in the language never occurs on a foot aligned with the right edge of the word. Note that foot extrametricality is blocked by consonant extrametricality in forms such as word-final superheavy CVVC/CVCC syllables. It is also important to note that extrametricality itself is restricted by the Nonexhaustivity Condition. That is, foot extrametricality does not apply in forms such as CVCV syllable words due to the Nonexhaustivity Condition.

- (8) Nonexhaustivity (Hayes, 1995: 58):
An extrametricality rule is blocked if it would render the entire domain of the stress rules extrametrical.

The ban on degenerate feet in the language is evidenced through the fact that stress never occurs on what would be a unary foot, which indicates a strong ban on degenerate feet. Independent evidence for the ban on degenerate feet in SSA also comes from the Word-Minimality requirements, where prosodic words are required to be minimally bimoraic (McCarthy and Prince, 1993). In SSA, prosodic minimality requires words to be minimally disyllabic or bimoraic (Alahmari, 2018). As such, this constitutes independent evidence that degenerate feet are strongly banned in this dialect. However, while this will ban the formation of degenerate feet after a binary foot is constructed, a unary foot could also be created prior to a binary foot not only after it. In such a case, this is prohibited by the Priority Clause condition, which works as a prohibition on degenerate feet that might be formed prior to a binary foot. Taken all together, i.e. the ban on degenerate feet and the Priority Clause, this means that a degenerate foot is absolutely banned. Syllables that cannot be incorporated into a foot are left astray.

(9) Priority Clause (Hayes, 1995: 95):

If at any stage in foot parsing the portion of the string being scanned would yield a degenerate foot, the parse scans further along the string to construct a proper foot where possible.

While parsing could leave behind unparsed syllables, stray syllables that can form a binary foot are constructed into a foot. This property is expressed with the persistent footing mode, which simply means that metrical rules will apply persistently in order to achieve well-formedness (Hayes, 1995). Evidence for persistent footing in SSA comes from the stress clash environment, which blocks persistent footing that would apply in the elsewhere environment as will be demonstrated in the upcoming analysis. The core idea is that while footing is persistent, it is more important for the stress system in SSA to avoid a stress clash than to be persistent. This is further illustrated through the OT analysis that shows that this is only a matter of constraint ranking.

(10) Persistent Footing (Hayes, 1995: 115):

- a. Single stray syllables are adjoined to existing feet if the result is well formed.
- b. Otherwise, sequences of stray syllables may be converted into feet.

By taking these principles of metrical theory into account, we are not only able to account for the general stress pattern in the language, but we can also account for other related phenomena such as pre-antepenultimate stress and high vowel syncope. The actual application of this analysis is shown with three syllable words such as [ˈra.ga.ba] ‘neck’ and [ˈmak.ta.ba] ‘library’ as well as with four syllable words such as [ˈra.ga.ba.ti] ‘my neck’ and [ˈmak.ta.ba.ti] ‘my library’. The analysis will

The combined effect of weak local parsing, foot extrametricality, and the ban on degenerate feet is able to account for the general stress pattern in the language. Note that although we do not show the derivation for forms in which stress shift occurs, the presented analysis will naturally account for that. In what follows, I will discuss other interesting issues related to the general stress pattern in this dialect. More evidence for the application of foot extrametricality can be found in the analysis of the quality of word-final vowels in the language. More evidence for the weak local parsing analysis will be shown in the analysis of high vowel syncope. Challenging issues of geminates and long vowel effects will be taken into account as well.

3.1 Word-final vowels

As far as the quality of word-final vowels is concerned in stress assignment, I address the question of what is the quality of word-final vowels in the language and the issue of how they are analyzed in correlation to the general stress pattern in this dialect. First, since a word-final long vowel does not surface in the language and neither does it receive stress word-finally, I assume that word-final vowels in SSA are underlyingly short and that they lengthen after suffixation as opposed to McCarthy (2005). The quality of word-final vowels could also be analyzed as underlyingly long under a shortening analysis. Nonetheless, a lengthening analysis, as assumed here, has two major advantages. First, the analysis of word-final vowels as underlyingly short is more economical since it does not need to invoke extra processes to get the right stress assignment such as final long vowel extrametricality. Second, a lengthening analysis could follow from the tendency that stress is always non-final in the language as a NONFINALITY requirement. Thus, not only the lengthening analysis has to do with how stress is directly assigned, but is also more consistent with the observed tendency of the general stress pattern in this dialect where syllables with long vowels always attract stress. Under this analysis, word-final short vowels become long in a derived environment, where a consonant initial suffix induces lengthening of the immediately prior vowel with stress shifting onto that syllable as the forms in (13) show.

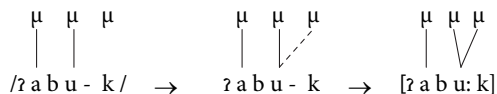
(13) Word-final vowel lengthening: CV# → CVV/SUFFIX:

- | | | | |
|----|---------------|------------------|------------------|
| a. | /ʔabu/ | [ʔa.bu] | father |
| b. | /ʔabu-k/ | [ʔa.'buuk] | your father |
| c. | /ʔakal-na/ | [ʔa.'kal.na] | we ate |
| d. | /ʔakal-na-ha/ | [ʔa.kal.'naa.ha] | we ate it |
| e. | /ʔakal-tu/ | [ʔa.'kal.tu] | you (pl.) ate |
| f. | /ʔakal-tu-ha/ | [ʔa.kal.'tuu.ha] | you (pl.) ate it |

In the forms in (13), stress is located on the initial syllable prior to suffixation. After suffixation, vowel lengthening takes place and stress shifts to the penultimate syllable. Note that stress in these forms never occurs as word-final since it is either located on a penultimate such as (13a) or before an extrametrical consonant such as (13b). This observation is more consistent with a vowel lengthening analysis rather than a vowel shortening one if it were analyzed as underlyingly long, but never attracts stress.

Under this assumption, lengthening is motivated through the floating mora analysis. The floating mora analysis comes in the spirit of the concept of mora augmentation as developed by Davis and Ueda (2002, 2006). Principally, it is assumed that in the studied dialect, consonant-initial suffixes contain a floating mora in their underlying forms such as that /-^μSUFFIX/. Thus, word-final vowel lengthening under the floating mora analysis occurs as a result of mora augmentation in the first place. This is indicated as the representation in (14) shows, where the suffix is shown to contain a floating mora in its underlying form, and through vowel lengthening it becomes realized in the output form.

(14) Suffixal floating mora & vowel lengthening:



In (14), it is assumed that the suffix /-k/ ‘your’ contains a floating mora in its underlying form. Prior to suffixation, stress is located on the penultimate syllable. After suffixation, the suffixal floating mora adjoins the word-final short vowel, resulting in vowel lengthening and stress shift. Hence, the floating mora analysis precisely accounts for the quality of word-final vowels and processes pertinent to that such as vowel lengthening. This analysis also independently accounts for other phenomena in the language such as vowel epenthesis and vowel shortening (Alahmari, 2018). However, these issues are beyond the scope of the present paper.

3.2 High vowel syncope

High vowel syncope in SSA interacts with the general stress pattern in the language. Generally speaking, syncope is applied if the target is a non-final unstressed high vowel (15a,b), but it is blocked when the target high vowel occurs in a stressed syllable (15c). This generalization, however, is unable to capture the lack of application of high vowel syncope in forms such as (15d) since the high vowel in such a form is not stressed and at the same time it does not delete.

- (15) High vowel syncope:
- | | | | | |
|----|------------|-------------|--------------|----------|
| a. | /mudiir/ | ['mɗiir] | *[mu.ɗiir] | manager |
| b. | /bilaad-i/ | ['blaa.di] | *[bi.laa.di] | my land |
| c. | /mudara/ | ['mu.da.ra] | *[mda.ra] | managers |
| d. | /kalima/ | ['ka.li.ma] | *[kal.ma] | word |

The presented analysis in this paper independently accounts for high vowel syncope in the language. Under this analysis, a high vowel that is parsed by a foot is immune to syncope while a high vowel that is left unparsed after foot construction is the one that is prone to delete. Consider the following examples in (16).

- (16) High vowel syncope under foot parsing:
- | | | | |
|----|---------------------|-----------------|---------------|
| a. | (‘mu.da).ra | ['mu.da.ra] | managers |
| b. | (‘ka.li).ma | ['ka.li.ma] | word |
| c. | mu.(‘ɗiir) | ['mɗiir] | manager |
| d. | (‘maʕ).ri.fa | ['maʕ.rfa] | knowledge |
| e. | (mih).ta.ri.(‘fiin) | [mih.tar.'fiin] | professionals |

In (16a), a stressed high vowel occurs in a footed syllable, thus it is retained in the output form. In (16b), an unstressed high vowel occurs in a footed syllable, and it is preserved in the output form. In (16c), a high vowel occurs in an unfooted syllable that is left unparsed by the Priority Clause Condition, and thus it is deleted in the output form. In (16d), a high vowel occurs in a syllable that is left unparsed by WLP since it occurs directly after a binary foot is constructed. Thus, the high vowel in such a form is deleted in the output form. In (16e), a high vowel occurs in an unfooted syllable, and thus it is deleted in the output form. Note that in (16e), foot extrametricality is blocked by consonant extrametricality. Also note that two medial light syllables trapped between two binary feet are left unparsed, i.e. *ta.ri*. The first light syllable *ta* is skipped by WLP while the second light syllable *ri* is skipped by the Priority Clause. Persistent footing is blocked here by the stress clash environment. This shows that syncope only applies to underparsed high vowels.

This analysis accurately accounts for how high vowel syncope works in the language. It also provides an insight into the status of abstract foot structure in Arabic. That is, to the extent that SSA does not have secondary stress and that secondary stress is not well motivated in Arabic, foot construction then becomes an issue of nothing more than being there to get the main word stress assigned properly. However, under the proposed analysis that high vowel syncope is applicable whenever the target high vowel occurs in an underparsed position, be it WLP, the Priority Clause, or the ban on degenerate feet, then this brings evidence for the reality of abstract foot structure in Arabic since it neatly accounts for the difference between forms such as (16b) where high vowel syncope does not apply to forms such as (16d) where it does.

3.3 Geminates and long vowels

The special behavior of geminates and long vowels in attracting stress reveals that there are different degrees of syllable weight in the language. This is attributed to the inherently bimoraic weight of geminates and long vowels in comparison with non-final heavy CVC syllables that become heavy only after the application of Weight-By-Position. This is referred to as a difference between distinctive weight such as that of geminates or long vowels and coerced weight such as that of non-final CVC syllables (Morén, 1999). This means that while CVG/CVV syllables are inherently bimoraic, non-final CVC syllables are not underlyingly bimoraic at some point. Based on this prosodic difference between these types of heavy syllables, the special behavior of geminates and long vowels can be accounted for by reference to prominence rather than to foot structure (Davis, 1989; Hayes, 1995; Watson, 2002). Hayes (1995) proposes Prominence Projection as a device that detects different degrees of syllable weight based on which stress is assigned according to the particular End Rule of the language. The prominence analysis conjectures grid marks that are added to syllables to detect their degree of prominence. This is done along the lines in (17).

- (17) Prominence Projection (Hayes, 1995: 284):
 Project a prominence grid as follows:
 **: heavy syllables
 *: light syllables

The following example in (18) shows an implementation of the prominence projection adopted from Watson (2002) in her account for stress assignment in San'ani Arabic where CVG/CVV syllables are marked with two asterisks on the prominence grid while non-final CVC syllables are indicated with one asterisk. This is shown in the following example for the word [daw.war] 'I searched'. There are two heavy bimoraic syllables in this word. First, moraic trochee feet are constructed over the two bimoraic syllables in the word (18a), i.e. CVG and non-final CVC syllables. Then, after prominence projection, stress is assigned to the CVG syllable rather than to the non-final CVC syllable based on the prominence grid in which a CVG syllable is more prominent than a CVC syllable (18b).

- (18) Prominence-based stress assignment:
 a. Foot construction.

(*) (*)

d a ww a r ‹t

- b. Projection of prominence grid, ERR.

$$\begin{array}{c} (* \quad \quad) \\ (*) \quad (*) \\ \\ d \ a \ w w \ a \ r \ \langle t \rangle \\ * \quad \quad * \\ * \end{array}$$

While the final CVCC syllable is eligible for stress assignment since foot extrametricality is blocked by consonant extrametricality, after the projection of the prominence grid, stress is assigned to the CVG syllable since it is marked on the prominence grid as the most prominent syllable. Thus, End Rule Right in such forms refers to the prominence grid rather than to the foot layer. This prominence-based analysis is able to account for the special behavior of geminates and long vowels in stress assignment. It shows us that prominence-based stress assignment reflects an unbounded stress system, at least with regard to forms with geminates and long vowels.

3.4 Summary

To summarize, the general stress pattern in SSA is accounted for under weak local parsing, foot extrametricality, and the ban on degenerate feet. Not only this can account for pre-antepenultimate stress in the language, it can also account for other related phenomena such as high vowel syncope. It has been shown that high vowels that are parsed by a foot are not deleted while high vowels that are left unparsed are those prone to deletion. It has also been shown that geminates and long vowels can be accounted for in a prominence-based analysis such as that stress is first prominence-based, in which stress is assigned to geminates and long vowels if there are any in the word. Otherwise, stress is foot-based where it is assigned to the rightmost non-final foot in the word.

4. Optimality-theoretic analysis

The OT analysis in this section aims to account for the stress pattern in SSA through constraint interactions (Kager, 1999; Prince and Smolensky, 1993; McCarthy and Prince, 1993). Three major issues are to be taken into account throughout this analysis. First, there is foot construction and stress assignment under the effect of Weak Local Parsing (WLP). Second is the interaction between high vowel syncope and the general stress pattern in the language. And third there is the effect of geminates and long vowels on stress assignment. To that end, I will argue that this is a binary stress system with local ternary alternations and that these alternations aim to keep

feet apart (Kager, 1994, 2001, 2007; Gordon, 2002; Elenbaas, 1999; Elenbaas and Kager, 1999). It will be shown that unparsed light syllables under WLP primarily serve to satisfy the rhythm constraint *CLASH. It will also be shown that all elements of the presented analysis can be accounted for through standard stress assignment and foot construction constraints.

4.1 Foot construction and stress assignment

As demonstrated earlier, feet are binary and the direction of parsing is from left to right with End Rule Right (ERR). These major elements of the analysis are accounted for through standard foot construction constraints presented below. An important constraint that is assumed in this analysis is the foot type constraint FTTYPE=TROCHEE (Kager, 1999).

(19) Rhythm, Directionality, and foot alignment constraints:

- a. ALL-FT-LEFT
Align (Ft, Left, PrWD, Left)
Every foot stands at the left edge of the PrWd.
- b. ALL-FT-RIGHT
Align (Ft, Right, PrWD, Right)
Every foot stands at the right edge of the PrWd.
- c. FT-BIN
Feet are binary under moraic or syllabic analysis.
- d. PARSE-SYLL
Syllables are parsed by feet.
- e. *CLASH
No stressed syllables are adjacent.

The direction of parsing is accounted for by the ranking that ALL-FT-LEFT is ranked higher than ALL-FT-RIGHT. Thus, foot construction from left to right is accounted for as the following tableau shows.

(20) [ra.ga.ba] 'neck'

| /ragaba/ | ALL-FT-LEFT | ALL-FT-RIGHT |
|-----------------|-------------|--------------|
| → a. (ra.ga).ba | | * |
| b. ra.(ga).ba | *! | |

In this tableau, directionality of parsing from left to right is demonstrated as candidate (a) shows, which is the optimal output form. Candidate (b) is ruled out by the constraint ALL-FT-LEFT since it fails to align the constructed foot to the left edge of the PW.

Foot binarity is achieved through the ranking that the constraint FT-BIN is ranked higher than the constraint PARSE-SYLL. This ranking is shared by all binary stress systems (Kager, 1994). This is illustrated in the following tableau.

(21) [ra.ga.ba] ‘neck’

| /ragaba/ | FT-BIN | PARSE-SYLL |
|------------------|--------|------------|
| → a. (ra.ga).ba | | * |
| b. (ra.ga).(ba) | *! | |

In (21), candidate (a) is the optimal output form in which the final syllable is left unfooted. Candidate (b) fails since it parses the final light syllable into a foot, ruled out by FT-BIN.

Through the ranking that PARSE-SYLL is ranked higher than ALL-FT-LEFT, persistent footing is achieved. These are gradient constraints that are violated by the number of involved syllables. This is better shown with a four light syllable word as the following tableau shows.

(22) [ra.ga.ba.ti] ‘my neck’

| /ragaba-ti/ | PARSE-SYLL | ALL-FT-LEFT |
|----------------------|------------|-------------|
| → a. (ra.ga).(ba.ti) | | ** |
| b. (ra.ga).ba.ti | *!* | |

In this tableau, candidate (a), which achieves exhaustive parsing, is selected as the optimal output form. Candidate (b) is ruled out because of its multiple violations of the constraint PARSE-SYLL.

Another alignment constraint is needed to account for the observation that in SSA word-initial stress is prioritized over word-final stress. These are the word alignment constraints ALIGN-PRWD-LEFT and ALIGN-PRWD-RIGHT (McCarthy and Prince, 1993). These constraints ensure edge stress assignment. With the strong tendency for left edge stress assignment in this dialect, the needed ranking is that ALIGN-PRWD-LEFT (ALL-WD-L) is ranked higher than ALIGN-PRWD-RIGHT. As demonstrated earlier, due to the strong ban on degenerate feet in the language, the constraint FT-BIN must be ranked higher than ALIGN-PRWD-LEFT. This ranking also accounts for the Priority Clause condition. This is shown in the following tableau.

(23) [ka.riim] ‘generous’

| /kariim/ | FT-BIN | ALIGN-PRWD-LEFT |
|----------------|--------|-----------------|
| → a. ka.(riim) | | * |
| b. (ka).(riim) | *! | |

In (23), candidate (a) is the optimal form in which the first light syllable is left unparsed. If it is parsed as a unary foot, as candidate (b) shows, this is ruled out by the constraint FT-BIN, an OT interpretation of the Priority Clause condition.

As was explained in the presented analysis, pre-antepenultimate stress in the language is attributed to Weak Local Parsing. WLP is accounted for through the ranking of the rhythm constraint *CLASH above the parsing constraint PARSE-SYLL. This is shown in the following tableau. Note that while final CVC syllables are not bimoraic due to the high-ranked constraint *FINAL-C- μ , non-final CVC syllables become bimoraic as a result of the constraint Weight-By-Position (WBP), which in turn violates the low-ranked constraint DEP-LINK-MORA[C] (Hayes, 1989; Kager, 1999; Morén, 1999).

(24) ['mak.ta.ba] 'library'

| /maktaba/ | *CLASH | PARSE-SYLL |
|------------------|--------|------------|
| → a. (mak).ta.ba | | ** |
| b. (mak).(ta.ba) | *! | |

In this tableau, the optimal output form is candidate (a), in which only one binary foot is constructed over the first heavy syllable while the final two light syllables are left unparsed. Candidate (b) is ruled out by the anti-clash constraint since it constructs a foot over the last two light syllables, resulting in a stress clash. These constraints show that WLP is only a matter of constraint ranking.

Stress assignment and foot extrametricality are accounted for through standard head-foot alignment and non-finality constraints (Kager, 1999, 2007; McCarthy and Prince, 1993; Prince and Smolensky, 1993). These constraints are listed below.

(25) Head foot alignment and stress assignment constraints:

- a. ALIGN-HEAD-R Align (PrWd, R, Head/PrWd, R)
The PrWd ends with the primary stress foot.
- b. ALIGN-HEAD-L Align (PrWd, L, Head/PrWd, L)
The PrWd begins with the primary stress foot.
- c. NONFINALITY
No stress foot is final in PrWd.

Stress assignment under ERR is accounted for through the ranking that the constraint ALIGN-HEAD-R is ranked higher than the constraint ALIGN-HEAD-L. Foot extrametricality is accounted for as the ranking that the constraint NONFINALITY is ranked above ALIGN-HEAD-R. This is illustrated in the following tableau.

(26) ['ra.ga.ba.ti] 'my neck'

| /ragaba-ti/ | NONFINALITY | ALIGN-HEAD-R |
|-----------------------|-------------|--------------|
| → a. ('ra.ga).(ba.ti) | | ** |
| b. (ra.ga).('ba.ti) | *! | |

If stress is assigned to a word-final foot as candidate (b) shows, the constraint NONFINALITY is violated, resulting in its elimination. Candidate (a), in which stress is assigned to the rightmost non-final foot in the word, is selected as the optimal output form.

Now we consider the relative ranking of the constraints discussed thus far through the following constraint ranking: FT-BIN >> NONFINALITY, ALL-PRWD-LEFT, *CLASH >> PARSE-SYLL >> ALL-FT-LEFT >> ALIGN-HEAD-R. Based on this ranking, a comprehensive tableau is sketched below.

(27) ['mak.ta.ba.ti] 'my library'

| /maktaba-ti/ | FT-BIN | NON-FIN | ALL-WD-L | *CLASH | PARSE-SYLL | ALL-FT-L | ALL-HD-R |
|------------------------|--------|---------|----------|--------|------------|----------|----------|
| → a. ('mak).ta.(ba.ti) | | | | | * | ** | *** |
| b. ('mak).ta.ba.ti | | | | | **!* | | *** |
| c. ('mak).(ta.ba).ti | | | | *! | * | * | *** |
| d. mak.('ta.ba).ti | | | *! | | ** | * | * |
| e. (mak).ta.('ba.ti) | | *! | | | * | ** | |
| f. (mak).('ta.ba).(ti) | *! | | | * | | * , *** | * |

This tableau shows that the optimal output form based on this ranking is candidate (a). Candidate (b) is ruled out by the constraint PARSE-SYLL. Candidate (c) fails because it results in a stress clash. Candidate (d) violates the word alignment constraint ALL-PRWD-LEFT since it fails to parse the first bimoraic syllable as a foot. Candidate (e) assigns stress to the final foot in the word, violating the constraint NONFINALITY. Candidate (f) is ruled out by the foot binarity constraint because it constructs a unary foot over the final light syllable.

4.2 High vowel syncope

As noted earlier, high vowel syncope in the language is subject to foot parsing. Under this OT analysis, this claim is further supported through the interactions of the proposed constraint ranking. That is, high vowel syncope comes naturally

through constraint interactions to the extent that no specific anti-high vowel constraint is invoked. Consider the following tableau for the input form /bulud-i/ ‘my lands’. Note that syncope does not apply to word-final high vowels.

(28) [‘bu.lu.di] ‘my lands’

| /bulud-i/ | ALL-FT-L | ALL-HD-R |
|------------------|----------|----------|
| → a. (‘bu.lu).di | | * |
| b. bu.(‘lu.di) | *! | |

In (28), the optimal form is candidate (a), in which the first two high vowels are parsed by a foot. Hence syncope is blocked. That is why we do not get *[bul.di] for instance. If candidate (b) were the optimal form, then the first high vowel is supposed to delete since it is not parsed by the foot such as *[blu.di]. However, candidate (b) is ruled out by the foot alignment constraint ALL-FT-L. For the application of high vowel syncope to underparsed high vowels, consider the following tableau for the input form /maʃrifa-ti/ ‘my knowledge’.

(29) [‘maʃ.rfa.ti] ‘my knowledge’

| /maʃrifa-ti/ | ALL-WD-L | *CLASH | PARSE-SYLL |
|------------------------|----------|--------|------------|
| → a. (‘maʃ).ri.(fa.ti) | | | * |
| b. (maʃ).(‘ri.fa).ti | | *! | * |
| c. maʃ.(‘ri.fa).ti | *! | | ** |

In (29), the optimal form is candidate (a), in which syncope targets the underparsed high vowel, which is skipped by WLP, leaving the high vowel trapped between two binary feet. That is why we end up having [‘maʃ.rfa.ti] as the optimal output form after high vowel syncope. If candidate (b) were the optimal form, syncope would be blocked, but this is ruled out by the anti-clash constraint *CLASH. Similarly, if candidate (c) were the optimal output form, syncope would be blocked, but this is ruled out by the word alignment constraint ALL-WD-L. This shows that high vowel syncope in the language can be accurately accounted for by reference to foot structure rather than to stress assignment.

4.3 Geminate and long vowel effects

The effect of geminates and long vowels on stress assignment is accounted for through the constraint Peak-Prominence (PK-PROM), which expresses the condition that if stressed, then heavy (Kager, 2007; Walker, 1997; Prince and Smolensky, 1993). This constraint simply means that stress is assigned to the highest peak or to the most prominent element in the word.

(30) Peak-Prominence (PK-PROM):

Peak (x) ϕ Peak (y) if $|x| > |y|$.

The element x is a better peak than y if the intrinsic prominence of x is greater than that of y.

Based on the notion that geminates and long vowels are inherently bimoraic while nonfinal CVC syllables become bimoraic only after the application of WBP (Morén, 1999), the following heaviness scale is proposed. It aims to measure different degrees of syllable weight such as that geminates and long vowels are the heaviest syllables.

(31) Heaviness scale:

 $|CVG, CVV| > |CVC, CV|$

Stress assignment under PK-PROM will assign stress to a heavy geminate or a long vowel if there is any in the word. To ensure stress assignment under ERR, this constraint is ranked higher than the constraint ALIGN-HEAD-R. Consider the following example in which two heavy syllables, i.e. CVG and non-final CVC, are eligible for stress assignment. Note that prominence-based stress assignment does not refer to foot construction (Hayes, 1995).

(32) ['dar.ra.sat.ni] 'she taught me'

| /darrasat-ni/ | PK-PROM | ALIGN-HEAD-R |
|---------------------|---------|--------------|
| → a. 'dar.ra.sat.ni | | *** |
| b. dar.ra.'sat.ni | *! | * |

If stress is assigned to the non-final CVC syllable, as candidate (b) shows, it is ruled out by the constraint PK-PROM. This leaves candidate (a) as the optimal output form in which stress is assigned to the CVG syllable.

When the word has more than one heavy syllable with the same degree of heaviness, stress is then assigned to the rightmost prominent peak according to ALIGN-HEAD-R. Consider the following example in which two heavy syllables, i.e. CVV and CVG, are equal in prominence.

(33) [tuf.'faa.ha.ti] 'my apple'

| /tuffaaha-ti/ | PK-PROM | ALIGN-HEAD-R |
|---------------------|---------|--------------|
| → a. tuf.'faa.ha.ti | | ** |
| b. 'tuf.faa.ha.ti | | ***! |
| c. tuf.faa.'ha.ti | *! | * |

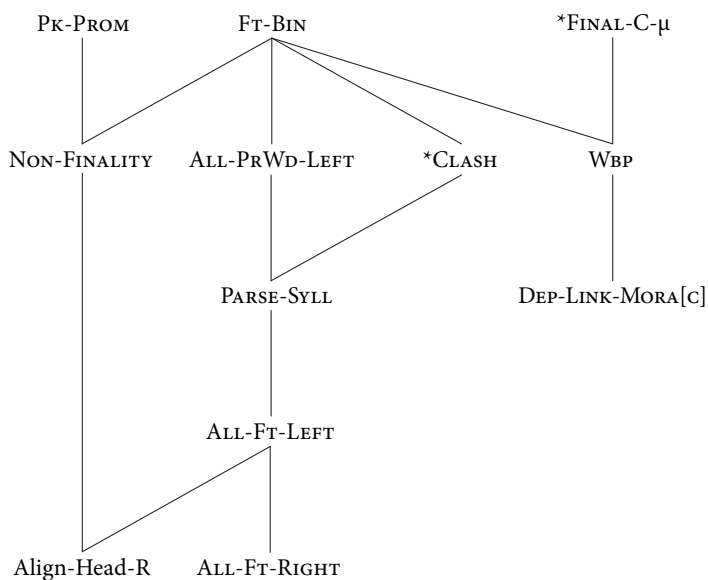
The optimal output form here is candidate (a), which respects PK-PROM and incurs the least number of violations of ALIGN-HEAD-R. Candidate (b) also respects

PK-PROM for stress assignment on the first CVG syllable, but it is ruled out by its multiple violations of the constraint ALIGN-HEAD-R. Candidate (c) is ruled out by PK-PROM because it ignores the most prominent syllables in the word.

4.4 Summary

The presented OT analysis accounts for the general stress pattern in SSA based on constraint interaction. It has been shown that all elements of the presented analysis are accountable through standard constraints of stress assignment and foot construction. Through these constraint interactions, the analysis independently accounts for other related phenomena such as high vowel syncope to the extent that we do not even need to invoke a specific anti-high vowel constraint. It also shows that local ternary intervals primarily avoid stress clashes. The effect of geminates and long vowels on stress assignment is accounted for through the prominence constraint PK-PROM (Prince and Smolensky, 1993). According to this constraint, stress is assigned to the most prominent syllable in the word, i.e. CVG/CVV. Otherwise, stress is assigned to the rightmost non-final foot. Based on the general stress pattern of non-final stress assignment, the constraint PK-PROM is assumed to outrank NON-FINALITY, but since word-final long vowels do not surface in the language there is no way this ranking argument could be demonstrated. An overall constraint ranking of the relevant constraints is given below.

(34) Final constraint ranking:



5. Conclusion

In this paper, I have presented an analysis that accounts for the general stress pattern in the dialect of SSA. The analysis specifically accounts for issues such as pre-antepenultimate stress, geminate and long vowel effects, the quality of word-final vowels, and high vowel syncope. It has been shown that the general stress pattern can be accounted for under Weak Local Parsing (WLP) (Hayes, 1995). The analysis under WLP shows that feet are kept apart minimally by skipping light syllables after a binary foot is constructed, particularly in stress clash environment. This accounts for pre-antepenultimate stress in the language which is further supported by the OT analysis that shows that local ternary intervals result from the violation of the constraint PARSE-SYLL for the sake of satisfying the high-ranked constraint *CLASH. The analysis also reveals that degenerate feet are absolutely banned due to the high-ranked constraint FT-BIN and that foot extrametricality is applied due to the high-ranked constraint NON-FINALITY. It has been shown that stress is assigned according to ERR and that the effect of geminates and long vowels can be accounted for in a prominence-based analysis through projection of prominence grid (Hayes, 1995). In OT terminology, this is interpreted as the constraint P_K-PROM through which stress is assigned to the most prominent syllable in the word in a scalar-quantity sensitivity stress system such as that heavy geminates and long vowels are inherently bimoraic while heavy non-final CVC syllables are not.

The analysis has a major advantage as it naturally extends to account for other related phenomena such as high vowel syncope and the quality of word-final vowels. It has been shown that high vowel syncope is subject to foot parsing. That is, high vowels that are parsed by a foot are immune to syncope while high vowels that are not parsed by a foot are the ones that undergo deletion. The proposed analysis has implications for the general theory of metrical stress with regard to the role of abstract foot structure in stress assignment. In particular, it brings evidence for the reality of abstract foot structure in Arabic through the patterns of high vowel syncope and foot construction in this dialect. The analysis also sheds light on the role of the word-alignment constraint ALL-PRWD-LEFT in foot construction. It gives an insight into the role of this constraint in triggering foot construction while at the same time accounting for the strong tendency of non-final word stress in the language. And finally, it provides an additional example beyond San'ani Arabic (Watson, 2002) where syllables with geminates and long vowels are special in attracting stress over other CVC syllables.

Acknowledgements

I would like to specially thank Stuart Davis for his valuable comments on earlier versions of this paper. I would also like to thank two anonymous reviewers for their helpful comments and notes on various aspects of the paper. Thanks also to the participants of the linguistics graduate students' presentations at Indiana University, Bloomington in Spring 2018 for their input, and to the audience of the Arabic Linguistics Symposium at Arizona State University (ASAL 32) for useful discussion on this paper. All errors are my own responsibility.

References

- Alahmari, Musa M. (2018). An optimality-theoretic analysis of some aspects of the phonology and morphology in a southwestern Saudi Arabian Arabic dialect. Ph.D. Dissertation. Indiana University Bloomington.
- Bamakhramah, Majdi A. (2009). Syllable structure in Arabic varieties with a focus on super-heavy syllables. Ph.D. Dissertation. Indiana University Bloomington.
- Davis, Stuart. (1989). Stress, syllable weight hierarchies, and moraic phonology. *Eastern States Conference on Linguistics*, 6, 84–92.
- Davis, Stuart. (2011a). Quantity. In John Goldsmith, Jason Riggle and Alan Yu (eds.), *The Handbook of Phonological Theory*, 2nd edition (pp. 103–140). Oxford: Wiley-Blackwell. <https://doi.org/10.1002/9781444343069.ch4>
- Davis, Stuart. (2011b). Geminates. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell Companion to Phonology* (V. 2, pp. 873–897). Malden, MA & Oxford: Wiley-Blackwell. <https://doi.org/10.1002/9781444335262.wbctpo037>
- Davis, Stuart and Isao Ueda. (2002). Mora augmentation processes in Japanese. *Journal of Japanese Linguistics*, 18, 1–23. <https://doi.org/10.1515/jjl-2002-0104>
- Davis, Stuart and Isao Ueda. (2006). Prosodic vs. morphological mora augmentation. *Lexicon Forum*, 2, 121–143.
- Elenbaas, Nine. (1999). A unified account of binary and ternary stress. Ph.D. Dissertation. Utrecht University.
- Elenbaas, Nine and René Kager. (1999). Ternary rhythm and the *LAPSE constraint. *Phonology*, 16, 273–330. <https://doi.org/10.1017/S0952675799003772>
- Gordon, Matthew. (2002). A factorial typology of quantity-insensitive stress. *Natural Language & Linguistic Theory*, 20, 491–552. <https://doi.org/10.1023/A:1015810531699>
- Hayes, Bruce. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, 253–306.
- Hayes, Bruce. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press: Chicago.
- Kabrah, Rawiah S. (2004). Opacity and transparency in the phonology of Makkan Arabic: A stratal optimality-theoretical approach. Ph.D. Dissertation. Boston University.
- Kager, René. (1994). *Ternary rhythm in alignment theory*. Ms, Utrecht University. Available as ROA-35 from the Rutgers Optimality Archive.
- Kager, René. (1999). *Optimality Theory*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511812408>

- Kager, René. (2001). Rhythmic directionality by positional licensing. Available as ROA-514 from the Rutgers Optimality Archive.
- Kager, René. (2007). Feet and metrical stress. In De Lacy, Paul (ed.), *The Cambridge Handbook of Phonology* (pp. 195–227). Cambridge, MA: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486371.010>
- Kager, René. (2009). Stress. *Encyclopedia of Arabic Language and Linguistics* (IV, pp. 344–353). Leiden, The Netherlands: Brill.
- Kenstowicz, Michael. (1980). Notes on Cairene Arabic syncope. *Studies in the Linguistic Sciences*, 10, 39–53.
- McCarthy, John. (1979). On stress and syllabification. *Linguistic Inquiry*, 10, 3, 443–465.
- McCarthy, John. (2005). The length of stem-final vowels in Colloquial Arabic. In Mohammad T. Alhawary, and Elabbas Benmamoun (eds.), *Current Issues in Linguistic Theory 267: Perspectives on Arabic Linguistics* (pp. 1–26). Philadelphia: John Benjamins Publishing Company.
- McCarthy, John and Alan Prince. (1993). Generalized alignment. In G. E. Booij and J. van Marle, eds. (pp. 79–153). *Yearbook of Morphology*. Dordrecht, Kluwer.
- Morén, Bruce T. (1999). Distinctiveness, coercion and sonority: A unified theory of weight. Ph.D. Dissertation. University of Maryland at College Park.
- Prince, Alan and Paul Smolensky. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell Publishing.
- Walker, Rachel. (1997). *Mongolian stress, licensing, and factorial Typology*. Ms., University of California at Santa Cruz. Available as ROA-172 from the Rutgers Optimality Archive.
- Watson, Janet. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
- Watson, Janet. (2011). Word stress in Arabic. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell Companion to Phonology* (V. 5, pp. 2990–3018). Oxford: Blackwell. <https://doi.org/10.1002/9781444335262.wbctpo124>

What triggers *'imāla*

Focus on a Palestinian variety with phonological analysis

Noor Abo Mokh and Stuart Davis
Indiana University

'imāla is a term used to describe the raising of the low vowel /a/ to [e] or [i], a feature of several Arabic dialects, including Levantine varieties where conditional raising occurs. This paper discusses various issues related to final *'imāla* focusing on a Palestinian dialect. Final *'imāla* typically affects the feminine marker suffix /-a/ in nouns and adjectives (the equivalent of *tā marbūʿa* in Standard Arabic). Some of the issues to be discussed concern consonants that block *'imāla* and their phonological features, the exceptional classes of colors and mental/physical challenges, and the lack of *'imāla* to loanwords. One important observation is that short nouns whose consonantal roots start with a /w/, such as [sʰif-a] 'characteristic' and [hib-a] 'gift', systematically fail to undergo *'imāla*. Phonological analyses will be given.

Keywords: *'imāla*, raising, feature geometry, feminine suffix, morphosemantic class, loanwords, Palestinian Arabic

1. Introduction

'imāla is the traditional Arabic term for the raising of the low vowel /a/ to [e] or [i] that is widespread in some Arabic varieties including Levantine varieties. *'imāla* was already observed in some of the Iraqi dialects in the 8th century by Sībawayhi, Kitāb II, 281.4–5, 281.9–10.¹ The early grammarians divided *'imāla* into final and non-final *'imāla* with more observations on non-final *'imāla*. Both final and non-final *'imāla* occur in contemporary dialects. For example, for the dialect to be considered in this paper, the raising typically only affects the word-final

1. Details of Sībawayhi's discussion and that of other early Arabic grammarians can be found in Levin (2011).

suffix *-a* that marks the feminine on nouns and adjectives and not other instances of the final vowel *-a* where the *-a* is not morphological (as in [ʃafa] ‘a dinner’ for instance).² *‘imāla* has been the focus of several recent studies. One study by Habib (2012) examines non-final *‘imāla* in a rural Syrian variety, while another study by Shachmon (2011) mentions three variants of *‘imāla* in central Palestinian dialects: urban (Jerusalem), rural Triangle dialects, and rural *a*-dialects. Given the classification presented in Shachmon (p. 146) that divides these dialects with respect to degree of raising (see also Shahin 2012), the dialect to be examined in this paper (a rural Palestinian Triangle variety spoken in Israel), fits Shachmon’s categorization of raising to the highest vowel, [i]. In addition to this classification, Shachmon (2011) discusses pausal *‘imāla* where raising only occurs before a pause and not within a syntactic phrase. In the dialect to be described here *‘imāla* occurs even in nonpausal position, for example within a noun phrase. A somewhat different phenomenon in Central Palestinian Arabic is described and analyzed by Shachmon and Faust (2017) whereby there is vowel lowering in a prepausal final syllable as exemplified by a word like /kútub/ ‘books’, which has the prepausal pronunciation of [kútob] (where the vowel in the final syllable is lowered and somewhat lengthened, but not stressed). They distinguish this phenomenon from *‘imāla*; we note that such prepausal lowering does not seem to be characteristic of the dialect to be discussed.³

While final vowel raising is a salient feature of many varieties of Levantine Arabic, even reflected by the pronunciation of some Arabic loanwords in English (for example, *tahina* vs *tahini* or *knafa* vs. *knafi*, where the latter pronunciation of each pair displays *‘imāla*), the details of final *‘imāla* in a given variety of Arabic are rarely discussed in the literature, although Younes (1982), Herzallah (1990), Al-Wer (1998, 2007), and Al-Wer et al. (2015) contain pertinent information on West Bank Palestinian varieties.⁴ The details of final raising/*‘imāla* become

2. The final vowel in [ʃafa] is lengthened when followed by a suffix as exemplified by [ʃafaa-na] ‘our dinner’; however in the case of words with the word final feminine suffix, the *tā marbūʿa* resurfaces, for instance, [saafa] ‘watch’ becomes [saafi-t-na] ‘our watch’ and [miʃlaka] ‘spoon’ becomes [miʃlaka-t-na] ‘our spoon’. The precise phonetic realization of the vowel immediately before the *tā marbūʿa* in forms with a possessive suffix is a complicated issue that is beyond the scope of the present paper, but see Younes (1982) for discussion.

3. Other studies on *‘imāla* include Owens (2006) where non-final *‘imāla* is discussed, Bellem (2007) discusses other aspects of *‘imāla* such as the issue of emphatics and Durand (2011) where phonological rules that govern *‘imāla* after /r/ are discussed, but the details of these studies will not be discussed in this paper.

4. Additionally, Cotter (2016) and Cotter and Horesh (2015) studied the realization of the raising vowel in Gaza City and Jaffa Palestinian varieties and its variation across generations. They maintain that the default (or underlying) vowel of the feminine ending is *-e*, the raised variant rather than the low vowel. This is an issue that we address at the end of Section 6.

interesting for two reasons. First, in most (if not all) of the Levantine varieties that display *'imāla*, the raising of the final /a/ is prevented when the consonant preceding it is of a certain type. Second, there are certain word classes where raising is prevented even though the phonological form of the word would suggest that raising should apply. This paper aims at examining the factors that trigger *'imāla* in Palestinian Arabic, specifically in the rural Palestinian dialect spoken in the triangle area in Israel (henceforth, PA-TA), the native variety of the first author, and to provide phonological analysis.

In Section 2 of this paper we will present representative data from PA-TA showing the phonological environments where *'imāla* applies and where it is prevented from applying. As will be seen, *'imāla* is prevented from applying when the consonant immediately before the feminine suffix /-a/ is either an emphatic, back velar, pharyngeal or laryngeal. In Section 3 we will present a detailed analysis of the feature representation that captures emphatics, back velars, pharyngeals and laryngeals as a natural class of consonants. Sections 4–6 discuss other issues of phonology and morphology that relate to the PA-TA dialect-specific pattern of *'imāla*. Section 4 focuses on the issue of the phonetic fronting of historical /q/ to [k] and its interaction with *'imāla*. Section 5 deals with the analysis of the exceptional morpho-semantic class of colors and mental/physical challenges that systematically fail to undergo *'imāla* in this dialect. Section 6 discusses other classes of exceptions to *'imāla* that have not been previously observed in the literature to the best of our knowledge. These include short nouns whose consonantal roots start with an unpronounced /w/, such as [s^hif-a] 'characteristic' and [hib-a] 'gift' and also loanwords including those from Modern Hebrew. We will discuss pertinent phonological issues as they arise in the various subsections of the paper. The conclusion in Section 7 will summarize the main points of our study.

2. Exemplification of *'imāla* in PA-TA

In this section, we present representative data that illustrate final *'imāla* (i.e. vowel raising) in the rural PA-TA dialect that is the native variety of the first author. In this dialect, the final suffix -a that typically indicates the singular feminine (or nouns of instance) is raised to -i and applies in both pausal position and prepausal position (i.e. within a noun phrase). Examples of words where the vowel of the feminine marker is raised (*'imāla* applies) are in (1).

(1) Examples of words where *'imāla* occurs

| <u>word</u> | <u>gloss</u> |
|----------------|-----------------------|
| a. [ʃatab-i] | 'a step (to a house)' |
| b. [sitt-i] | 'six' |
| c. [wirθ-i] | 'inheritance' |
| d. [dʒaadʒ-i] | 'chicken' |
| e. [sidd-i] | 'loft' |
| f. [ʃaaðð-i] | 'different f.' |
| g. [mooz-i] | 'banana' |
| h. [haass-i] | 'sense' |
| i. [tʰooʃ-i] | 'fight' |
| j. [hiff-i] | 'side' |
| k. [malik-i] | 'queen' |
| l. [naml-i] | 'ant' |
| m. [rʰasm-i] | 'painting' |
| n. [muun-i] | 'supplies' |
| o. [kahw-i] | 'coffee' |
| p. [kubbaaj-i] | 'cup' |

The data in (1) show the regular pattern where *'imāla*/raising applies in PA-TA. As can be seen, *'imāla* applies regularly to the feminine suffix -a where the low vowel /a/ is raised to [i]. For instance, in the word /ʃatab-a/ in (1a), the final vowel -a is raised to the high vowel [i] ([ʃatab-i]). This also occurs in a noun phrase as in [ʃatab-i dʒdiid-i] ('a new step (to a house)').

While this is the regular pattern of *'imāla*, there are cases where *'imāla* is blocked because of the type of phoneme immediately preceding the feminine suffix. Specifically, raising is prevented systematically from applying when the immediately preceding consonant is either a coronal emphatic (i.e. a pharyngealized coronal) /t^ʕ s^ʕ ð^ʕ r^ʕ/, a back velar /x ɣ K/, where /K/ is underlyingly the historical /q/,⁵ a pharyngeal /ħ ʕ/ or a laryngeal /h ʔ/, as exemplified in (2). Bellem (2007) has classified these sounds in one group including emphatics due to the *'imāla* blocking feature (note that in (2) and elsewhere, low-level vowel quality differences are not indicated in the transcription of words unless germane).

5. To be clear, by back velar, we mean a consonant that is pronounced as velar in the PA-TA dialect under consideration but is uvular in Classical Arabic and various other dialects. The analysis of back velars will be discussed in Sections 3 and 4. We realize that we are using the symbol /K/ in a non-conventional manner since it is not an IPA symbol

(2) Words that end in coronal emphatics, back velars, pharyngeals, and laryngeals

| <u>word</u> | <u>gloss</u> |
|---|--------------------|
| a. [bat ^ʕ t ^ʕ -a] | 'duck' |
| b. [furs ^ʕ -a] | 'break' |
| c. [beeð ^ʕ -a] | 'egg' |
| d. [fið ^ʕ ð ^ʕ -a] | 'silver' |
| e. [ħur ^ʕ r ^ʕ -a] | 'free f.' |
| f. [faxx-a] | 'trap' |
| g. [t ^ʕ abx-a] | 'dish' |
| h. [luy-a] | 'language' |
| i. [milʕak-a]/ [miʕlak-a]/ | 'spoon' |
| j. [s ^ʕ ulħ-a] | 'reconcilia-tion' |
| k. [kaħħ-a] | 'coughing' |
| l. [waaħ-a] | 'oasis' |
| m. [mbahbaħ-a] | 'wide' |
| n. [d̤ʒaruuʕ-a] | 'a gift of sweets' |
| o. [riʔ-a] | 'lung' |
| p. [faakh-a] | 'fruit' |

In all of the examples in (2) *'imāla* is prevented from applying because of the consonant preceding the feminine suffix. In (2a)–(2e) are examples of words that end in a coronal emphatic. In (2f)–(2i) are examples of words that end in a back velar. Note that in (2i), the [k] is underlyingly the historical /q/ which surfaces as [k] in this dialect as opposed to the underlying sound /k/ which can surface as [tʃ] where the vowel undergoes *'imāla* (to be discussed in Section 4). In (2j)–(2n) are words that end with a pharyngeal and in (2o)–(2p) are words that end in a laryngeal. An analysis of *'imāla* needs to account for why coronal emphatics, back velars, pharyngeals, and laryngeals act as a natural class of sounds in blocking raising. We turn to this issue in Section 3. In Section 4 we discuss the related problem as to why the velar stop [k] that historically comes from the uvular stop blocks *'imāla* while [k] that historically comes from a velar stop does not.

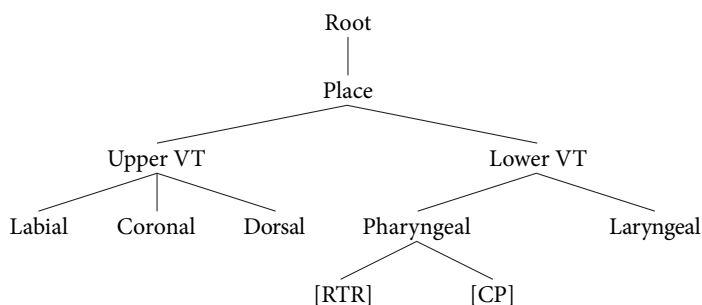
3. Feature geometric representation and the blocking of *'imāla*

We take the position that final *'imāla* involves a process of vowel raising in PA-TA that raises the low vowel /-a/ of the feminine suffix to [i] (in the specific dialect being investigated and ignoring possible low-level variation in the realization of -i). As seen by the data in (2) this process is blocked by the presence of an emphatic, back velar, pharyngeal or laryngeal consonant immediately before the suffix, thus

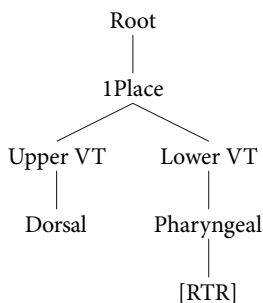
suggesting that this group of consonants acts as a natural class of sounds. In this section, we will offer a detailed feature geometric representation of these consonants in which they can pattern as a natural class. We will then offer an analysis of *'imāla* that accounts for why these consonants block raising.

While the representation of back consonants has been the subject of some debate in the feature geometric literature (see in particular Zawaydeh 1999; Bin-Muqbil 2006, and Sylak-Glassman 2014), here we maintain the feature geometric representations for these consonants as illustrated in (3) mainly following Davis (1995) and Zawaydeh (1999).

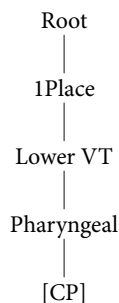
(3) a.



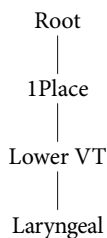
b. *The representation of uvulars*



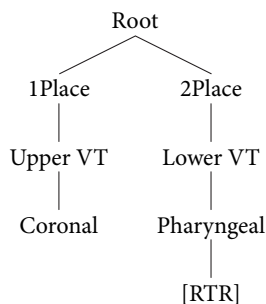
c. *The representation of pharyngeals*



d. *The representation of laryngeals*



e. *The representation of emphatics*



The first diagram in (3a) shows a general feature geometric representation for consonants proposed in Vaux (1993) and adopted by Davis (1995); the other diagrams in (3) display the specific feature geometric representations for uvulars (or back velars), pharyngeals, laryngeals and emphatics. In these diagrams, the node labeled Lower VT stands for the lower vocal tract and is active (or specified) for sounds made with an articulation in the back of the mouth. The node labeled 1Place indicates the primary place of articulation of consonants and is distinguished from 2Place, which would be activated for consonants having a secondary place of articulation. The representation of the uvular shown in (3b) characterizes the uvular as a complex segment having both an Upper VT component (Dorsal) and the Lower VT component Pharyngeal annotated with the retracted tongue root feature [RTR]. The representations shown for pharyngeals and laryngeals in (3c) and (3d) are similar to each other in that their primary place of articulation consists of just the Lower VT node but they differ in that pharyngeals have the node Pharyngeal with the constricted pharynx feature [CP] while the laryngeal consonants have the Laryngeal node. As shown in the diagram (3e), 2Place (i.e. secondary place of articulation) is pertinent for emphatic consonants, which have secondary pharyngealization. Emphatics possess a feature geometric representation in which the primary place of articulation is Coronal located under the Upper VT node while the secondary articulation is indicated with the component Pharyngeal (under the Lower VT node in (3e)), which is annotated with the retracted tongue root feature [RTR], similar to the uvular representation. Consequently, as detailed in Davis (1995), the class of emphatics, uvulars, pharyngeals, and laryngeals (often referred to as gutturals) constitute a natural class of sounds since they all share the Lower Vocal Tract node, meaning that they all involve an articulation of the lower vocal tract.

Given the feature geometric representations shown in (3) above we can offer a formal account as to why *'imāla* (raising) is blocked when the suffixal vowel -a is immediately after an emphatic, uvular (post-velar), pharyngeal, or laryngeal. The idea is that the low vowel /a/ assimilates to the immediately preceding guttural consonant. Assimilation involves the autosegmental spreading of the Lower VT node from the guttural to the low vowel. The low vowel being articulated by a low tongue body position is compatible with having the Lower VT node while non-low vowels would be resistant to it. Evidence that the low vowel /a/ assimilates to the Lower VT node of the preceding guttural comes from the observation that the final low vowel has a retracted pronunciation after an emphatic consonant as in [beeð^ɣ-a] 'an egg' (since the emphatic has the retracted tongue root feature similar to a uvular as shown above) and the low vowel has a more front non-retracted pronunciation after a laryngeal as in [faakh-a] 'a fruit' since laryngeals do not have the retracted tongue root feature. Thus, the exact pronunciation of the final low vowel in data like that in (2) depends on the specific features under the Lower Vocal Tract (VT)

node. This is diagnostic of a (progressive) assimilation process. Given this, the rule or process that raises the final /a/ to [i] only applies to /a/ that is not assimilated to a preceding consonant, that is, one that does not share the Lower Vocal Tract node. In (4), we exemplify our analysis with the word in (2p), [faakh-a] 'fruit', where the suffixal [-a] fails to undergo raising because of the shared Lower VT node with the preceding consonant.

(4) Lack of raising in [faakh-a] 'fruit' (LVT = Lower Vocal Tract Node)

- | | |
|---|---|
| a. Underlying form: | f aa k h-a LVT |
| b. Progressive assimilation: | f aa k h-a LVT |
| c. Final vowel raising (<i>'imāla</i>): | Does not apply (blocked by the linked structure as shown in (b)) ⁶ |
| d. Phonetic form: | [faakha] |

Consequently, final *'imāla*, the raising of /a/ to [i], only applies to a low vowel that is not assimilated to the lower VT node, as in the examples in (1) (where the consonant immediately before the suffix lacks the Lower VT node). But raising would be blocked in the examples in (2) where the consonant immediately before the suffix has the Lower VT node.

As a final matter it should be pointed out that our analysis of final *'imāla* described above is different from that of Younes (1982) and McCarthy (1994) who argue that final *'imāla* involves a process whereby /-i/ is the underlying form of the feminine suffix and it lowers to [a] when adjacent to a guttural. We will return to this matter in Section 6 offering evidence that the formal analysis of *'imāla* involves raising as conceived here and not lowering.

4. The historic uvular /q/ and the back velar problem

Given the analysis presented in the previous section in which the class of gutturals is characterized by the presence of the Lower Vocal Tract (LVT) node in its feature geometric representation, an issue comes up in the PA-TA dialect under consideration that relates to the pronunciation of the historic uvular /q/ as well as the back velar/uvular fricatives /x/ and /χ/. As seen in data like in (2f)–(2i), these sounds pattern as if they are uvular in preventing *'imāla*. However, in the PA-TA dialect, these

6. The notion that a rule can be blocked by a linked structure goes back to Hayes (1986) where a wide variety of cases are provided.

historical uvulars are pronounced as velars. For instance, in this dialect the word /qalb/ 'heart' and the word /kalb/ 'dog' can be pronounced identically as [kaleb] (but mostly as [tʃaleb] for 'dog'). Nonetheless, when the historic uvular is in final position immediately before the suffix -a as in the word /waraq-a/ [waraka] 'paper', or /ʕamiiq-a/ [ʕamiik-a] 'deep', *'imāla* is still prevented even though the sound is pronounced as a velar stop. This is exemplified by the words in (5).

- (5) Examples of words with /q/ that is pronounced as /k/, but blocking *'imāla*

| | <u>word</u> | <u>gloss</u> |
|----|-------------|-----------------------|
| a. | [rkiik-a] | 'soft f.' / 'thin f.' |
| b. | [sirk-a] | 'theft' |
| c. | [sadak-a] | 'charity' |
| d. | [baak-a] | 'a bunch' |

The examples in (5) should be compared to instances of words that end in a historical velar such as /samak-a/ [samak-i] 'fish' (also [samatʃ-i] with /k/-palatalization) where *'imāla* occurs even though the last consonant can be pronounced as a velar /k/. This is illustrated in the examples in (6) below. Especially noteworthy is the comparison between [samak-i] 'fish' in (6a) with [sadak-a] 'charity' in (5c). Both have a velar stop before the final vowel, but *'imāla* is blocked in [sadak-a] since the velar [k] comes from a historical uvular (this will be discussed further towards the end of the section). Our contention is that all the back velars, [k], [x], and [ɣ], are phonologically or underlyingly uvular in PA-TA and it is only a later rule of the phonology that turns them into phonetic velars.

- (6) Examples of words that end in /k/ or /tʃ/ where *'imāla* is not blocked (the treatment of the status of the variation between these two sounds is beyond the scope of this paper)

| | <u>word</u> | <u>gloss</u> |
|----|-----------------------|----------------------------|
| a. | [samak-i]/ [samatʃ-i] | 'fish' |
| b. | [ħarak-i]/ [ħaratʃ-i] | 'movement' |
| c. | [birk-i]/ [birtʃ-i] | 'swimming pool' |
| d. | [friik-i]/ [fritʃ-i] | 'Farik/grits-type of food' |
| e. | [barak-i]/ [baratʃ-i] | 'blessing' |

One way that we can distinguish a [k] sound that we consider a back velar as in (5) from a [k] sound that is not a back velar as in (6) is that the latter can typically palatalize as shown in (6), but not the former.⁷

7. We do note that in the word [malik-i] 'queen' in (1k), the velar /k/ cannot palatalize unlike the words in (6). The form *[malitʃ-i] 'queen' is not found in this dialect. A reviewer added that the term /malika/ occurs in this form invariably in the modern dialects in this region, while in traditional dialects it is mal[a]ka. /malika/ is a 'special term' that people hear on the media repeatedly.

Given our contention that the back velars like [k], [x], [ɣ] are phonologically or underlyingly uvular in PA-TA and it is a later rule that fronts uvulars to velars, we show a derivation in (7) for the word [waraka] ‘paper’ from underlying /waraq-a/ that illustrates how the ordering of processes accounts for the lack of *‘imāla* for words like those in (5) and (6).

(7) Lack of raising in [waraka] from underlying /waraq-a/ ‘paper’

- | | |
|---|--|
| a. Underlying form: | w a r a q -a |
| | |
| | LVT |
| b. Progressive assimilation: | w a r a q a |
| | |
| | LVT |
| c. Final vowel raising (<i>‘imāla</i>): | Does not apply due to the linked structure |
| d. Uvular fronting: | w a r a k a |
| e. Phonetic form: | [waraka] |

Crucially, in the derivation in (7) the raising rule (*‘imāla*) applies before uvular fronting. Consequently, like in the derivation in Example (4), raising is blocked because it does not apply to the low vowel that shares the LVT node with the preceding consonant. The subsequent rule of uvular fronting shown in (7d) has the effect of deleting the LVT node. In this way, we are able to account for the velars exemplified in (6), which are historically velars (lacking the LVT node) and trigger raising, from those that are historically uvular as exemplified in (5) that do not trigger raising. This suggests that the historical uvulars still pattern as uvular even though they are pronounced as velar and so should be represented as such phonologically or underlyingly. One additional piece of evidence for their representation as phonologically uvular is the observation that in PA-TA, as in the Palestinian variety discussed by Herzallah (1990), the historical uvular stop may trigger de-pharyngealization of an emphatic earlier in the word. Thus, as indicated by the transcription in (5c), a historical form like /s^ʕadaqa/ ‘charity’ is pronounced as [sadaka] in the PA-TA dialect under consideration. The de-pharyngealization of /s^ʕ/ in [sadaka] can be seen as a type of OCP (Obligatory Contour Principle) effect against having two guttural consonants in the same word that are too similar, in both having a Lower Vocal Tract node with the feature [RTR]. While we do not deal with the specifics of it, the depharyngealization found with [sadaka] is consistent with the abstract underlying representation where the surface [k] is underlyingly uvular.⁸

8. Another way in which a surface [k] that is underlyingly uvular is kept distinct from the surface [k] that is underlyingly velar comes from an observation that emerges from Al-Wer and Herin (2011) that these two forms are sociolinguistically different, where the [k] that is underlyingly uvular can be realized as a glottal stop, but not the [k] that is underlyingly velar.

5. The semantic class of colors and mental/physical challenges

Another issue pertinent to PA-TA and probably to most dialects regarding *'imāla* relates to the exceptional semantic class of colors and mental/physical challenges (which perhaps can be viewed as being nongradable adjectives; see Grano & Davis 2018). In many Levantine varieties the feminine form of these words, although ending in the low vowel /a/, does not undergo *'imāla*. Examples are shown in (8) below where the first column shows the masculine singular, the second column the feminine singular where *'imāla* fails to apply, and the third column the plural.⁹

(8) Examples of words of colors and mental/physical challenges

| | <u>Word (sg. m.s)</u> | <u>Word (sg. f.)</u> | <u>Word (pl.)</u> | <u>gloss</u> |
|----|-------------------------------------|--|-------------------------------------|---------------------|
| a. | [iswad] | [sood-a] | [suud] | 'black' |
| b. | [ixð ^ʕ ar ^ʕ] | [xað ^ʕ r ^ʕ -a] | [xuð ^ʕ ur ^ʕ] | 'green' |
| c. | [ibjað ^ʕ] | [beeð ^ʕ -a] | [biid ^ʕ] | 'white' |
| d. | [izrak] | [zark-a] | [zuruk] | 'blue' |
| e. | [ihmar ^ʕ] | [ħamr ^ʕ -a] | [ħumur ^ʕ] | 'red' |
| f. | [is ^ʕ far ^ʕ] | [s ^ʕ af ^ʕ r ^ʕ -a] | [s ^ʕ ufur ^ʕ] | 'yellow' |
| g. | [iʕma] | [ʕamj-a] | [ʕimi] | 'visually impaired' |
| h. | [ihbal] | [habl-a] | [hubul] | 'simple-minded' |
| i. | [ihwal] | [ħool-a] | [ħuul] | 'cross-eyed' |
| j. | [it ^ʕ r ^ʕ aʕ] | [t ^ʕ ar ^ʕ ʕ-a] | [t ^ʕ ur ^ʕ uʕ] | 'deaf' |
| k. | [ifkam] | [fakm-a] | [fukum] | 'toothless' |
| l. | [ixr ^ʕ as] | [xars-a] | [xur ^ʕ us] | 'mute' |
| m. | [iʕr ^ʕ aðʕ] | [ʕar ^ʕ ðʕ-a] | [ʕur ^ʕ uðʕ] | 'limping' |
| n. | [iʕwar ^ʕ] | [ʕoor ^ʕ -a] | [ʕuur ^ʕ] | 'one-eyed' |
| o. | [ir ^ʕ waʕ] | [r ^ʕ ooʕ-a] | [r ^ʕ uuʕ] | 'clumsy' |

The words in (8a)–(8f) above are basic color terms whereas the examples in (8g)–(8o) express mental/physical challenges. As seen in the feminine forms of these words in the second column above, all of these words are exceptions to *'imāla*. While some of these words would be exceptional because of their phonology (i.e. their final consonant is a guttural), examples like (8a) and (8h)–(8k) among others show that this class of words systematically fails to undergo raising. In the subsections below, we discuss two possible (synchronic) analyses for the exceptional behavior of these words: a morpho-semantic analysis and a morphophonological analysis.¹⁰

9. There are some dialects with *'imāla*, such as those in Upper Egypt (Sa'iidi) in which words of this class do undergo *'imāla* (Khalafallah, 1969).

10. See Haddad (1984) for a discussion on the different possible analyses of these in Lebanese Arabic.

5.1 Morpho-semantic class

One analysis of the failure of raising (*'imāla*) in the feminine forms in the words shown in (8) is to note that these words constitute a separate morphosemantic class characterized by apparent partial suppletion in the masculine singular (e.g. [ixð^ʕar^ʕ] 'green, m.s.', [ifkam] 'toothless, m.s.', and [ihbal] 'simple-minded-m.s.') and a broken plural having a specific templatic shape (CuCuC for words with strong consonants, e.g. [hubul] 'simple-minded pl.' and predictable variation for words with weak roots, e.g. [suud] 'black pl.'). We view the masculine form of these words as displaying partial suppletion (in the sense of Haspelmath and Sims 2013) since, based on the feminine, the expected masculine singular for (8a) and (8d), for example, would be the ungrammatical [sood] and [zark], respectively, instead of [iswad] 'black' and [izrak] 'blue'. One can then offer an analysis of the morphosemantic class in (8) in the framework of Construction Morphology following Booij (2010) whereby the class of (basic) color words and mental/physical challenges can be viewed as a morphological construction having its own phonological, morphosyntactic, and semantic properties that would be specific to the construction itself. As part of the phonological aspect of the construction, the rule of final vowel raising (*'imāla*) would not be specified as applying to this construction. Thus, the lack of *'imāla* can be viewed as part of a complex morphological construction in the spirit of Booij (2010). We leave the details of this analysis for future research.

5.2 A morphophonological analysis

A second analysis for the lack of *'imāla* in (8) is a morphophonological one. Under such an analysis, words of the morpho-semantic class of colors and mental/physical challenges can be analyzed as having a separate feminine ending /-aaʔ/,¹¹ which undergoes vowel shortening in word-final position. *'imāla* is prevented because it does not apply to a word-final long vowel. Evidence that some words that end in a short final /a/ have a long /aa/ underlyingly come from alternations like that in (9) where the final short vowel surfaces as long when a consonant-initial suffix is added.

11. Note that some traditional grammars of Classical Arabic, such as Wright (1896/1974), consider the ending /-aaʔ/ to be a separate feminine suffix. The Classical Arabic cognate forms of the feminine in (8) would have this ending. However, in most modern dialects there is no evidence for the presence of a final glottal stop. See Ali (2020) for pertinent discussion.

(9) Evidence for final vowel length alternation

| | <u>Underlying forms</u> | <u>Unsuffixd word</u> | <u>Suffixd word</u> | <u>gloss</u> |
|----|-------------------------|-----------------------|---------------------|---------------------------|
| a. | /samaa/ | [sama] | [samaa-na] | 'sky' / 'our sky' |
| b. | /ʕafaa/ | [ʕafa] | [ʕafaa-na] | 'dinner' / 'our dinner' |
| c. | /masaa/ | [masa] | [masaa-na] | 'evening' / 'our evening' |
| d. | /hawaa/ | [hawa] | [hawaa-na] | 'air' / 'our air' |

Thus, one can argue that the underlying form of the vowel resurfaces under suffixation. This is consistent with McCarthy (2005) who discusses in detail alternations like that in (9) and argues from an optimality-theoretic perspective that they support an analysis where the stem-final vowel is underlyingly long and shortens when in word-final position.¹²

Within an analysis using a rule-based phonological framework, the rule of raising (*'imāla*) is restricted so that it only applies to a short final -a. There would then be an ordering restriction such that raising (*'imāla*) precedes the shortening of the final long vowel; thus, from the underlying form of a word like /xars-aa/ 'mute, f' in (8), *'imāla* would fail to apply since the final vowel would be long. Vowel shortening would subsequently apply resulting in the opaque [xars-a]. This phonological analysis assumes that *'imāla* is ordered relatively early in the derivation, which is consistent with the derivation shown in (7) where *'imāla* applies before the historical uvular becomes a velar. The rule ordering analysis is shown below in (10) for the word [xarsa] 'mute' where we posit /xars-aa/ as the underlying form.

| | |
|--|-------------------------------|
| (10) Underlying form: | /xars-aa/ |
| Final vowel raising (<i>'imāla</i>): | does not apply to long vowels |
| Vowel shortening: | xars-a |
| Phonetic form: | [xarsa] |

In the derivation in (10), final vowel raising and vowel shortening are in a counterfeeding relation. While this rule ordering analysis accounts for the exceptionality of the class of colors and physical/mental challenges, it does so at the expense of positing a second feminine suffix synchronically.¹³ The analysis in Section 5.1 using Booij's construction morphology framework does not have to posit a separate feminine suffix for such words. We do not decide on the matter of which is the

12. The vowel length alternation shown in (9) could hypothetically apply to the feminine singular forms in (8), but such word forms would be semantically odd when a suffix like /-na/ 'our' is added; they would require a very creative context for the meaning to be felicitous.

13. As mentioned in footnote 11, some traditional grammars of Classical Arabic, such as Wright (1896/1974), do consider the ending /-aaʔ/ to be a separate feminine suffix.

better analysis here, but just show that the exceptional class in (8) can either be accounted for phonologically as in this subsection or through the perspective of Construction Morphology.

6. Other exceptional word groups

In this section, we point out two additional word types that are exceptional to *'imāla*: the first are very short nouns that have a root-initial /w/ which does not surface, and the second group are loanwords. First, consider the nouns in (11).

(11) Short nominal form with root initial /w/

| | <u>noun</u> | <u>verb</u> | <u>gloss</u> |
|----|-----------------------|-----------------------|---------------------------|
| a. | [s ^ʕ if-a] | [was ^ʕ af] | 'characteristic/describe' |
| b. | [s ^ʕ il-a] | [was ^ʕ al] | 'relation/connected' |
| c. | [hib-a] | [wahab] | 'give /gift' |
| d. | [siʕ-a] | [wisiʕ] | 'capacity/became wide' |
| e. | [θiq-a] | [wiθiq] | 'trust/trusted' |

The nouns in (11) are very short; they contain roots with an initial /w/, which does not surface. For instance, (11a), the root of the word [s^ʕif-a] is w.s^ʕ.f. The root-initial /w/ would occur in the perfective form of the verb, as shown in (11), and in certain derived nouns. Roots with initial /w/ occasionally have nominalizations lacking /w/ and are quite short for a nominalized form as exemplified in (11). We make the observation that such nominalized forms seem to systematically fail to undergo *'imāla*. These words are not like the feminine forms in (8) that arguably end in /-aa/ since there is no evidence from a length alternation that nominals like those in (11) have an underlying long vowel. (Compare [s^ʕif-a] with [s^ʕifat-na] 'our characteristic' where unlike in (9) there is no length alternation.). Moreover, while the words in (11), of the pattern [CVC-a], are short, their shortness is not what makes them immune to *'imāla* given that other words of this shape do undergo *'imāla* as exemplified by /san-a/ 'year', which is pronounced as [san-i] with raising in the PA-TA dialect under consideration (and as [sin-i] in closely related varieties). As far as we are aware, the systematic absence of *'imāla* for words like those in (11) has not been previously discussed in the literature. We suspect that these words are probably resistant to *'imāla* in other varieties of Levantine Arabic as well, but the matter needs further investigation.¹⁴

14. A reviewer suggests that the lack of *'imāla* in words like [hib-a] 'gift' and [s^ʕif-a] 'characteristic' in (11) is because they are loanwords from Classical Arabic; however, in PA-TA other loanwords from Classical Arabic, such as /θaqa:fa/ realized as [θaqa:fi], do undergo *'imāla*. Therefore, we

A final group of words we will discuss here that are typically exceptional to *'imāla* are loanwords. Examples of loanwords can be found in (12) and (13) below. The words in (13) are Modern Hebrew words borrowed into the PA-TA dialect under consideration while the words in (12) are borrowed from other languages.

(12) The case of Loanwords – the systematic non-application of *'imāla*

| | <u>word</u> | <u>gloss</u> | <u>noun-adjectives</u> | <u>gloss</u> |
|----|--|----------------|---|-----------------------|
| a. | [ooð ^s -a] | 'room' | [ooð ^s a kbiiri] | 'a big room' |
| b. | [makiin-a] | 'machine' | [makiina zdiidi] | 'a new machine' |
| c. | [t ^s ar ^s abeez-a] | 'coffee table' | [t ^s ar ^s abeeza t ^s awiili] | 'a long coffee table' |
| d. | [mooð ^s -a] | 'fashion' | [mooð ^s a yariibi] | 'a strange fashion' |
| e. | [kamar ^s -a] | 'camera' | [kamar ^s a mniiha] | 'a good camera' |
| f. | [kundar ^s -a] | 'shoes' | [kundar ^s a ſaalji] | 'high-heel shoes' |

(13) Words with feminine suffix borrowed from Hebrew (word stress is indicated)¹⁵

| | <u>Word in Hebrew</u> | <u>Word in Arabic</u> | <u>gloss</u> |
|----|--------------------------|---|--------------|
| a. | [ʔisk-á] | [ʔisk-á:] | 'deal' |
| b. | [glí:d-a]/[glid-á] | [glí:d-a] | 'ice cream' |
| c. | [mift ^s aʁ-á] | [mift ^s ar ^s -á:] | 'police' |
| d. | [misʔad-á] | [misʔad-á:] | 'restaurant' |

The loanwords in (12) are borrowed into Arabic with the final [a] of the source form interpreted as being a marker of feminine gender even if the source language does not have a gender distinction. In the words in (12), the final [a] does not raise even if the phonological conditions are favorable for raising as in (12b)–(12c). Moreover, as shown by the noun-adjective phrases in the third column in (12), the borrowed noun that lacks raising (*'imāla*) can be modified by an Arabic adjective that undergoes *'imāla*.¹⁶

do not consider the exceptionality of the words in (11) as necessarily being due to their possible borrowing from Classical Arabic. That said, the issue of application or non-application of *'imāla* to Classical Arabic words needs further study.

15. The transcriptions of the Hebrew source words reflect the intuitions of the first author who speaks an understudied minority dialect of Modern Hebrew. Also, in the transcription, in order to make clear the stress on a long vowel, we transcribe the long vowels in (13) with a colon.

16. A different phenomenon involves English loanwords that end in a consonant but get assigned feminine gender as reflected by the addition of a final vowel when borrowed into the PA-TA dialect under consideration. Such words do seem to undergo *'imāla* and include examples like [dazziin-i] 'dozen' and [sfindʒ-i] 'sponge'. The difference between these words and those that are borrowed with a final [a] that is interpreted as feminine like in (12b)–(12c) is in need of further investigation.

The borrowings from Modern Hebrew in (13) are interesting because the final [-a] of the Hebrew words is also a feminine suffix, which is cognate with the Arabic feminine suffix. Despite the cognate nature of the suffixes *'imāla* fails to occur, and the suffixal vowel of the borrowed noun is realized as low. In addition, if the borrowed Hebrew noun is modified by an Arabic adjective as in the phrase [gliid-a zdiid-i] 'a new ice cream' the adjective undergoes *'imāla* while the modified noun borrowed from Hebrew does not. The understanding of the failure of *'imāla* to the cognate suffix in Hebrew almost certainly relates to the issue of how stress and vowel length are assigned to Hebrew loanwords. As described by Abu Elhija Mahajna (2019), based on her native lower Galilean (Iksal) Palestinian variety of Arabic, but applicable to other varieties spoken by Palestinian citizens of Israel, Hebrew loanwords into Arabic preserve the Hebrew source language stress even if there is an apparent conflict with the dialectal stress pattern. As Abu Elhija Mahajna (2019: 55) notes, a Hebrew word like [mataná] 'gift' with stress on a final short vowel is borrowed into Arabic with the location of source language stress preserved (i.e. on the final syllable), but with lengthening of the final vowel (e.g. as [mataná:] so that the borrowed word can fit into the metrical pattern of Palestinian Arabic. As seen by the data items in (13), words with the cognate feminine ending are almost always borrowed with a final long stress vowel [-á:] in Arabic. We suggest here that it is the realization of the feminine ending as a final long vowel on most Hebrew loanwords that prevents *'imāla* from occurring. In this way, it is similar to the derivation shown in (10), except that in most Hebrew loanwords the final vowel length has not yet been lost, although this may be starting to happen in Hebrew loans like [glí:d-a] in (13b) where final vowel shortening and stress retraction can occur. This retraction might occur with words of higher frequency. A more detailed discussion of this is left for future research.

The exceptional data presented in this section, especially the short nominalizations shown in (11) and the loanword data like in (12b)–(12c) have a theoretical implication for the analysis of the Arabic feminine suffix. In our discussion of *'imāla* so far in this paper we have been making a standard assumption that the underlying form of the suffix is /-a/, and that it raises under certain conditions. An alternative analysis of *'imāla* has been put forward by Younes (1982) and McCarthy (1994) and maintained by Cotter and Horesh (2015) and Cotter (2016) whereby *'imāla* involves a lowering rule in which the feminine suffix is underlyingly /-i/ and lowers to [-a] when the immediately preceding consonant is a guttural. We would contend that the exceptions in (11) reflect /-a/ as a default feminine suffix. Under the lowering perspective of Younes (1982) and McCarthy (1994), the occurrence of a form like [s^ʕif-a] 'characteristic' in (11a) above, would be difficult to explain given that on a lowering analysis it would be underlyingly /s^ʕif-i/. There is no phonological reason for lowering to occur if the underlying form was /s^ʕif-i/ given that the consonant

adjacent to the suffix is a labiodental and not a guttural. As a side point, one might speculate under the lowering analysis, that given an example like /s^ʕif-i/ 'description', the suffix [-i] lowers to [-a] because of the spreading of emphasis (i.e. pharyngealization) from the word-initial emphatic consonant. (The PA-TA dialect under consideration does have rightward spreading of emphasis.) However, as shown by some of the data in (1) and (2), it is only the presence of a final guttural consonant that favors the realization of the suffix [-a]. In a word like [t^ʕooʃ-i] 'fight' in (1i) or [ħiff-i] 'side' in (1j) the presence of a word-initial guttural does not block *'imāla* from applying. Thus, the exceptional words in (11), as well as the loanwords in (12), support a raising analysis as we have maintained in this paper in which /-a/ can be viewed as the underlying (or default) form of the feminine suffix. In essence, the exceptional data in (11)–(13) treat /-a/ as the default form of the suffix.

7. Conclusion

This paper has detailed the final *'imāla* pattern in one variety of Palestinian Arabic. Final *'imāla* is viewed as an instance of vowel raising of the underlying /-a/ of the feminine suffix to [-i]. We also discussed a variety of issues that pertain to final raising and provided various analyses. These include the feature geometric representation of the consonants that block final raising as discussed in Section 3, the patterning of historical uvulars as phonologically uvular even though they surface as velar in Section 4, the analysis of the exceptional morpho-semantic class of colors and physical/mental challenges in Section 5, and the indication of other exceptional classes in Section 6 that include short nominal forms whose initial root consonant is /w/ as in (11a)–(11e) and loanwords as systematic exceptions. Important aspects of our analysis include the role of the Lower Vocal Tract node in understanding why the class of gutturals blocks *'imāla*, the abstract nature of the back velars /x ɣ K/ in PA-TA that still pattern as phonological uvular reflecting their diachronic origins rather than their synchronic pronunciation, the observation that short nouns whose consonantal roots start with an unpronounced /w/, such as [s^ʕif-a] 'characteristic' and [ħib-a] 'gift', systematically fail to undergo *'imāla*, and an explanation for why words borrowed from Modern Hebrew fail to undergo *'imāla* even though the feminine suffix in the two languages are cognate. Our hope is that this paper will serve to foster the systematic analysis of *'imāla* in other dialects so that an understanding of cross-dialectal variation in *'imāla* patterns can be achieved.

Acknowledgments

We thank Enam Al-Wer, Noam Faust, Youssef Haddad, Uri Horesh, and Reem Khamis-Dakwar for discussion on specific aspects of the paper. We also acknowledge the volume editors and anonymous reviewers for their helpful comments. The usual disclaimers apply.

Funding

We are grateful for the Center for the Study of the Middle East and the Islamic Studies Program at Indiana University for their financial support.

References

- Abu Elhija Mahajna, D. 2019. A Study of Loanwords and Code Switching in Spoken and Online Written Arabic by Palestinian Israelis. Ph.D. dissertation, Indiana University, Bloomington.
- Ali, A. 2020. On the realization of the glottal stop in Sudanese Arabic. Paper presented at the 34th Annual Symposium on Arabic Linguistics, University of Arizona, Tucson, February 28 – March 1.
- Al-Wer, E. 1998. Raising of /a/ and related vocalic movements in the emerging dialect of Amman. *Proceedings of the Third International Conference of AIDA*, 29, 77–82.
- Al-Wer, E. 2007. The formation of the dialect of Amman: From chaos to order. In Catherine Miller, Enam Al-Wer, Dominique Caubet, & Janet Watson (eds.), *Arabic in the City: Issues in Dialect Contact and Language Variation*, (pp. 55–76). London: Routledge.
- Al-Wer, E., & Herin, B. 2011. The lifecycle of Qaf in Jordan. *Langage et Société*, 4, 59–76. <https://doi.org/10.3917/ls.138.0059>
- Al-Wer, E., Horesh, U., Herin, B., & Fanis, M. 2015. How Arabic regional features become sectarian features. Jordan as a case study. *Zeitschrift für Arabische Linguistik*, (62), 68–87.
- Bellem, A. 2007. Towards a Comparative Typology of Emphatics. PhD dissertation, SOAS, University of London.
- Bin-Muqbil, M. 2006. Phonetic and Phonological Aspects of Arabic Emphatics and Gutturals. Ph.D. dissertation, University of Wisconsin, Madison.
- Booij, G. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Cotter, W. M. 2016. A sociophonetic account of morphophonemic variation in Palestinian Arabic. In *Proceedings of Meetings on Acoustics 171ASA*, 26(1), 060001. <https://doi.org/10.1121/2.0000213>
- Cotter, W., & Horesh, U. 2015. Social integration and dialect divergence in coastal Palestine. *Journal of Sociolinguistics*, 19(4), 460–483. <https://doi.org/10.1111/josl.12135>
- Davis, S. 1995. Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry*, 26, 465–498.
- Durand, E. P. 2011. Word-Final *imaala* in Contemporary Levantine Arabic: A Case of Language Variation and Change. MA thesis, the University of Texas at Austin.
- Grano, T. & Davis, S. 2018. Universal markedness in gradable adjectives revisited. *Natural Language & Linguistic Theory*, 36, 131–146. <https://doi.org/10.1007/s11049-017-9365-0>

- Habib, R. 2012. 'Imala and rounding in a rural Syrian variety: Morpho-phonological and lexical conditioning. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 57(1), 51–75.
- Haddad, G. 1984. Problems and Issues in the Phonology of Lebanese Arabic. Ph.D. dissertation, University of Illinois at Urbana-Champaign, Urbana.
- Haspelmath, M. & Sims, A. 2013. *Understanding Morphology* (2nd edition). London and New York: Routledge. <https://doi.org/10.4324/9780203776506>
- Hayes, B. 1986. Inalterability in CV phonology. *Language* 62, 321–351. <https://doi.org/10.2307/414676>
- Herzallah, R. 1990. Aspects of Palestinian Arabic Phonology: A Non-linear Approach. Ph.D. dissertation, Cornell University, Ithaca.
- Khalafallah, A. A. 1969. *A Descriptive Grammar of Sa'idi Egyptian Colloquial Arabic*. Janua Linguarum Series Practica 32, de Gruyter Mouton. <https://doi.org/10.1515/9783111357355>
- Levin, A. 2011. 'imāla. In *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong. Brill Online.
- McCarthy, J. 1994. The phonetics and phonology of Semitic pharyngeal. In Patricia Keating (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, (pp. 191–233). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511659461.012>
- McCarthy, J. 2005. The length of stem-final vowels in Colloquial Arabic. In M. Alhawary & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XVII–XVIII* (p. 1–26). Amsterdam/Philadelphia: John Benjamins Publishing Co.. <https://doi.org/10.1075/cilt.267.03mcc>
- Owens, J. 2006. *A Linguistic History of Arabic*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199290826.001.0001>
- Shachmon, O. 2011. Pausal final 'imāla in central Palestinian dialects. *Jerusalem Studies in Arabic and Islam*, 38, 145–162.
- Shachmon, O., & Faust, N. 2017. The spread of prepausal lowering in Galilean Arabic. *Journal of Semitic Studies*, 62(2), 371–387. <https://doi.org/10.1093/jss/fgx008>
- Shahin, K. 2012. Palestinian Arabic. In: *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong. Brill Online.
- Sibawayhi, *Kitāb* = 'Abū Bišr 'Amr ibn 'Uṭmān Sibawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881–1889.
- Sylak-Glassman, J. 2014. Deriving Natural Classes: The Phonology and Typology of Post-Velar Consonants. Ph.D. dissertation, University of California, Berkeley.
- Vaux, B. 1993. *Is Atr a laryngeal feature? Atr-voice interactions and the geometry of the lower vocal tract*. Ms., Harvard University, Cambridge, Mass.
- Wright, W. 1896/1974. *A Grammar of the Arabic Language* (3rd edition, 1974, impression). Beirut: Librairie du Liban.
- Younes, M. A. 1982. Problems in the Segmental Phonology of Palestinian Arabic. Ph.D. dissertation, The University of Texas, Austin.
- Zawaydeh, B. A. 1999. The Phonetics and Phonology of Gutturals in Arabic. Ph.D. dissertation, Indiana University, Bloomington.

A study of the place of articulation of the Arabic voiceless dorsal fricative

Noor Abo Mokhⁱ, Steven M. Lulichⁱⁱ, Abdullah Alfaifiⁱ,
Sarah Robinsonⁱ, Sherman Charles^{i,ii} and Kenneth De Jongⁱ

ⁱDepartment of Linguistics, Indiana University / ⁱⁱDepartment of Speech,
Language and Hearing Sciences, Indiana University

Previous descriptions of the place of articulation of the Arabic dorsal fricative suggest that it ranges from velar to uvular, and there remains uncertainty in terms of identifying the place of articulation of this fricative. If the dorsal fricatives are velar in articulation, this presents a mismatch between their phonological class membership and their phonetic manifestation, since the velar-uvular split is precisely where the guttural class boundary is. If the natural class of gutturals excludes velars, then the common patterning of the dorsal fricative with the gutturals suggests that it is not underlyingly velar. This study seeks to empirically determine the place of articulation of the voiceless dorsal fricative (/x/ or /χ/) in a speaker of each of the following six Arabic dialects: Moroccan Darija, Algerian Darija (West side of Algeria), Egyptian (Cairene), Triangle-Area Palestinian, Syrian (Aleppo), and Faiḍi (Saudi Arabia). The study uses a state-of-the-art 3D/4D ultrasound machine, digitized 3D palate impressions, and audio recordings to analyze the dorsal fricatives of Arabic speakers. The results show a general tendency toward uvular articulation or pre-uvular articulation across dialects, with some variation in some speakers.

Keywords: Arabic dialects, gutturals, dorsal fricative, articulatory phonetics, tongue ultrasound

1. Introduction

Previous descriptions of the place of articulation of the Arabic dorsal fricative, which were mostly impressionistic, suggest that it ranges from velar to uvular. Sibawayh's traditional description, for instance, classifies the dorsal fricative within the same category as /β/ and /q/, leading to its typical transcription as a voiceless uvular fricative /χ/ (Al-Nassir, 1985). More recent studies show contradictory

descriptions of this fricative even within the same Arabic varieties. Watson (2002) and Zawaydeh (1997) argue that the Standard Arabic voiceless uvular fricative /χ/ has become either velar or post-velar (in Cairene Arabic and some Palestinian dialects). This variation is also recorded for regional varieties spoken within the same dialect. In Najdi dialect, for instance, Abboud (1979) describes this fricative as post-velar, while Ingham (2008) describes it as uvular. In Modern Standard Arabic, Saiegh-Haddad (2003) describes it as uvular /χ/, while McCarus (2008) describes it as velar /x/. For the Turkish (Jastrow, 2006), Baghdadi (Abu-Haidar, 2006), Bahraini (Holes, 2006), Cairene (Woidich, 2006), and Lebanese (Feghali, 1999) dialects, the dorsal fricative is described as velar. These studies show that there is uncertainty in terms of identifying the place of articulation of this fricative.

This variation is not very surprising. Previous literature studying Semitic languages has traditionally grouped together uvulars, laryngeals and pharyngeals as one natural class, namely the guttural class, as in McCarthy (1991). The guttural class includes sounds with a primary place of articulation, or constriction, near the back of the vocal tract. McCarthy (1994) further argues that Modern Standard Arabic and most Arabic dialects retain the full set of gutturals which consists of uvulars (/χ/), pharyngeals (/ħ/ and /ʕ/), and laryngeals (/h/ and /ʔ/). However, there is a disagreement on whether velar fricatives belong to the same natural class. Some studies have indicated that velars belong to the guttural class since they have the pharyngeal node in their feature geometry representation. This was exemplified by McCarthy (1994) and Davis (1995), where they argued that the velar sounds in a Palestinian variety discussed in Herzallah (1990) are underlyingly uvular, since they interact with other phonological phenomena, such as *'imala*. This phenomenon involves a lowering rule in which the feminine suffix is underlyingly /-i/ and lowers to [a] except in cases where it is preceded by a pharyngeal, a laryngeal, a velar or a uvular (also discussed in Davis (1995), and Zawaydeh (1999)).

If the dorsal fricatives are velar in articulation, this presents an interesting mismatch between the phonological class and the phonetic manifestation of the class. The velar-uvular split is precisely where the guttural class boundary is. If the natural class of gutturals excludes velars, then the common patterning of the dorsal fricative with the gutturals suggests that it is not underlyingly velar. Also relevant to this mismatch is the role of contrast. While there is a velar-uvular contrast in stops (/k/ vs. /q/), there is only one dorsal fricative (the Consonant Inventory of Arabic is attached in Appendix 1). Without contrast, it is possible that the fricative place of articulation might freely vary, giving rise to regional and dialectal differences and contributing to the confusion apparent in the relevant phonetics and phonological literature, and also, then, giving rise to a mismatch between the phonological classification and the phonetic content of the dorsal fricative segment.

This study seeks to empirically determine the place of articulation of the voiceless dorsal fricative (/x/ or /χ/) in a speaker of each of the following six Arabic dialects: Moroccan Darija, Algerian Darija (West side of Algeria), Egyptian (Cairene), Triangle-Area Palestinian, Syrian (Aleppo), and Faiifi (Saudi Arabia). The study uses a state-of-the-art 3D/4D ultrasound machine, digitized 3D palate impressions, and audio recordings to analyze the dorsal fricatives of Arabic speakers. The results show a general tendency toward uvular articulation or pre-uvular articulation across dialects, with some variation in some speakers. For example, the Moroccan speaker consistently articulated the dorsal fricative with a place of maximal constriction very similar to the place of articulation of uvular stops. For the Algerian speaker, by contrast, there was a wide range of variation, with most articulations of the dorsal fricative being anterior to those of the uvular stop. To quantify this variation, we estimated the relative distance of the dorsal fricative tongue posture to the uvular and velar stop articulations, verifying the general observations from the ultrasound images themselves.

In Section 2 we present the methodology, providing details on the participants and experimental design. Next, we provide the results of the study (Section 3), followed by analysis and discussion of the results (Section 4).

2. Methods

2.1 Participants

The data were collected from six native speakers (three males, three females). The data collection process took place in the Speech Production Laboratory in the Department of Speech and Hearing Sciences at Indiana University, Bloomington. The speakers' ages ranged from 26 to 40 years. Each of the participants grew up speaking one of the following six dialects: Moroccan Darija (Fez, Morocco), Algerian Darija (West Algeria), Egyptian (Cairene, Cairo), Palestinian (Triangle-Area, Israel), Syrian (Aleppo, Syria), and Faiifi (South West Saudi Arabia). In addition to their native dialects, the participants had been exposed to other dialects, as well as Modern Standard Arabic (MSA). The speakers reported no history of hearing or speaking impairments. All participants had experience with English as a second language, in addition to other languages. In this paper, participants will be referred to by the name of their dialect.

2.2 Stimuli

The stimuli included real words in which dorsal fricatives appeared in a variety of phonetic environments. In addition, real words containing the palatal glide /j/, the voiceless velar stop /k/, the voiceless uvular stop /q/, and the voiced pharyngeal fricative /ħ/ were recorded as reference words in order to compare the dorsal fricatives to sounds with contrasting places of articulation. In the analysis, we focus on comparing the place of articulation of the dorsal fricatives with the velar and uvular stops, /k/ and /q/.

Due to lexical and phonetic differences across dialects and subdialects of Arabic, one stimulus list was designed for the Egyptian, Palestinian, Syrian and Faifi speakers, and separate stimulus lists were designed for the Moroccan and Algerian speakers. All of the stimuli, regardless of the speaker's dialect, had similar vowel environments, usually the low front vowel /a/. The stimuli were confirmed by additional native speakers of the corresponding dialects, who did not serve as subjects for the experiment.

The following tables show the reference words for the Egyptian, Palestinian, Syrian, and Faifi speakers (Table 1), the Moroccan speaker (Table 2), and the Algerian speaker (Table 3).

Table 1. PoA reference words for the Egyptian, Faifi, Syrian, and Palestinian speakers. The vowels in [jVdrVs] were [i] or [u], depending on the dialect

| | Place of articulation | Transcription | Gloss |
|----|-----------------------|---------------|---------------|
| a. | Palatal | [jVdrVs] | 'he studies' |
| b. | Velar | [ka:mil] | 'complete' |
| c. | Uvular | [qa:ʕa] | 'hall, class' |
| d. | Pharyngeal | [ħa:zim] | 'strict' |

Table 2. PoA reference words for the Moroccan speaker

| | Place of articulation | Transcription | Gloss |
|----|-----------------------|---------------|---------------|
| a. | Palatal | [jVqra] | 'he studies' |
| b. | Velar | [ka:mil] | 'complete' |
| c. | Uvular | [qa:ʕa] | 'hall, class' |
| d. | Pharyngeal | [ħa:ɗ̣ˤa] | 'old lady' |

Table 3. PoA reference words for the Algerian speaker

| | Place of articulation | Transcription | Gloss |
|----|-----------------------|---------------|---------------|
| a. | Palatal | [jVqra] | 'he studies' |
| b. | Velar | [ka:mil] | 'complete' |
| c. | Uvular | [qa:ʕa] | 'hall, class' |
| d. | Pharyngeal | [ħa:mi] | 'hot' |

Tables 4 (the Egyptian, Palestinian, Syrian, and Faifi speakers), 5 (the Moroccan speaker), and 6 (the Algerian speaker) provide the Arabic words including the dorsal fricatives in four different environments for each one of the speakers: word-initially, intervocalically, word-finally, and geminated. We hypothesized that these different environments might affect the articulation of the dorsal fricatives.

Table 4. PoA target words for the Egyptian, Faifi, Syrian, and Palestinian speakers

| | Transcription | Gloss |
|---------------------|---------------|--------------------------|
| a. Word-initially | [xa:lɪd] | <i>'eternal, Khalid'</i> |
| | [xa:ɪmɪd] | <i>'still, inactive'</i> |
| b. Intervocalically | [ʔaxaf] | <i>'lighter than'</i> |
| | [ʔaxað] | <i>'he took'</i> |
| c. Word-finally | [fax] | <i>'trap'</i> |
| | [matˈbax] | <i>'kitchen'</i> |
| d. Geminated | [ʔaxxar] | <i>'he delayed'</i> |
| | [baxxar] | <i>'he steamed'</i> |

Table 5. PoA target words for the Moroccan speaker

| | Transcription | Gloss |
|---------------------|---------------|--------------------------|
| a. Word-initially | [xa:led] | <i>'eternal, Khalid'</i> |
| | [xa:jb] | <i>'disappointment'</i> |
| b. Intervocalically | [ʔaxaff] | <i>'lighter than'</i> |
| | [ʔaxað] | <i>'he took'</i> |
| c. Word-finally | [fax] | <i>'trap'</i> |
| | [fsax] | <i>'broke (v)'</i> |
| d. Geminated | [jsaxxar] | <i>'to give'</i> |
| | [baxxar] | <i>'he steamed'</i> |

Table 6. PoA target words for the Algerian speaker

| | Transcription | Gloss |
|---------------------|---------------|--------------------------|
| a. Word-initially | [xa:led] | <i>'eternal, Khalid'</i> |
| | [xa:ɟef] | <i>'afraid'</i> |
| b. Intervocalically | [maxadda] | <i>'pillow'</i> |
| | [ʔaxað] | <i>'he took'</i> |
| c. Word-finally | [fax] | <i>'trap'</i> |
| | [fsax] | <i>'broke'</i> |
| d. Geminated | [waxxar] | <i>'he delayed'</i> |
| | [baxxar] | <i>'he steamed'</i> |

2.3 Experimental procedure

One challenge when designing experiments of Arabic vernaculars is accounting for the linguistic differences between MSA and Arabic vernaculars, known as diglossia. Diglossia is a term used to describe a language that has two varieties used by its speakers under two different communicative conditions (Ferguson, 1959). In the case of Arabic, the dialects are not taught or written; they are rather used as the main means of communication. On the other hand, MSA is taught in schools, and it is the language of news and formal communications (sometimes also used by speakers of North African varieties to communicate with speakers of other dialects). Any written form of Arabic will therefore usually be associated with MSA. When performing the experiments, we asked the participants to stay as faithful as possible to their dialects, and to ignore the spelling of the words. Prior to the experiment, we tested the stimuli on another participant (who did not take part in this study) in order to see how bias toward MSA would affect speakers when presented with the written stimuli. Similarly, here we asked the participant to produce the words in his own dialect. The test was successful, in that the participant was able to ignore MSA associations with the written stimuli, and instead produced the words in his own dialect.

Palate impressions for each participant were made using dental impression alginate. The impressions were then digitized with a NextEngine 3D laser scanner. These palate impressions provide anatomical information about the point of contact between the tongue and the palate and help contextualize the tongue shape data within the participants' vocal tracts. An ultrasound machine was then used to capture the tongue contours. The ultrasound images were recorded using a Philips EPIQ 7G ultrasound system with an xMatrix x6-1 digital 3D/4D transducer. The transducer was placed under the chin using an Articulate Instruments ultrasound stabilization headset (Scobbie, Wrench, and van der Linden, 2008). Recording rates varied between 9 and 16 volumes per second. Subjects were seated in a sound-treated room, and were presented with each stimulus written in Arabic script and viewed on a computer monitor. The participants read the four reference words first and then they were asked to read the 8 target words. The process was repeated twice for each participant. The stimuli were recorded using a microphone placed in front of the speakers. The total number of tokens produced was 216.

2.4 Analysis

Ultrasound files and audio recordings were exported along with the digitized palate impressions to a Windows computer for analysis. An analysis of the data was performed using WASL, a MATLAB toolbox for multimodal speech research developed in the Speech Production Laboratory at Indiana University. For each

speaker, the ultrasound, audio, and palate data were aligned using the stop consonants as an indication of tongue-palate contact. Then using a tracing feature within WASL, sagittal ultrasound images of the reference and target sounds were manually traced. The traced mid-sagittal images were plotted for visualization and analysis using MATLAB.

3. Results

3.1 Ultrasound traces – 2D midsagittal profiles

Figure 1 shows 2D midsagittal tongue contours for the Moroccan reference sounds. Contact between the tongue and palate does not appear to be complete because the fast-moving closure gesture was not fully captured at the recorded ultrasound frame rates. Similar figures were made for each speaker.

While all reference sounds are interesting, the velar and uvular stops were of special interest due to their relationship to the two predicted places of articulation of the target sound. As can be seen in Figure 1, for the Moroccan speaker, the uvular stop (green) has a more posterior constriction to the velar stop (brown). In all participants, the uvular reference words showed a shallow curve indicating a point of contact with the soft palate or uvula, while the same speaker's velar reference word showed a steeper curve and a more anterior point of contact.

Figure 2 shows the 2D images of the dorsal fricative tested in all 8 words for the Moroccan (left), the Algerian (center), and the Egyptian (right) speakers. For each one of the participants, the mid-sagittal image of the two (velar = thick brown line; uvular = thick green line) reference words is plotted along with the palate

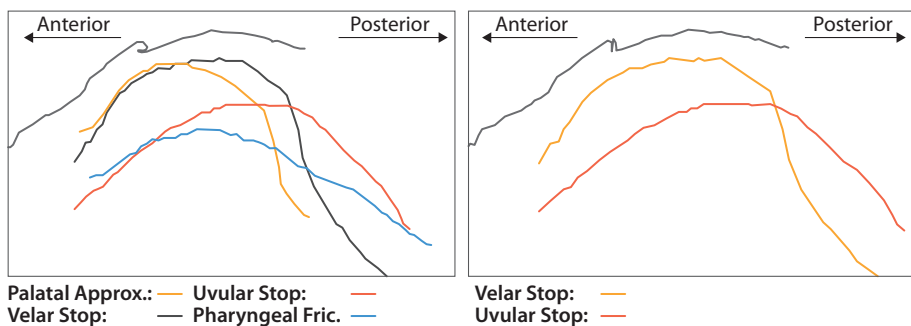


Figure 1. Moroccan speaker – PoA references: Velar and uvular stops; the palate appears as a black line above the tongue traces. The palatal glide is red, the velar stop is brown, the uvular stop is green and pharyngeal fricative is blue

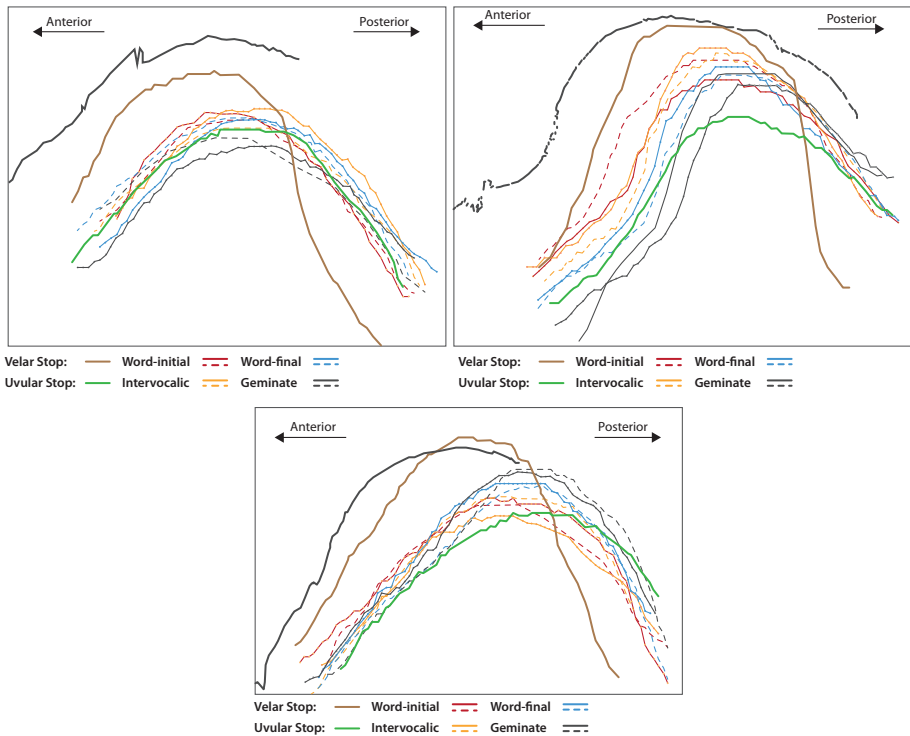


Figure 2. To the far left is the 2D image of the Moroccan speaker, next is the Algerian speaker, and below is the Egyptian speaker – Dorsal fricatives

(black line) and the dorsal fricatives. The dorsal fricatives in each of the different environments is shown in a different color (the solid line is for the first repetition, and the dashed line is for the second repetition). The pink line shows the dorsal fricative tested word-initially, the light green line shows the fricative intervocalically, the blue line shows the fricative word-finally, and the orange line shows the fricative geminated.

Speakers showed remarkable consistency in articulation, with the dorsal fricative usually aligning with the uvular stop. For the Moroccan speaker, for instance, there is a consistent articulation of the dorsal fricative which aligns consistently with the uvular stop and not the velar stop, which is more anterior to the fricatives. This consistency follows across all tested environments, although there is some variation within this individual. For instance, the intervocalic dorsal fricative has a higher tongue position than the geminated dorsal fricative.

For the Algerian speaker, unlike the Moroccan speaker, the fricative constriction is more variable, and the tongue surfaces for the target sounds fall somewhere between the velar and uvular stops. In general, the tongue shapes for the target sounds are more similar to the uvular stop than to the velar stop.

For the Egyptian speaker the tongue is closely aligned with the uvular stop, albeit with some variation, similar to the Moroccan speaker.

Figure 3 shows the 2D images of the Syrian, Palestinian, and Faifi speakers. The three speakers show consistency in the alignment of the dorsal fricatives with the uvular stop, with some variation.

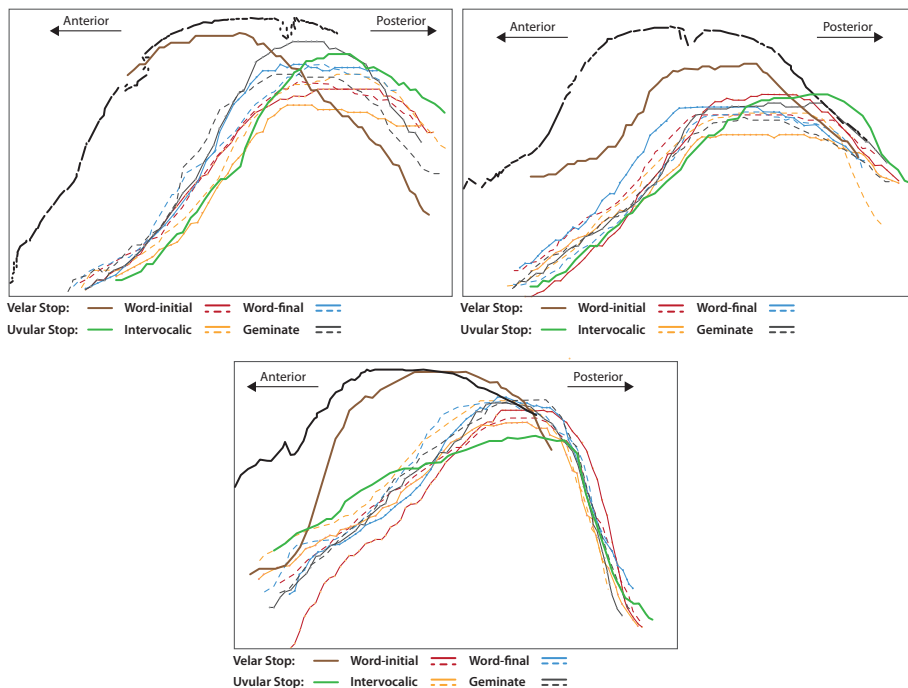


Figure 3. To the far left is the 2D image of the Syrian speaker, next is the Palestinian speaker, and below is Faifi speaker – Dorsal fricatives

3.2 Nearest Neighbor Distance (NND) from Velar and Uvular Stops to Dorsal Fricatives

The results described thus far are all qualitative and based on visual inspection of the tongue shapes. The advantage of this type of qualitative visualization is that it captures the variations within individuals while still showing the overall tendency of all speakers to prefer a uvular constriction. To quantify the similarity between tongue shapes for the dorsal fricatives and both the velar and uvular stops, we used the nearest-neighbor distance measure, NND (Zharkova et al., 2009, Csapó et. al, 2017). The NND measure was used to quantify average distances between pairs of individual tongue shapes, which were represented as x and y values. The NND is not a symmetric measure, i.e. the distance from the dorsal fricative to the uvular

stop is not the same as the distance from the uvular stop to the dorsal fricative. This asymmetry is illustrated with a simple example in Figure 4.

We therefore calculated the NND twice, from the dorsal fricative to the velar stop, and from the dorsal fricative to the uvular stop. The NND values for the fricative-to-velar stop distance were normalized by the NND from the uvular stop to the velar stop, while NND values for the fricative-to-uvular stop distance were normalized by the NND from the velar stop to the uvular stop. Thus, normalized NND values close to 1 indicate close similarity with either the uvular or velar stops, while NND values close to zero indicate poor similarity.

We then plotted the results on a graph, as shown in Figure 5, where the x-axis represents the distance to the uvular stop, and the y-axis represents the distance to the velar stop. The dashed line indicates equal distances. Thus, points that fall below the dashed line are those for which the dorsal fricative is more similar to the velar stop, and points that fall above the dashed line are those for which the dorsal fricative is more similar to the uvular stop. Based on the data provided in Figure 5, the qualitative analysis of the sagittal profiles is in close agreement with velar/uvular similarity using quantitative NND measures. For instance, for the Moroccan speaker the points are clustered closer to the 1 on the y-axis, indicating a more uvular, or less-anterior articulation than the velar sound. Comparing this to the Algerian speaker, the plot corresponds neatly with the variation observed in the 2D images: while the fricatives are clustered at a more anterior point than uvular, they are generally more similar to the uvular stop than to the velar stop, with some variability. In each of the remaining four speakers (Egyptian, Palestinian, Syrian, Faifi) the dorsal fricative is quantitatively more similar to the uvular stop than to the velar stop.

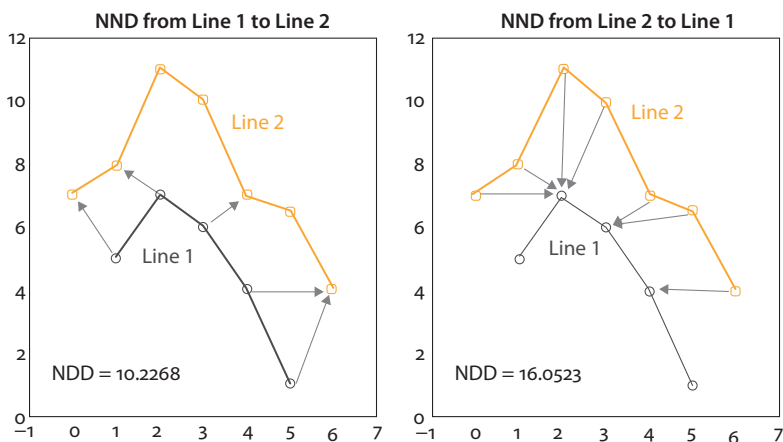


Figure 4. Illustration of the asymmetrical nature of NND calculations using two simple lines

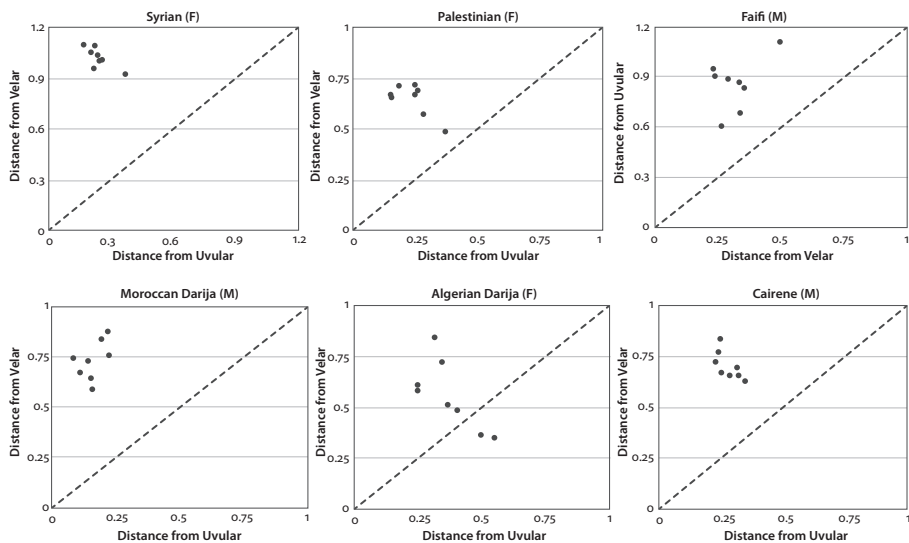


Figure 5. Distance from velar and uvular stops Nearest Neighbor Distance (NND)

4. Discussion and conclusions

Figures 3 and 4 give a clear perspective on the variation of place of articulation among the six participants. Our main research question was: What is the nature of the dorsal fricative in six Arabic dialects (Faifi, Algerian, Egyptian, Syrian, Moroccan, Palestinian)? Following both qualitative and quantitative analyses for all six speakers, we conclude that the dorsal fricative is produced with a posterior place of articulation similar to uvular stops. It is noteworthy, that there is some variation among speakers, especially for the Algerian speaker. The data also showed that none of the speakers produced the dorsal fricative with a velar place of articulation.

To conclude, while there were some variations, there is a strong tendency across the speakers in this study, who represent six geographically distinct dialects, to favor a uvular articulation for the dorsal fricative. The most uvular articulation was recorded for the Moroccan speaker, and the least uvular articulation was for the Algerian speaker, suggesting that production of this fricative does not vary along a geographical continuum, although further studies are needed with greater numbers of speakers within each dialect before strong conclusions regarding dialect differences can be drawn.

The results may reflect the fact that the fricatives do not contrast velar and uvular places of articulations, as do the stops. Given the lack of contrast, the dorsal fricative is free to encroach upon the velar region, as is most clearly seen for the Algerian speaker. The fact that the dorsal fricative patterns with the gutturals is consistent with its overall similarity to the uvular stop.

One of the limitations of this work is the design of the stimuli and testing procedures. Arabic dialects have many linguistic variations. Lexical variations created challenges in some cases, because it was difficult to develop examples with similar phonetic environments. Therefore, we had to consult with native speakers of the tested dialects in order to control the stimuli. Another crucial factor was the number of speakers. We plan to expand this study by recruiting more speakers from the same dialects, and from additional dialects such as Upper Egyptian, Libyan, and other Gulf dialects. This will allow further examination of the dorsal fricative in terms of variation across speakers of the same dialects. We are also planning to vary the vowel environment of the dorsal fricative, e.g. using the high front vowel /i/, to investigate its effects on dorsal fricative articulation, and to examine the patterns of variation with more imposing sources of variation. Similarly, emphatic consonants will also be added to the stimuli to examine the spread of emphasis and its influence on the dorsal fricative. In addition, we plan to carry out acoustic analyses of velar and uvular stops, and the dorsal fricatives, to see how the differences in tongue position affect the acoustics.

References

- Abu-Haidar, F. 2006. Baghdad Arabic. *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong, 1, 222–231. Leiden – Boston: Brill.
- Abboud, P. F. 1979. The verb in Northern Najdi Arabic. In *BSOAS*, 42, ii, 467–99.
<https://doi.org/10.1017/S0041977X0013575X>
- Al-Nassir, A. A. 1985. Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise al-Kitab. (Doctoral dissertation). University of York, Heslington. Bouamor, Houda, Nizar Habash, and Kemal Oflazer. A Multidialectal Parallel Corpus of Arabic. LREC. 2014.
- Csapó, T. G., Deme, A., Gráci, T. E., & Markó, A. 2017. Comparison of distance measures in tongue contour traces of ultrasound images. Poster presented at Ultrafest VIII, Potsdam, Germany.
- Davis, S. 1995. Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry*, 465–498.
- Feghali, M. N. 1999. *Spoken Lebanese*. Parkway Publishers, Inc., Box 3678, Boone, NC.
- Ferguson, Charles A. 1959. Diglossia. *Word* 15.2 325–340.
<https://doi.org/10.1080/00437956.1959.11659702>
- Herzallah, R. 1990. Aspects of Palestinian Arabic phonology: A non-linear approach. Doctoral dissertation, Cornell University, Ithaca, N.Y.
- Holes, C. 2006. Bahraini Arabic. *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong, 1, 241–255. Boston, MA: Brill.
- Ingham, B. 2008. Najdi Arabic. *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong, 3, 326–334. Boston, MA: Brill.
- Jastrow, O. 2006. Anatolian Arabic. *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong, 1, 87–96. Boston, MA: Brill.

- McCarthy, J. J. 1991. Semitic gutturals and distinctive feature theory. *Perspectives on Arabic Linguistics*, 3, 63–91. <https://doi.org/10.1075/cilt.80.o6mcc>
- McCarthy, John. (1994). The phonetics and phonology of Semitic pharyngeal. In Patricia Keating (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511659461.012>
- McCarus, E. 2008. Modern Standard Arabic. *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong, 3, 238–262. Boston, MA: Brill.
- Saiegh-Haddad, E. 2003. Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, 24 (03). <https://doi.org/10.1017/S0142716403000225>
- Scobbie, J. M., Wrench, A. A., van der Linden, M. 2008. Head-Probe stabilization in ultrasound tongue imaging using a headset to permit natural head movement. *Proceedings of the 8th International Seminar on Speech Production*, 373–376.
- Watson, J. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
- Woidich, M. A. 2006. Cairo Arabic. *Encyclopedia of Arabic Language and Linguistics*, Managing Editors Online Edition: Lutz Edzard, Rudolf de Jong, 1, 323–333.
- Zawaydeh, B. 1997. An acoustic analysis of uvularization spread in Ammani-Jordanian Arabic. *Studies in the Linguistic Sciences* 27.1, 185–200.
- Zawaydeh, B. 1999. The phonetics and phonology of gutturals in Arabic. Doctoral dissertation, Indiana University Bloomington.
- Zharkova, N., & Hewlett, N. 2009. Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /a/. *Journal of Phonetics*, 37 (2), 248–256. <https://doi.org/10.1016/j.wocn.2008.10.005>

Appendix 1. Phoneme inventory of Arabic

| | Labial | Dental | Alveolar | Emphatic | Palatal | Velar | Uvular | Pharyngeal | Glottal |
|-----------|--------|--------|----------|-------------------------------|---------|-------|--------|------------|---------|
| nasal | m | | n | | | | | | |
| stop | b | | d t | d ^ʕ t ^ʕ | ð̤ | k | q | | ʔ |
| fricative | f | ð θ | z s | ð ^ʕ s ^ʕ | ʃ | χ x | | ħ | h |
| tap | | | r | | | | | | |
| approx. | | | l | | j | w | | | |

PART II

Syntax and semantics

On the syntax of correlation

Evidence from Egyptian Arabic

Usama Soltan

Middlebury College

The goal of this paper is to describe and analyze the syntax of two types of correlation structures (CSs) in Egyptian Arabic (EA): A structure introduced by the quantifier *kull* ('every'), and another introduced by the measure PP *ṣalā ʔadd* (literally, 'on/by amount'). More specifically, it is shown that both types of structures involve subordination, where the first clause is subordinate to the second. A syntactic analysis along the lines of den Dikken (2005) and Taylor (2013) for comparative correlatives is first shown to account for the main grammatical properties of both types of CSs in EA. A cartographic implementation of the analysis is then proposed to account for a wider range of empirical data related to CSs, including their information structure properties. The data and analysis presented in this paper thus add to the body of empirical literature on CSs cross-linguistically, while showing that their grammatical properties do follow under standard syntactic analysis, and that analyzing such structures as 'constructional' primitives of the grammar is unnecessary.

Keywords: correlation structures, comparative correlatives, left-periphery, clausal cartography, Arabic syntax, Egyptian Arabic

1. Introduction

There are a variety of grammatical structures that human languages utilize to express correlation between two properties or two situations. For example, one can use conditional or temporal clauses, as in (1) from English.

- (1) a. If you study linguistics, you'll find it fascinating.
- b. When it is hot, I stay home.

The correlation can also involve proportionality in degree, as in (2).

- (2) If/When you read more, you'll understand more.

Significantly notable in this regard is that human languages seem to have a structure whose specific function is to express correlation, particularly correlation in degree.¹ In English, this structure has the schema in (3), and is illustrated by examples such as those in (4).

- (3) [[the + comparative form ...], [the + comparative form ...]]
- (4) a. The more I study linguistics, the more fascinating it becomes.
b. The hotter it is, the more likely I will stay home.

In the relevant literature, the structure has been referred to by different names over the years. One notable name is *the comparative conditional*, as in Fillmore (1987), McCawley (1988), and Beck (1997), among others. Since Culicover and Jackendoff (1999), the structure has come to be known as the *comparative correlative*, and its two component clauses are referred to as C1 and C2, as illustrated by the English example in (5) below.²

- (5) [_{C1} The more you read], [_{C2} the more you understand].

Often cited for their grammatical idiosyncrasies, these structures have been argued to be underivable under standard assumptions of syntactic theory, and, as such, to provide evidence for the need to posit ‘constructions’ as syntactic primitives in the grammar (Fillmore 1987; McCawley 1988; Culicover and Jackendoff 1999; see also Borsley 2003, 2004, 2011, for a Head-driven Phrase Structure Grammar approach to comparative correlatives). An alternative approach has been to attempt to explain away these syntactic idiosyncrasies by analyzing their macro-syntax and micro-syntax within a theory of syntax that is strictly compositional, where lexical and functional morphemes are the only building blocks for the combinatorial rules of syntactic structure, as argued for in den Dikken (2005) and Taylor (2013), among a few others. If correct, this latter approach dispenses with any need to posit ‘form-meaning constructions’ as primitives in language design.

1. Taylor (2013) provides a list of languages for which a comparative correlative (or a *correlation structure*, as it will be referred to in this paper) has been identified, either in the literature or via Taylor’s inquiries with native speakers of a variety of different languages. A potential exception in this regard might be languages with *exceed*-comparatives, based on informal conversations I had with speakers of three such languages. If it turns out that *exceed*-comparative languages do indeed lack comparative correlatives, this raises the interesting question as to why this is so and whether it can be shown to follow from the general behavior of comparatives in such languages.

2. Other names are ‘*the ... the ...*’ or ‘*the Xer ... the Xer ...*’ construction. The structure has also been always discussed in descriptive grammars of many languages. In the generative literature, it seems that its first mention was in Ross (1967) in the context of his discussion of island constraints on movement.

Building on the basic discussion in Soltan (2019), the goal of this paper is to describe and analyze two types of structures that express correlation in Egyptian Arabic (EA): A structure introduced by the quantifier *kull* (= ‘every’), and another introduced by the measure PP *ṣalā ʔadd* (literally, = ‘on/by amount’).³ More specifically, I show that both types of correlation structures involve subordination, where C1 is an adverbial subordinate clause and C2 is a main clause. I then show that a syntactic analysis along the lines proposed by den Dikken (2005) and Taylor (2013) for comparative correlatives can be extended to account for the main grammatical properties of the EA structures. To overcome some empirical and theoretical challenges facing that analysis, I then proceed to implement it within the cartographic approach to the left periphery of clause structure, first proposed in Rizzi (1997). The data and analysis presented in this paper thus add to the body of empirical literature on correlation structures, while showing that their grammatical properties do follow under standard syntactic analysis, and that analyzing such structures as ‘constructional’ primitives of the grammar is unnecessary (cf. footnote 3).

Before I proceed to the next section, an important terminological point is in order. Since, as will be illustrated momentarily, a comparative morpheme is not always required in correlation structures in EA, the term ‘comparative correlative’ becomes rather inadequate, descriptively speaking. I have, therefore, chosen to refer to the structures under study with the more generic term *correlation structures* (CSs, henceforward), a subtype of which is the standard comparative correlative represented in English by examples such as those in (4) and (5) and in EA by what will be referred to in the paper as *scalar CSs*.⁴

The paper is organized as follows: In Section 2, I describe two types of CSs in EA and the grammatical properties associated with each. In Section 3, I discuss diagnostic tests showing that the macro-syntax of both types of CSs involves subordination, where the first clause is subordinate to the second. In Section 4, I offer a syntactic analysis of the macro-syntax and micro-syntax of both types of CSs in light of the main insights of den Dikken (2005) and Taylor (2013). In Section 5,

3. Alqurashi and Borsley (2014) provide an analysis of the basic facts of comparative correlatives in Modern Standard Arabic cast within the framework of Head-driven Phrase Structure Grammar, challenging the approach advocated in den Dikken (2005) regarding cross-linguistic variation in comparative correlatives. As will be discussed in this paper, the analysis proposed here, which shares the main insights of den Dikken’s (2005) and Taylor’s (2013) accounts, shows that such structures in Arabic dialects do indeed follow from a theory of grammar where heads are the only atoms in the syntax and where invoking ‘constructions’ is shown to be unnecessary.

4. While the term ‘correlative’ could have been an appropriate label, it is already used in the literature to refer to structures with a sentence-initial relative clause in languages like Hindi-Urdu, as first discussed by Srivastav (1991). I provide an illustrating example of a Hindi-Urdu correlative structure in Section 4 of this paper.

I point out a number of challenges with the proposed analysis, and argue that a cartographic implementation of the analysis overcomes these challenges and also promises to account for the information structure properties of CSs. Section 6 includes some brief remarks on the implications of the analysis for the debate on the place of ‘constructions’ in the theory of grammar. Section 7 sums up the conclusions of the paper.

2. The grammatical properties of correlation structures in EA

As Taylor (2013) points out, in the relevant literature, CSs have been characterized by a number of grammatical properties across different languages. First, they are biclausal. Second, they typically contain a comparative morpheme and/or a quantificational element, which may appear inside a PP. Third, they often exhibit a non-canonical word order. Fourth, they express a correlation between two properties or situations that involve incrementality of degree. These grammatical properties can be illustrated by the English example in (6) below (recall that we are following Culicover and Jackendoff’s (1999) convention in referring to the first clause of a CS as C1 and the second clause as C2).

(6) [_{C1} The more you exercise], [_{C2} the better you’ll feel].

The English sentence in (6) is biclausal, has an overt comparative Degree Phrase (DegP) displaced to the left periphery of each clause, and expresses a proportional correlation in degree between ‘exercising’ and ‘feeling good.’ While modern varieties of English do not use an overt quantificational element in CSs, archaic English did, as shown by the following example, where ‘how much’ and ‘so much’ appear inside a PP.

(7) “*By how much the better man you are yourself, by so much the more will you be inclined to believe me.*” (Fielding, *Tom Jones*; 18th c., cited in den Dikken 2005)

Cross-linguistically, we generally see the same grammatical properties in CSs. I illustrate this below with examples from a variety of languages cited in the literature.⁵

5. The following abbreviations are used in the glosses of the data in the paper: 1, 2, 3 for first, second, and third person, respectively; SG = singular; PL = plural; M = masculine; F = feminine; NEG = negation marker; FUT = future; COMP = complementizer; IPFV = imperfective; and IMP = imperative; CMPR = comparative; NOM = nominative; ACC = accusative; PRES = present. EA emphatic consonants are represented with a dot underneath the symbol (e.g., ʔil-) in the transcription of the data. Also, for convenience, I represent the EA definite article as the citation form ʔil-, even in contexts where it is not pronounced as such (e.g., due to assimilation). Finally, as noted in the text, I follow Culicover and Jackendoff’s (1999) convention of referring to the first clause of a CS as C1 and the second clause as C2.

Spanish (Borgonovo and Valmala 2010)

- (8) Cuanto más comes, (tanto) más contento estás
 how.much more eat.you (that.much) more happy are.you
 ‘The more you eat, the happier you are.’

Polish (Borsley 2003)

- (9) Im bardziej zmęczony jesteś, tym gorzej pracujesz
 IM more tired you-are TYM worse you-work
 ‘The more tired you are, the worse you work.’

Welsh (Borsley 2003)

- (10) Mwy a dw i 'n ddarllen, mwy a dw i 'n ddeall
 most am I PROG read most am I PROG understand
 ‘The more I read, the more I understand.’

Swedish (Dooley 2014)

- (11) Ju mer privat-a vi bli-r, desto personlig-are bli-r
 JU more private-PL we become-PRES DESTO personal-CMPR become-PRES
 vi också
 we also
 ‘The more private we become, the more personal we become also.’

Greek (Kapetangianni and Taylor 2009)

- (12) Oso pio poli zahari tros toso pio poles
 as.much.as more sugar eat.2SG.PRES that.much more
 thermides pernis
 calories get.2SG.PRES
 ‘The more sugar you eat, the more calories you get.’

German (Beck 1997)⁶

- (13) Je müder Otto ist, desto aggressiver ist er
 JE tired-er Otto is DESTO aggressive-er is he
 ‘The more tired Otto is, the more aggressive he is.’

Mandarin Chinese (Lin 2007)

- (14) nǐ yuè shēngqì tā jiù yuè gāoxìng
 You more angry he JIU more happy
 ‘The angrier you are, the happier he is.’

French (Beck 1997)

- (15) Plus quelqu’un est grand, plus il a de grand pieds
 more somebody is tall more he has of big feet
 ‘The taller somebody is, the bigger his feet are.’

6. Beck (1997) uses ‘the’ in glossing the markers *je* and *desto*. For consistency with the glossing across similar data in this section, I changed the glosses as shown in the example in (13).

Russian (den Dikken 2005)

- (16) Naskoľ'ko luchshe mashina, nastoľ'ko ona dorozhe
 by-how-much better car-NOM by-that-much it-F.NOM more.expensive
 'The better the car, the more expensive it is.'

Hungarian (den Dikken 2005; ADESS = adressive; VM = verbal modifier)

- (17) Minél többet olvasol, annál többet {megértesz /
 what-ADESS more-ACC you.read that-ADESS more-ACC VM-you.understand
 értesz meg}
 you.understand VM
 'The more you read, the more you understand.'

The CSs illustrated above are all biclausal, they contain a comparative/degree morpheme, and they may be overtly marked by a special marker in the first clause and sometimes in the second clause as well. Some of these markers tend to be quantificational (generally meaning 'much') and may occur inside a PP.

Against this background of cross-linguistic presentation of CSs, let us now consider CSs in EA. In particular, in the rest of this section, I describe and illustrate two types of CSs in EA, which I will refer to as Type A and Type B. I discuss the grammatical properties of both types, highlighting their structural similarities as well as how they differ from each other. I also point out some ways in which they differ from CSs in the other languages discussed in the literature.

2.1 Type A of correlation structures in EA

The most commonly used CS in EA is marked at the left periphery by the two elements *kull* and *mā*, where *kull* is the universal quantifier in the language, and *mā* is the complementizer typically used in adverbial subordinate clauses. I will refer to this type of CSs as *Type A*. An example is given in (18).⁷

- (18) [_{C1} kull mā ti-ʔrā ʔaktar] [_{C2} (kull mā)
 every COMP IPFV-read.2SGM more every COMP
 ha-ti-fham ʔaktar]
 FUT-IPFV-understand.2SGM more
 'Every increase in the degree/amount of your reading correlates with a proportional increase in the degree/amount of your understanding.'
 'The more you read, the more you will understand.'

7. In this section, I provide English glosses that are closer to the literal meaning of the EA examples, as well as the closest equivalent English structure, on separate lines. For ease of exposition, I do not always provide these semi-literal glosses for every example in the paper, but it is important to note that the English translations are approximations of the meanings associated with the EA structures.

Type A of CSs, as exemplified in (18), is characterized by a number of grammatical properties, which I discuss below.

First, the comparative DegP *ʔaktar* surfaces in its *in situ* position within the lexical domain in both clauses, unlike its fronted position in languages such as English, for example. Displacement of DegP to the left periphery is strictly disallowed, as the ungrammaticality of (19) shows.⁸

- (19) *ʔaktar (kull) (mā) ti-ʔrā ʔaktar (kull) (mā)
 more every COMP IPFV-read.2SGM more every COMP
 ha-ti-fham
 FUT-IPFV-understand.2SGM

Second, as shown in example (18), while the two-element marker *kull mā* is obligatory in C1, it is optional in C2. I will refer to this as the *lexical doubling* property of CSs.⁹ As shown by the cross-linguistic data cited earlier in this section, the two clauses of a CS tend to exhibit parallel structure, and lexical doubling in EA CSs seems to reflect this property.¹⁰

Third, examples such as (18) receive the same interpretation as their equivalents in languages such as English: They express a proportional correlation in degree between the two scales associated with the properties or situations expressed in the two clauses, as the gloss of the sentence indicates. I will refer to this type of meaning as *scalar correlation*, and to these instances of Type A as *scalar CSs*.¹¹

8. This *in situ* property has been noted for CSs in other languages as well, e.g., in Mandarin Chinese, as reported in McCawley 1988, Lin 2007, and E 2014; similarly, Leung 2003, cited in den Dikken 2005, notes the same for Cantonese, Thai, and Japanese.

9. My native speaker consultants and I do not find a clear semantic difference between CSs with a lexically doubled correlation marker and those without it, but there seems to be a stylistic preference among speakers to use lexical doubling. I should also point out that lexical doubling is not allowed in certain contexts in EA CSs, an important fact that I return to in Section 4.

10. While the modern Arabic dialects that I checked all seem to allow lexical doubling, Standard Arabic (the variety known as *Fushā*) does not seem to allow that (*pace* Alqurashi and Borsley 2014), as shown by attested data from Standard Arabic texts as well as explicit statements by traditional grammarians that lexical doubling is not allowed. Interestingly, this lexical doubling property is often transferred from a speaker's vernacular grammar to their formal writing in *Fushā*, leading to numerous instances where *kull mā* appears doubled within a CS in modern written Arabic. Given Arabic grammarians' position on this, it is safe to conclude that lexical doubling developed as a grammatical innovation, and that it was not allowed in early varieties of Standard Arabic, at least those that were documented by the early grammarians of Arabic.

11. Other possible terms that can be used to describe this type of correlation are 'proportional' or 'incremental' correlation. Needless to say, nothing hinges on the descriptive label used.

Fourth, scalar CSs in EA readily allow the occurrence of a standard of comparison (SoC), as in (20), a property that has been reported to be ungrammatical in languages such as English and German, for example (cf. Beck 1997).¹²

- (20) kull mā ti-zākir ʔaktar min Ahmad (kull mā)
 every COMP IPFV-study.2SGM more than Ahmad every COMP
 ha-t-gīb magmūf ʔaḥsan minn-u
 FUT-IPFV-get.2SGM score better than-him
 ‘Every increase in the degree/amount of your studying more than Ahmad correlates with a proportional increase in the degree/amount of your getting a better score than him.’
 ‘*The more you study than Ahmad, the better grade you will get than him.’

Fifth, CSs of Type A do not have to include a comparative DegP. Rather, a CS introduced by *kull mā* may contain non-comparative DegPs (e.g., the Deg head can be in the positive or the superlative form), or no overt DegP altogether, as the examples in (21)–(23) show.¹³

- (21) kull mā ti-ʔrā kitīr (kull mā) ha-ti-tʔallim kitīr
 every COMP IPFV-read.2SGM much every COMP FUT-IPFV-learn.2SGM much
 ‘Every instance of your reading much correlates with an instance of your learning much.’
 Roughly: ‘Every time you read much, you will learn much.’
- (22) kull mā Ahmad kān ʔašṭar ʔālib fī ʔil-ʔaṣl (kull mā)
 every COMP Ahmad was.3SGM best student in the-class every COMP
 ʔil-ustāz kān mabsūt minn-u
 the-teacher was.3SGM pleased.SGM from-him
 ‘Every instance of Ahmad’s being the best student in class correlates with an instance of his teacher’s satisfaction with him.’
 Roughly: ‘Every time Ahmad is the best student in the class, the teacher is pleased with him.’
- (23) kull mā ti-ʔrā (kull mā) ha-ti-fham
 every COMP IPFV-read.2SGM every COMP FUT-IPFV-understand.2SGM
 ‘Every instance of your reading correlates with an instance of your understanding.’
 Roughly: ‘Every time you read, you will understand.’

12. Bhatt (2009) cites a 2007 handout by E. A. Smith with examples showing that the ban on a SoC in comparative comparatives is not absolute in English (e.g., *The faster the cat is than the dog, the more likely it is to get away*). I will get back to this issue in Section 4.

13. That is precisely why the more generic term ‘correlation structure’ is more descriptively adequate than ‘comparative correlative’ for such structures in EA, as pointed out earlier in this paper.

While structures such as those in (21)–(23) still express a correlation, they do not express scalar incrementality in degree, as the glosses of the given examples indicate.¹⁴ I will refer to these instances of Type A as *non-scalar CSs*.

In sum, CSs of Type A in EA are marked by the two-element marker *kull mā*, which is obligatory in C1, but can be optionally lexically doubled in C2. A comparative DegP, when available, appears in situ and gives rise to a scalar correlation interpretation. This scalar CS can contain a SoC phrase. A DegP in a CS can also be non-comparative or missing altogether, in which case the CS has a non-scalar correlation interpretation.

2.2 Type B of correlation structures in EA

Type B of CSs in EA is marked at the left periphery by the marker *ʕalā ʔadd mā*, where *ʕalā ʔadd* is a PP literally meaning ‘on/by amount,’ and *mā* is the complementizer commonly used in adverbial subordinate clauses.¹⁵ As in Type A, the marker is obligatory in C1 and can be optionally lexically doubled in C2. Consider the illustrating example in (24).

14. The optionality of an overt DegP has been reported in *oso*-CSs in Greek by Kapetangianni and Taylor (2009). They do not say, however, if a semantic difference in proportionality obtains in Greek. Also, Bhatt (2009) cites data from CSs in Bulgarian (due to Roumyana Pancheva) and Hindi-Urdu, where no comparative DegP occurs.

(i) Bulgarian:

Kolkoto četeš tokova naučavaš
how.much read.2SG that.much learn.2SG
‘As much you read, that much you learn.’

(ii) Hindi-Urdu (PFV = perfective)

Jitnii der ho-tii gayii, utnii ʔhaṇḍ baṛh-tii gayii
how.much.F late be-IPFV.F go.PFV.F that.much.F cold increase-IPFV.F go.PFV.F
‘The later it got, the colder it became.’ (Literally: As much late it got, that much cold it became.)

It is hard to draw any generalizations based on the short discussion by Bhatt, but it seems from the gloss of the Bulgarian example in (i) that absence of a comparative DegP is associated with a correlation in fixed degrees, which would be similar to Type B of CSs in EA. The gloss of the Hindi-Urdu example in (ii), by contrast, seems to indicate a scalar correlation in degree, even in the absence of a comparative morpheme. Given the limited cross-linguistic data available, it is hard for one to draw a generalization regarding the type of DegP used in a given CS and its associated interpretation.

15. In this Arabic dialect, the noun *ʔadd* (= ‘amount’) also combines with the wh-word *ēh* (= ‘what’) to form the compound wh-phrase *ʔadd ēh*, which is used in both interrogative and exclamative structures in the context of questions/exclamations about quantity, age, or prices, sometimes in combination with a preceding preposition.

- (24) [_{C1} ṣalā ʔadd mā ti-ʔrā] [_{C2} (ṣalā ʔadd mā)
 on amount COMP IPFV-read.2SGM on amount COMP
 ha-ti-fham]
 FUT-IPFV-understand.2SGM
 ‘The degree/amount of your reading correlates with an equal degree/amount
 of your understanding.’
 Approximate translation: ‘By how much you read, by so much you will
 understand.’

As the gloss of (24) shows, Type B expresses a *non-proportional* (or *non-variable*) correlation between fixed points on two given scales: *The degree/amount of X correlates with an equal degree/amount of Y*.¹⁶ Given its non-proportional correlation interpretation, Type B does not permit the occurrence of a comparative DegP, as shown by the ungrammaticality of (25).

16. A similar non-proportional interpretation has been reported for CSs in Latin and Mongolian. Michaelis (1994) points out that the *quanto...tanto...* CS in Latin is polysemous in that it can be used to express proportional and non-proportional readings, as illustrated by (i) below.

- (i) a. Quanto inpectore hanc rem meo magis
 how-much.ABL in heart.ABL this.ACC matter.ACC my.ABL more
 voluto, tanto mi aegritudo auctior
 ponder.1SG.PRES.IND.ACT that-much.ABL me.DAT grief.NOM greater.NOM
 est in ahimo.
 is.3SG.PRES.IND in spirit.ABL
 ‘The more I turn this matter over in my mind, the greater the grief is in my soul.’
 Plautus, *Captivi* 781–2.
- b. Sed non statuendo felicitati modum, quanto altius elatus
 but not setting.ABL success.DAT limit.ACC how-much.ABL higher lifted.NOM
 erat tanto foedius conruit.
 was.3SG.IMPERF.IND that-much.ABL worse fell.3SG.PERF.IND
 ‘But by not setting a limit to his success, to the extent that he [M. Atilius] had risen
 high, he fell badly.’ Livy 30.30

Similarly, Hsiou (2003) points out that the *xedii...tödii...* CS in Mongolian can have either a proportional or a non-proportional reading, as the data in (ii) shows.

- (ii) a. xedii ix bol tödii sain
 how-much many TOPIC that-much good
 ‘The more the better.’
- b. Bat xedii calin av-dag bol tödii tögrög zarcuul-dag
 Bat how-much salary take-PRES TOPIC that-much money spend-PRES
 ‘Bat spends all the money he earns.’

The literature on CSs, however, has typically focused on those that express proportionality, commonly referred to as comparative correlatives, as noted earlier in this paper.

- (25) **ʕalā ʔadd mā ti-ʔrā ʔaktar* (*ʕalā ʔadd mā*)
 on amount COMP IPFV-read.2SGM more on amount COMP
ha-ti-fham ʔaktar
 FUT-IPFV-understand.2SGM more

In sum, CSs of Type B in EA are characterized by the marker *ʕalā ʔadd mā* at the left periphery, which is obligatory in C1, but optional in C2. Type B has a non-proportional/non-variable correlation interpretation, and is incompatible with an overt comparative DegP.

Given the descriptive account in this section, CSs in EA pose two main questions. First, there is the macro-syntax question: What is the syntactic relation linking the two clauses of a CS: Subordination or parataxis? Second, there is the micro-syntax question: What does the internal syntax of CSs look like, such that we are able to account for their grammatical properties, including the multiple types of correlation interpretations they give rise to in the language? I discuss both questions in Sections 3 and 4, respectively.¹⁷

3. The macro-syntax of correlation structures

As has been mentioned earlier, the syntax of CSs has typically been claimed to exhibit some idiosyncratic grammatical properties. For one thing, CSs tend to be marked by syntactic elements that are either specific to these structures or are used in a very different manner from their typical usage elsewhere in a language. As shown by the cross-linguistic data in (8)–(17), many languages use specific particles in such structures; e.g., Swedish *ju ... desto*, Mandarin Chinese *yuè ... yuè*, and Greek *oso ... toso*. Alternatively, a language can use an already existing morpheme in its lexicon but in a rather unusual manner. The classic example is the use of ‘the’ in English CSs, where the word does not serve in its usual function

17. A third question not discussed here is whether CSs are a subtype of conditionals, which has been debated in both traditional grammar and modern linguistic analysis. We have already mentioned the use of the label *comparative conditional* by several authors to describe CSs, as in Fillmore (1987), McCawley (1988), and Beck (1997), among others. In fact, Beck’s (1997) semantic analysis attempts to assimilate CSs to conditionals; cf. footnote 29. Huber (2007) provides arguments against treating CSs as conditionals, however. The conditionality (or lack thereof) of *kulla-mā* CSs in Standard Arabic has also been a topic for debate in the Arabic grammatical tradition since the 8th century. While Al-Khalil (718–786) and Sibawayh (ca. 765–796), among others, describe *kulla-mā* as non-conditional, Al-Zamakhšari (1075–1144) seems to treat it on par with conditional markers. Meanwhile, Abu Ḥayyān (1256–1344) and Ibn Hishām (1306–1360) describe it as conditional in its semantics.

as a determiner.¹⁸ Similarly, as we have seen in the previous section, in Type A of CSs in EA, the universal quantifier *kull* is unusually followed by a full clause, rather than by a nominal.

At the same time, the syntactic relation between C1 and C2 seems to also exhibit some idiosyncratic behavior. While there is good evidence that the relation is one of subordination, the whole structure has a ‘paratactic feel’ to it, especially given how parallel the two clauses are. Cross-linguistic evidence for subordination comes from the behavior of tag questions, the subjunctive mood, and word order in V2 languages. I illustrate each below.

Tag questions in languages like English are typically formulated with reference to the subject and verb of a main clause, but not with reference to the subject and verb of an adverbial subordinate clause.

- (26) a. When we eat, he gets angry, doesn’t he?
 b. *When we eat, he gets angry, don’t we?

As Culicover and Jackendoff (1999) point out, tag questions in English CSs can only be formulated with reference to C2, and not C1.

- (27) a. The more we eat, the angrier you get, don’t you?
 b. *The more we eat, the angrier you get, don’t we?

This contrast follows if C2 is the main clause of the structure, whereas C1 is an adverbial subordinate clause.

The use of subjunctive morphology also provides evidence that the relation between C1 and C2 is that of subordination. Consider the following English examples, again from Culicover and Jackendoff (1999).

- (28) It is imperative that/I demand that...
 a. the more John eats, the more he pay(s).
 b. *the more John eat, the more he pay(s).

As the contrast in (28) shows, when a CS occurs as a complement of a predicate that triggers subjunctive morphology, only the verb in C2, but not the one in C1, may appear in the subjunctive, again indicating that C2 is the main clause, whereas C1 is a subordinate clause in such structures.¹⁹ The same fact holds in languages such

18. This idiosyncratic use of ‘the’ is reflected in how it has been differently characterized in various syntactic analyses of English CSs: Culicover and Jackendoff (1999) treat it as a specifier of words such as *more/less*; den Dikken (2005) treats it as a Deg head; and Taylor (2013) treats it as a complementizer.

19. Notice that the subjunctive morphology is optional in English.

as Spanish, as shown by the contrast in (29) from Borgonovo and Valmala (2010) (IND = indicative, and SUBJ = subjunctive).

- (29) a. Quiero que cuanto mas lean, mas
 want.I that how.much more read.IND.they more
 entiendan
 understand.SUBJ.they
 ‘I want them to understand more the more they read.’
- b. *Quiero que cuanto mas lean, mas
 want.I that how.much more read.SUBJ.they more
 entienden
 understand.IND.they

A third argument in support of the main clause status of C2 and the subordinate clause status of C1 is based on facts of word order in V2 languages, as shown by the German and Swedish examples in (30a,b), respectively.²⁰

- (30) a. Je länger du in Deutschland wohnst, desto besser wirst du
 JE longer you in Germany live DESTO better will you
 Deutsch sprechen
 German speak
 ‘The longer you live in Germany, the better you will speak German.’
- b. Ju närm-are vi kom, desto orolig-are blev vi
 JU near-CMPR WE come.PAST DESTO worried-CMPR become.PAST WE
 ‘The closer we came, the more worried we became.’

As McCawley (1988) initially noted, in German CSs such as (30a), the verb in C2 occupies the second position, per the word order rules for main clauses; in C1, however, the verb appears in final position, which is typical of the word order in German subordinate clauses. The same holds in Swedish CSs, as shown by (30b).

At the same time, CSs seem to exhibit grammatical properties that we do not expect with subordination.²¹ First, since adverbial subordinate clauses are optional, we expect C2 in a CS to occur on its own without the need for C1, contrary to fact, as the ungrammatical English example in (31) shows.

- (31) *The more you understand.

20. The German example in (30a) is from McCawley (1988), whereas the Swedish example in (30b) is from Dooley (2014).

21. Culicover and Jackendoff (1999) use this paradoxical evidence between subordination and parataxis to argue that English comparative correlatives pose a serious challenge to standard syntactic analysis and that they should be posited as ‘constructions’ in the grammar. But see den Dikken (2005) and Taylor (2013) for detailed responses to that claim.

Second, while the linear order of a main clause and an adverbial clause is typically free without affecting the general meaning of a sentence, as shown in (32), that is not the case in CSs, where a change in the ordering of C1 and C2 leads to a complete change in meaning, as shown by (33).

- (32) a. After we saw the movie, we went home.
 b. We went home, after we saw the movie. (= (32a))
- (33) a. The more you read, the more you understand.
 b. The more you understand, the more you read. (≠ (33a))

A third argument marshaled by Culicover and Jackendoff (1999) has to do with extraction possibilities out of CSs. They note that relativization and topicalization are possible out of both clauses, citing the data in (34)–(36).

- (34) The sooner you solve this problem, the more easily you'll satisfy the folks up at corporate headquarters. [base sentence]
- (35) a. This is the sort of problem which_i the sooner you solve t_i , the more easily you'll satisfy the folks up at corporate headquarters. [extraction from C1]
 b. The folks up at corporate headquarters are the sort of people who_i the sooner you solve this problem, the more easily you'll satisfy t_i . [extraction from C2]
- (36) a. This problem_p, the sooner you solve t_p , the more easily you'll satisfy the folks up at corporate headquarters. [extraction from C1]
 b. The folks up at corporate headquarters_p, the sooner you solve this problem, the more easily you'll satisfy t_p . [extraction from C2]

While extraction out of C2 is unsurprising, it is extraction out of C1 that poses a challenge to the treatment of CSs as involving subordination: If C1 is indeed an adverbial subordinate clause, we expect extraction to be impossible due to the adjunct island constraint, contrary to the data in (35a) and (36a).²²

In the rest of this section, I focus on discussing the macro-syntax of CSs in EA, which, as will be discussed shortly, exhibit clear properties of subordination. More specifically, C1 exhibits the characteristics of an adverbial subordinate clause, whereas C2 behaves as the main clause of the structure. This is supported by three

22. An anonymous reviewer points out that some violations of the adjunct island constraint have been shown to be grammatical in English, citing Chaves (2012). The literature on the study of islands does indeed make reference to examples showing variation and subtlety in the judgement of different types of island violations. This is clearly a topic that is beyond the scope of this article, and nothing in this paper hinges on that particular issue. For the purposes of the discussion in this section, I simply follow Culicover and Jackendoff (1999) in assuming that extraction out of adjuncts is generally not allowed (at least in a considerable number of contexts) in English, particularly with regard to the comparative correlative structures that they cite.

pieces of empirical evidence: the complementizer *mā*, the behavior of tag questions, and the occurrence of imperatives. I discuss each in turn.

First, as noted in Section 2, both types of CSs are marked by the occurrence of the complementizer *mā* (cf. the examples in (18) and (24), for instance), which is a typical grammatical property of adverbial clauses in the language, as the temporal and locative adverbial clauses in (37) below show.

- (37) a. baʿd mā ʔitfašē-nā xarag-nā ʔitmašē-nā
 after COMP ate.dinner-1PL went.out-1PL walked-1PL
 ‘After we had dinner, we went out for a walk.’
 b. fēn mā /maṭraḥ mā trūḥ ha-nrūḥ maṣā-k
 where COMP /location COMP go.2SGM FUT-go.1PL with-you
 ‘Wherever you go, we will go with you.’

The occurrence of the complementizer *mā* in CSs thus clearly indicates that we are dealing with a case of subordination.²³

A second argument in support of C1 being an adverbial subordinate clause, and C2 being a main clause, comes from the behavior of tag questions. In EA, a tag question is expressed via the use of expressions such as *ṣaḥ* (= ‘correct’) or *miš kida* (= ‘not so’), as shown in (38).

- (38) a. ʔinti duktōra ṣaḥ?
 you.SGF doctor correct
 ‘You are a doctor, correct?’
 b. ʔinti duktōra miš kida?
 you.SGF doctor NEG so
 ‘You are a doctor, isn’t that so?’

What is relevant to our purposes here is that when a tag question is appended to a CS, it can only be understood as a question on the content of C2, and not C1, as shown by the felicity of the responses in (39b) and (40b), as opposed to the infelicity of (39c) and (40c).

- (39) a. kull mā Ahmad yi-tʔaxxar kitīr barra ʔil-bēt (kull
 every COMP Ahmad IPFV-be.late.3SGM much outside the-house every
 mā) mamt-u bi-ti-ʔla? ʔalē-h miš kida?
 COMP mother-his ASP-IPFV-worry.3SGF on-him NEG so
 ‘The more Ahmad is out late, the more worried his mother is about him,
 isn’t she/*isn’t he?’

23. This, however, does not explain why the complementizer *mā* also appears in C2, which is the main clause. I will get back to the lexical doubling property in Section 4.2 and provide an answer to that question.

- b. ʔaywa bi-ti-ʔlaʔ ʕalē-h ʔawī
 yes ASP-IPFV-worry.3SGF on-him very.much
 ‘Yes. She worries about him a lot.’
- c. #ʔaywa bi-yi-tʔaxxar kitīr ʔawī
 yes ASP-IPFV-be.late.3SGM much very
 #‘Yes. He is out late a lot.’
- (40) a. ʕalā ʔadd mā ʔil-wilād ha-yi-ʔr-ū (ʕalā ʔadd
 on amount COMP the-children FUT-IPFV-read-3PL on amount
 mā) ha-yi-fham-ū miš kida?
 COMP FUT-IPFV-understand-3PL NEG this?
 ‘The amount of reading the children do is equal to the amount of under-
 standing they will gain; isn’t that so?’
- b. ʔaywa ha-yi-fham-ū kuwayyis ʔin-šāʔa-llāh
 yes FUT-IPFV-understand-3PL well if-wanted.3SG-Allah
 ‘Yes. Hopefully, they will understand well.’
- c. #ʔaywa ha-yi-ʔr-ū kuwayyis ʔin-šāʔa-llāh
 yes FUT-IPFV-read-3PL well if-wanted.3SG-Allah
 #‘Yes. Hopefully, they will read well.’

Since tag questions cannot target the content of an adverbial subordinate clause, it follows from the data in (39) and (40) that C1 is indeed the subordinate clause of a CS in EA, whereas C2 is the main clause.

The third argument in support of the subordinate-clause status of C1 and the main-clause status of C2 comes from the occurrence of imperatives. As (41) below shows, imperatives are only possible in main clauses, but not in adverbial subordinate clauses.

- (41) a. baʕd mā t-ākl-ū ʔiʕmil-ū ʔil-wāgib
 after COMP IPFV-eat-1PL do.IMP-1PL the-homework
 ‘After you eat, do the homework.’
- b. *baʕd mā kul-ū ʔiʕmil-ū ʔil-wāgib
 after COMP eat.IMP-1PL do.IMP-1PL the-homework

As the data in (42)–(43) below shows, in both types of CSs in EA, an imperative verb can only occur in C2, but never in C1, which is what we would expect if C2 is indeed a main clause, whereas C1 is a subordinate clause.

- (42) a. kull mā ta-kul ʔaktar ʔilʕab riyāda ʔaktar
 every COMP IPFV-eat-2SGM more play.IMP.2SGM sports more
 ‘The more you eat, exercise more!’
- b. *kull mā kul ʔaktar ʔilʕab riyāda ʔaktar
 every COMP eat.IMP.2SGM more play.IMP.2SGM sports more

- (43) a. Ṣalā ṭadd mā b-t-ākul ṭilṣab riyāḍa
 on amount COMP ASP-IPFV-eat.2SGM play.IMP.2SGM sports
 ‘Exercise to the degree/amount equal to the degree/amount of your eating.’
 b. *Ṣalā ṭadd mā kul ṭilṣab riyāḍa
 on amount COMP eat.IMP.2SGM play.IMP.2SGM sports

The fact that imperatives can occur in C2, but not in C1, thus further supports the conclusion that in EA CSs, C2 is the main clause, whereas C1 is the subordinate clause.

To sum up this section, the presence of the complementizer *mā*, the behavior of tag questions, and the occurrence of imperatives, all support the claim made here that C1 of an EA CS is a subordinate clause, whereas C2 is a main clause. It is thus safe to conclude that the macro-structure of CSs in EA indeed involves subordination. The next question to address has to do with their micro-syntax: What is exactly the internal structure of CSs? I discuss this next.

4. The micro-syntax of correlation structures

In their investigations of cross-linguistic variation in CSs, both den Dikken (2005) and Taylor (2013) argue for a syntactic analysis of such structures whereby C1 is merged as an adjunct of C2, thus accounting for the subordination properties characterizing such structures cross-linguistically, as discussed in the previous section. While the two analyses differ in the details, they share the same main assumption regarding the adjunct status of C1, which I adopt here given the empirical evidence in support of the subordinate-clause status of C1 in EA CSs, as argued in Section 3. For the purposes of the discussion in this section, I present den Dikken’s proposed structure for CSs, and show how it can be modified to account for the facts of EA CSs discussed in this paper.²⁴

The core assumption of den Dikken’s (2005) analysis is that CSs are in essence *correlative* structures of the type attested in languages such as Hindi-Urdu, as first discussed in Srivastav (1991). In particular, in such languages, in addition to regular relative clauses, such as the one in (44a), we also find correlative structures such as (44b).²⁵

24. Taylor (2013) proposes a cartographic recursive CP structure to account for the facts of CSs in English. Later in the paper (cf. Section 5), I will indeed argue that a cartographic approach is more empirically adequate to account for certain facts about CSs in EA. For the purposes of the current section, however, an adjunction analysis of the type proposed by den Dikken (2005) suffices.

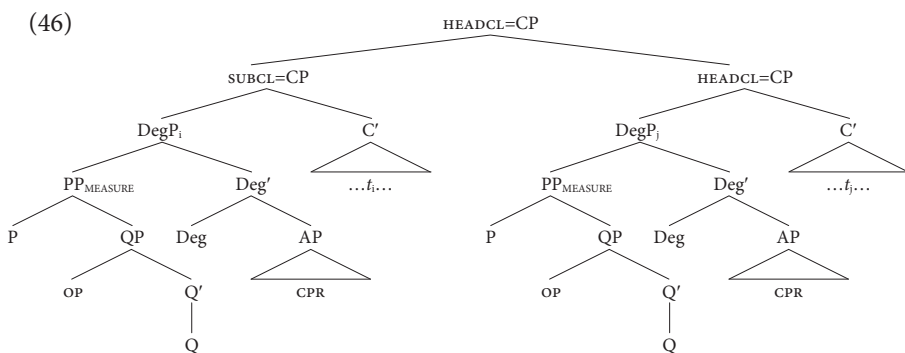
25. A more appropriate spelling would be ‘co-relative’, given that the subordinate clause in this case is a relative clause, but the use of ‘correlative’ is now standard.

- (44) a. [_{IP} [_{DP} *Vo* *laRkii* [_{CP} *jo* *khaRii* *hai*]] [_{I'} *lambii* *hai*]]
 DEM girl REL standing is tall is
 ‘The girl that is standing is tall.’
- b. [_{IP} [_{CP} *Jo* *laRkii* *khaRii* *hai*] [_{IP} *Vo* *lambii* *hai*]]
 REL girl standing is DEM tall is
 ‘Which girl is standing, that (one) is tall.’

A correlative structure is characterized by the adjunction of a relative clause, which is marked by a relativizing operator (REL in (44b)), to a main clause, which, in turn, is introduced by a demonstrative element (DEM in (44b)). A skeletal structure of correlatives is thus one along the lines of (45), where *Op* stands for ‘operator’ and *Dem* for ‘demonstrative.’

- (45) [_{main_clause} [_{subordinate_clause} *Op* ...] [_{main_clause} *Dem* ...]]

The key idea of den Dikken’s (2005) proposal is to extend the macro-structure in (45) to CSs cross-linguistically, where an overt comparative DegP typically appears at the left periphery of both clauses in the languages discussed in his article. A tree diagram of his proposed structure is given in (46), where HEADCL = head clause, whereas SUBCL = subordinate clause (C2 and C1, respectively, in the Culicover and Jackendoff’s labeling convention adopted in this paper); CPR stands for the ‘comparative’ morpheme, and OP for ‘operator.’²⁶



The structure in (46) is best illustrated by the CSs in (47) from archaic varieties of English, where every position of the proposed structure in (46) is filled with lexical material. A schematic structural representation of (47a), for instance, is as in (48).

26. In this section, I choose to adopt den Dikken’s proposed macro-structure for CSs, but I do not assume that CSs and correlatives have the same identical structure beyond their macro-syntax. In fact, EA has correlative structures of the Hindi-Urdu-type, but they differ from the CSs discussed here in their micro-syntax.

- (47) a. “By how much the lesse he looked for this discourse, by so much the more he liked it.” (Lyly, *Euphues*, 16th c.; from Jespersen 1961: 5.383, as cited in den Dikken 2005)
- b. “By how much the better man you are yourself, by so much the more will you be inclined to believe me.” (Fielding, *Tom Jones*; 18th c., as cited in den Dikken 2005)

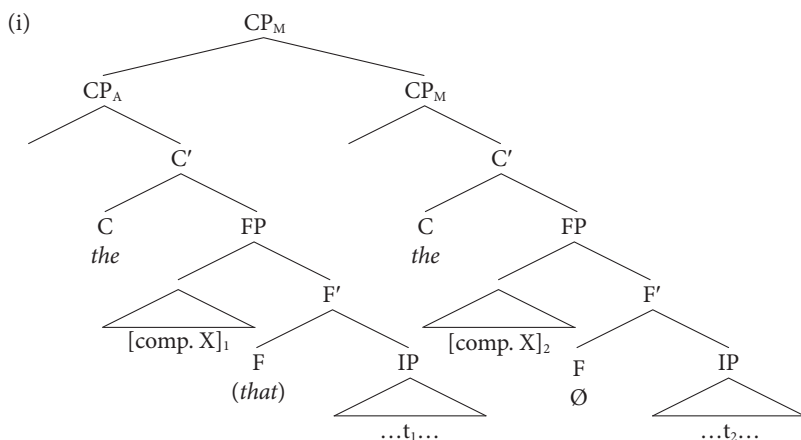
- (48) $[_{C1} [_{DegP} [_{PP} \text{by } [_{QP} \text{how much}]]] [_{Deg'} \text{the } [_{AP} \text{lesse}]]]_i [\dots t_i \dots]$,
 $[_{C2} [_{DegP} [_{PP} \text{by } [_{QP} \text{so much}]]] [_{Deg'} \text{the } [_{AP} \text{more}]]]_j [\dots t_j \dots]$

Archaic English CSs stand in contrast to those in today’s English varieties, where some of the positions in structure (46) appear null instead, as illustrated in (50) for the English example in (49).²⁷

- (49) The more you read, the more you understand.

- (50) $[_{C1} [_{DegP} [_{QP} \emptyset\text{-OP } \emptyset] [_{Deg'} \text{the } [_{AP} \text{more}]]]_i [\dots t_i \dots]$,
 $[_{C2} [_{DegP} [_{QP} \emptyset\text{-DEM } \emptyset] [_{Deg'} \text{the } [_{AP} \text{more}]]]_j [\dots t_j \dots]$

27. Taylor (2006, 2013) proposes a recursive CP structure for CSs in English. Under her analysis, ‘the’ is treated as a complementizer, with the comparative DegP occupying the specifier position of a functional projection FP, which may optionally be headed by the complementizer ‘that’. So, an English example such as ‘*The more (that) you eat, the happier you’ll feel*’ can be generated with the structure in (i) below, where CP_M = main clause CP; CP_A = adjunct CP; FP = a functional projection in the left periphery; and [comp. X] = a comparative DegP.



One advantage of Taylor’s analysis is that it allows for the explanation of extraction facts in CSs in languages like English, since the structure in (i) includes a SpecCP position in both C1 and C2 to which an XP (e.g., a wh-phrase) can move. This is not easy to accommodate under den Dikken’s analysis where SpecCP is already filled by a DegP. I refer the interested reader to Taylor (2013) for a detailed discussion of CSs in English and a few other languages.

Den Dikken further provides cross-linguistic data showing that various languages generally adhere to the skeletal structure in (46) for expressing correlation, but vary with regard to which syntactic heads are lexically realized. The question now is: Can den Dikken's analysis be extended to account for the facts of CSs in EA? More specifically, how can the structure in (46) accommodate the following facts holding in EA CSs?

- (51) a. In scalar CSs, the comparative DegP appears in situ.
 b. In scalar CSs, a standard of comparison is allowed.
 c. In addition to scalar CSs, the language also exhibits non-scalar as well as non-proportional CSs.

In the next section, I show that the analysis discussed in this section can indeed be extended to account for the properties in (51).

4.1 Deriving correlation structures in EA

In this section, I present an analysis of CSs in EA that accounts for the grammatical properties listed in (51) above, in light of the core insights of den Dikken's (2005) proposal (cf. also Taylor 2013 for a similar analysis of CSs in English and a few other languages).

First, I assume that C1, an adverbial subordinate clause, is adjoined to C2, a main clause. This directly accounts for the empirical facts discussed in Section 3 regarding the complementizer *mā*, tag questions, and imperatives.

Second, unlike the languages discussed by den Dikken (2005) and Taylor (2013), EA comparative DegPs in Type A are not displaced to the left periphery. This, however, follows straightforwardly from the fact that the quantifier *kull* is merged in SpecCP, rendering DegP displacement unavailable. Interestingly, in Maltese, which does not mark CSs with *kull*, the comparative DegP *aktar* (= 'more') undergoes overt movement, as the following example from Beck (1997) shows (notice that Beck glosses the complementizer *ma* as PART(ICLE)).²⁸

- (52) aktar ma jkun kiesah avukat aktar ikollu success
 more PART is cold attorney more has success
 'The colder an attorney is, the more success he has.'

Third, an in situ comparative DegP in a CS is expected to exhibit the same syntactic behavior of DegPs in the language, hence the occurrence of a SoC in CSs is not surprising (cf. the example in (20)). In fact, it is what we expect. This obviously raises a question regarding comparative CSs in languages such as English and German, where a SoC has generally been assumed to be disallowed. Beck (1997) argues that

28. Thanks to Maris Camilleri for drawing my attention to this fact in Maltese Arabic.

the ban on a SoC in such structures is due to their semantics requiring the presence of an obligatorily implicit SoC.²⁹ This proposal clearly cannot be extended to CSs in languages such as EA, among several others, where a SoC readily occurs.³⁰ In fact, Bhatt (2009) argues that positing a silent SoC in such structures is unnecessary, citing data from Hindi-Urdu and a variety of other languages, including English, where a SoC is shown to be possible in CSs (see footnote 12). Instead, Bhatt argues in favor of a semantic analysis of CSs in terms of the notion of a ‘reference world,’ rather than in terms of Beck’s quantification over pairs of worlds. While this does not provide a clear answer why some languages such as English and German do not usually allow SoCs in their CSs, it is worth pointing out that the EA facts are what we would expect. A ban on a SoC inside a DegP would actually be surprising because it would have to be stipulated as a construction-specific property, given that, as Bhatt argues, it cannot be derived from the semantics of such structures.

Fourth, assuming that the complementizer in each clause of a CS carries a [correlative] feature, which is licensed via merge of a QP or a measure PP in SpecCP, we directly account for Type A and Type B, respectively.³¹

Finally, the scalar correlation interpretation associated with Type A can be assumed to result from the QP at the left periphery binding a degree variable in the DegP via the mechanism of *unselective binding* (UB) (cf. Lewis 1975; Heim 1982; Pesetsky 1987; see also E 2014 for an UB proposal regarding in situ CSs in Mandarin Chinese).³² Under this analysis, a scalar CS such as (18), repeated in (53) for convenience, is assigned the structural representation in (54).³³

29. Under Beck’s (1997) semantic analysis, a CS such as ‘*The more you read, the more you understand*,’ in English, is assigned an interpretation along the lines of (i) below.

(i) *For all pairs of worlds w_1 and w_2 , if you read more in w_2 than in w_1 , then you understand more in w_2 than in w_1 .*

More generally, Beck’s analysis attempts to assimilate CSs to conditionals, hence her use of the term *comparative conditional* to refer to such structures.

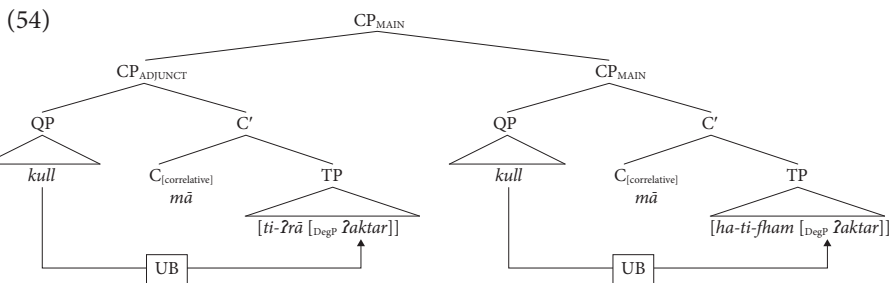
30. See Lin (2007) for a formal semantic analysis of CSs with focus on data from Mandarin Chinese.

31. The assumption regarding the presence of a [correlative] feature in both clauses does not derive the relationship between the two clauses in a straightforward manner, but it should suffice for our purposes in this section. In Section 5, I propose a cartographic implementation of the analysis, where [correlative] is a feature of the main clause, and the merge of QP and PP_{MEASURE} licenses a [quantificational] feature on the head *mā* in each clause.

32. It is worth pointing out here that the dependency between the QP *kull* and an in situ comparative DegP in EA is subject to a clausemate condition; as a result, long-distance dependencies in CSs are disallowed. It is not possible, therefore, to rely on island (in)sensitivity as a diagnostic to test for movement in such structures.

33. In the tree in (54), I use CP_{MAIN} to refer to the main clause and CP_{ADJUNCT} for the subordinate clause (C2 and C1, respectively, in the convention adopted in this paper).

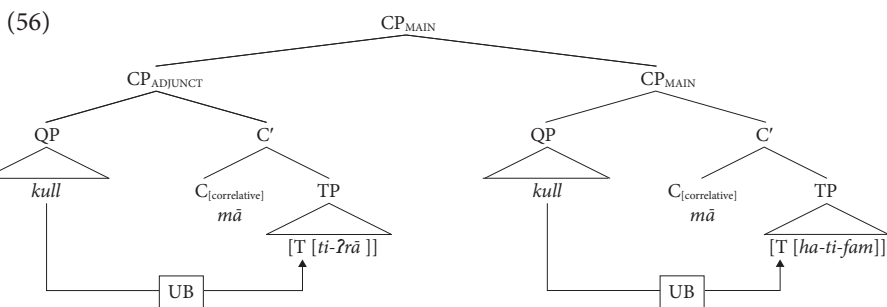
- (53) [_{C1} kull mā ti-ʔrā ʔaktar] [_{C2} (kull mā)
 every COMP IPFV-read.2SGM more every COMP
 ha-ti-fham ʔaktar]
 FUT-IPFV-understand.2SGM more
 ‘The more you read, the more you understand.’



The structure in (54) gives rise to the type of correlation associated with such structures: Every instance of reading at degree d correlates with an instance of understanding at degree d' , and d and d' are proportional.

When the DegP in Type A is non-comparative or missing altogether, the QP *kull* binds an event variable instead, presumably located in T, as represented in (56), for the sentence in (23), repeated below as (55).

- (55) kull mā ti-ʔrā (kull mā) ha-ti-fham
 every COMP IPFV-read.2SGM every COMP FUT-IPFV-understand.2SGM
 Roughly: ‘Every time you read, you will understand.’



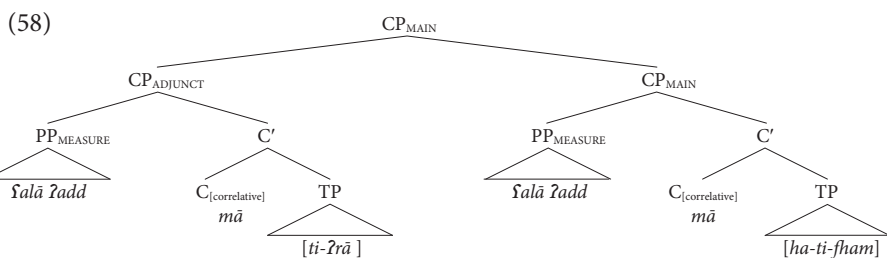
Such structures are associated with a non-scalar correlation reading, in what may be described as an ‘event correlation’: (56) states that every event of reading correlates with an event of understanding.

In contrast to Type A, the measure PP in SpecCP in Type B does not create a quantifier-variable dependency; rather the measure PP provides a fixed degree value in each clause. This is represented in (58), for the sentence in (24), repeated below as (57).

- (57) [_{C1} *ʕalā ʔadd mā ti-ʔrā*] [_{C2} (*ʕalā ʔadd mā*)
 on amount COMP IPFV-read.2SGM on amount COMP
 ha-ti-fham]
 FUT-IPFV-understand.2SGM

‘The degree/amount of your reading correlates with an equal degree/amount of your understanding.’

Literally (using archaic English): ‘By how much you read, by so much you will understand.’



The representation in (58) is associated with a meaning along the following lines: The degree of your reading is equal to the degree of your understanding. In such structures, degrees are expressed by the measure PP, not via variables, in contrast to scalar CSs. As such, they give rise to a non-variable/non-proportional correlation interpretation.

In sum, the proposed analysis accounts for the main facts of CSs in EA. Adjunction of C1 to C2 accounts for the subordination facts discussed in Section 3. Absence of overt displacement of a comparative DegP is due to the presence of an overt QP in SpecCP. In dialects where no such QP occupies SpecCP, overt displacement takes place, as in Maltese. Unselective binding of a degree variable results in a scalar type of correlation. Unselective binding of an event variable gives rise to a non-scalar correlation. The merging of a measure PP in SpecCP gives rise to a non-proportional correlation. Finally, in the scalar *kull mā* type, the comparative DegP can, unsurprisingly, host a standard of comparison. This leaves us with the lexical doubling property of CSs, which I discuss next.³⁴

34. An anonymous reviewer asks if an account that relies on the presence of a [correlative] feature of a functional head is actually very different from an analysis that accounts for the facts because some ‘construction’ requires certain components. This is a valid point that makes one wonder whether such different analyses may sometimes be merely notational variants of one another. Given that the primary goal in this paper is to show that the grammatical properties of CSs can indeed be accounted for within a theory that does not rely on constructions, I choose not to engage into a full discussion of the arguments that are typically made on both sides of this debate (but see footnote 47 for one such conceptual argument), referring the reader instead to the multiple works cited throughout this paper for further discussion.

4.2 Lexical doubling in CSs as a post-syntactic operation

Lexical doubling is puzzling because main clauses do not normally ‘double’ subordinators of an adjunct clause (e.g., subordinators such as ‘after,’ ‘before,’ or ‘if,’ are not lexically doubled in the main clause). A solution to the puzzle, however, is possible, given what we mentioned earlier about the parallelism between the two clauses of a CS. Cross-linguistically, parallelism seems to be a robust feature of CSs, which is probably responsible for their paratactic feel (cf. the cross-linguistic data in 8–17 in Section 2). Building on this observation, I would like to argue that lexical doubling in EA CSs is the result of an (optional) post-syntactic operation of phonological parallelism. More specifically, while the QP/PP and the complementizer are obligatory subordinators in C1, they are syntactically null in C2, but may undergo lexicalization at PF for phonological parallelism. There are three pieces of empirical evidence that support this claim. I discuss each below.

The first piece of evidence in support of a phonological parallelism account of lexical doubling comes from contexts where lexical doubling is obligatorily blocked. This is the case when C2 is in the imperative mood, as shown in (59).

- (59) a. kull mā ta-kul ʔaktar (*kull mā) ʔilʕab
 every COMP IPFV-eat-2SGM more every COMP play.IMP.2SGM
 riyāḏa ʔaktar
 sports more
 ‘The more you eat, exercise more!’
- b. ʕalā ʔadd mā b-t-ākul (*ʕalā ʔadd mā)
 on amount COMP ASP-IPFV-eat.2SGM on amount COMP
 ʔilʕab riyāḏa
 play.IMP.2SGM sports
 ‘Exercise to the degree/amount equal to the degree/amount of your eating!’

If lexical doubling is an optional post-syntactic rule, as proposed here, we expect it to be blocked if an obligatory operation targets the complementizer position of C2. This is exactly what happens in imperative contexts, where verbs in Arabic dialects are assumed to raise all the way up to the complementizer position (cf. Soltan 2009). If lexical doubling applies post-syntactically, as proposed here, then its absence in CSs whose C2 is imperative readily follows: In the presence of an imperative verb in C, the lexical doubling rule simply cannot apply in the post-syntax.³⁵

35. Whatever the exact formulation of the (optional) lexical doubling rule is, it must require the spell-out of both the complementizer *mā* and the XP in its Spec, which is reasonable if the rule is indeed driven by phonological parallelism. The trigger for phonological parallelism can simply be the presence of a [correlative] feature on C (or on Topic, as will be proposed in the next section).

A second piece of evidence that lexical doubling is due to parallelism comes from the fact that in so-called *inverted CSs* (ICSs), lexical doubling is rendered impossible.³⁶

- (60) (*kull mā) ha-ti-fham ?aktar kull mā
 every COMP FUT-IPFV-understand.2SGM more every COMP
 ti-?rā ?aktar
 IPFV-read.2SGM more
 ‘You will understand more, the more you read.’

Absence of lexical doubling in ICSs can be accounted for under the present analysis if we assume that it is triggered by a [correlative] feature in the structure, a feature that is available when the adverbial subordinate clause is a high adjunct, but not when it is merged as a low adjunct within *vP*. An implementation of this proposal is given under the cartographic analysis discussed in Section 5.³⁷

A third argument in favor of analyzing lexical doubling as a reflex of PF parallelism may be constructed with reference to one other case of subordination in the language. As it turns out, the subordinator *zayy* (= ‘as’), which marks manner adverbial clauses, can also be lexically doubled in contexts such as the one illustrated in (61) below (notice that *mā* occurs here as well).

- (61) zayy mā ti-?ul-l-ī (zayy mā) ha-šmil
 as COMP IPFV-say.2SGM-to-me as COMP FUT-IPFV.do.1SG
 Literally: ‘As you tell me, as I will do.’

The availability of lexical doubling in other cases of subordination gives further plausibility to the assumption that lexical doubling is better understood as driven by a general principle of phonological parallelism. While the exact conditions under which parallelism is licensed in certain subordination structures, and not in others, are not clear and beyond the scope of this study, the evidence presented here shows that such an analysis is indeed on the right track.

36. ICSs, while semantically very similar, have been argued to be non-derivationally related to CSs. In this paper, I have chosen, though implicitly, to adopt that assumption. For a discussion of this issue, see Culicover and Jackendoff (1999), den Dikken (2005), and Taylor (2013).

37. Lack of lexical doubling (or other instantiations of parallelism) in ICSs is shown in other languages as well. This is the case in Swedish, as in the example in (i), cited in Dooley (2014).

- (i) Och de smaka-de allt bättre ju mer han satte i sig
 and they taste-PAST even better JU more he put.PAST in REFLEXIVE
 ‘And they tasted even better the more he stuffed into himself.’

In English ICSs, the main clause exhibits regular word order, with the comparative DegP appearing in situ, as opposed to the subordinate clause, where DegP displacement still takes place.

- (ii) You’ll understand more, the more you read.

To sum up this section, there is empirical evidence from imperatives, ICSs, and *zayy*-subordination, that lexical doubling in CSs, though unexpected at first glance, is better understood as the result of a post-syntactic operation of phonological parallelism. As such, lexical doubling is not incompatible with our assumption throughout that, within CSs, C1 is a subordinate clause, while C2 is a main clause.

4.3 Summary

In this section, I proposed an account of two types of CSs in EA, whereby their macro-syntax is that of subordination and their micro-syntax follows from two types of XPs that license a correlative feature of the complementizer head (a QP in Type A and a measure PP in Type B). Since a QP is able to unselectively bind a degree variable, Type A can give rise to either a scalar correlation interpretation if there is a comparative DegP in the structure, or to a non-scalar correlation interpretation, if a DegP is non-comparative or missing altogether. A measure PP, by contrast, does not give rise to a quantifier-variable relation, and the interpretation is that of a non-proportional correlation. The presence of an overt QP in SpecCP in Type A explains why DegP displacement is disallowed in EA. By contrast, Maltese, which does not use a QP in SpecCP, resorts to DegP displacement instead. Additionally, the occurrence of a SoC in scalar CSs is what we expect, given the presence of a comparative DegP in this type of CSs. Finally, the lexical doubling property in CSs receives an explanation in terms of post-syntactic phonological parallelism, and, as such, is compatible with our assumption that in CSs, C1 is a subordinate clause, whereas C2 is a main clause.

5. A cartographic analysis of the micro-syntax of correlation structures

Even though the account presented in the previous section has been shown to account for the grammatical properties of CSs in EA, in this section, I point out three challenges to this analysis. I then show that implementing the analysis within a cartographic approach to clausal structure can overcome these challenges.

First, as Abeillé and Borsley (2008) point out, if CSs indeed involve adjunction of C1 to C2, we should expect C1 to be optional, given the general optionality of adjuncts. But this is not the case: C1 is an obligatory part of the structure (cf. example (31)). Similarly, we should expect free ordering of C1 and C2, which is not possible without inducing a total change in meaning (cf. the data in (32)–(33)). The adjunction analysis thus makes the wrong predictions in this regard.

A second empirical challenge has to do with word order. Under the analysis presented here so far, if a CS is embedded, we would expect C1 to fully precede

all of C2, including the complementizer. But this is not the case, as the following example from EA shows.

- (62) *ʔil-duktūr bi-y-ʔül ʔinn kull mā ti-ʔʕab riyāḏa*
 the-doctor ASP-IPFV-say.3SGM COMP every COMP IPFV-play.2SGM sports
ʔaktar (kull mā) ʕihhit-ak ha-t-kūn ʔahsan
 more every COMP health-your FUT-IPFV-be.3SGF better
 ‘The doctor says that the more you exercise, the better your health will be.’

The same problem arises when there is an overt root complementizer such as the question particle *huwwa* (glossed Q) or the declarative particle *da* (glossed DECL), as shown by the following two examples.³⁸

- (63) *huwwa kull mā ʔaštayał saḑāt ʔaktar (kull mā)*
 Q every COMP IPFV.work.1SG hours more every COMP
ha-ʔaxud murattab ʔaḑlā?
 FUT-IPFV.take.1SG salary higher
 ‘Is it the case that the more I work, the higher my salary will be?’
- (64) *ṣaddaʔ-ī-nī mā-fi-š dāḑi ti-fakkar-ī*
 believe.IMP-2SGF-me NEG-there-NEG reason IPFV-think-2SGF
fi-l-mawḑūf. da kull mā ti-fakkar-ī fi-l-mušmila ʔaktar
 in-the-issue DECL every COMP IPFV-think.2SGF in-the-problem more
(kull mā) ha-ti-ʔʕab-ī ʔaktar
 every COMP FUT-IPFV-be.tired-3SGF more
 ‘Believe me! There is no reason to think about the issue. [It is the case that] the more you think about the problem, the more tired you will get.’

In fact, given our treatment of *mā* so far as a complementizer, it is not even clear where the complementizers *ʔinn*, *huwwa*, and *da* in (62)–(64) would be merged in such sentences. An adjunction analysis of CSs of the type presented in the previous section is thus too limited to account for these additional facts.

Finally, a third challenge, albeit technical, is that the adjunction structure proposed in the previous section assumes the presence of [correlative] features on the complementizer of each clause, but does not specify a mechanism via which the two clauses become actually ‘correlated’ (cf. footnote 31).

These three challenges, however, can be overcome if the adjunction analysis is simply implemented within a cartographic approach to the left periphery of clausal structure, along the lines first proposed in Rizzi (1997). In fact, a cartographic approach has been proposed for CSs by Iwasaki and Radford (2009) for English, by Borgonovo and Valmala (2010) for Spanish, and by E (2014) for Mandarin Chinese.

38. See Soltan (2011) for a short discussion of these left-peripheral particles.

Taylor (2006, 2013) also implements her analysis of English CSs in terms of an articulated domain of the left periphery that allows for a recursive CP structure. In the rest of this section, I propose a cartographic analysis of CSs in EA that overcomes the three challenges raised above. In addition, I show that such an approach also promises to allow us to investigate the information structure properties of CSs. In particular, I offer suggestive evidence that C1 serves as a topic, encoding old information, whereas C2 is a comment, expressing new information.

Rizzi (1997) proposes an articulated structure of the left periphery of clauses, whereby the traditional ‘CP’ category is comprised of multiple syntactic projections: ForceP, TopP, FocP, and FinP.³⁹ This cartographic approach to clause structure has given rise to various lines of research on syntactic phenomena, particularly those with implications for the information structure of sentences. In the rest of this section, I provide a cartographic implementation of the analysis of CSs presented in Section 4 within a general framework of minimalist feature checking (Chomsky 1995, 2000, 2001). More specifically, I propose that C1 is a topic occupying SpecTopP of C2. The Top head is where the [correlative] feature is located, which thus serves to link the two clauses together.⁴⁰ The CS markers are located under FinP: *mā* is a Fin head, with a [quant]ificational feature that is checked by either the QP *kull* (Type A) or the measure PP *ʕalā ʔadd* (Type B). The Force head marks clause-typing and is where an (embedded or root) complementizer (such as *ʔinn*, *huwwa*, or *da*) occurs.

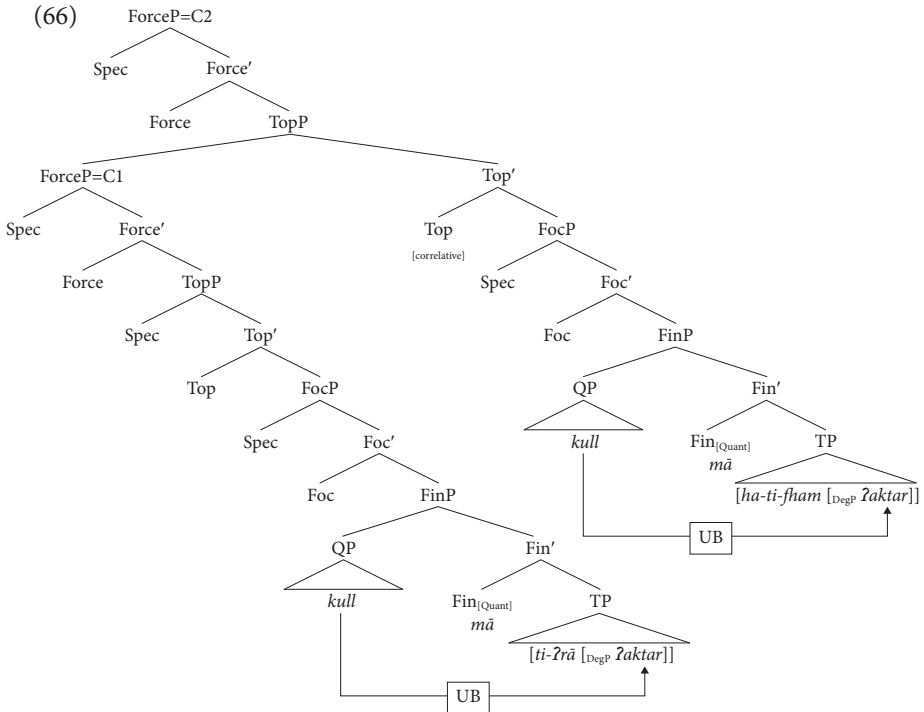
With these assumptions in mind, we can assign the cartographic structure in (66) to the scalar CS in (65), an example repeated from earlier for convenience.⁴¹

39. Force determines the clause type (e.g., declarative, interrogative, etc.). TopP and FocP host topics and foci, respectively. FinP is where finiteness is encoded.

40. This assumption is not meant to be a specific property of CSs; rather, it should be generalizable to other types of adverbial clauses such as temporal, locative, and conditional clauses, among several others. As such, the analysis proposed here promises to extend to cases of so-called high adjuncts (see Borgonovo and Valmala 2010 for this particular proposal with reference to data from Spanish). Notice further that this assumption should allow us to distinguish cases of high adjuncts from structures in which an adverbial clause appears as a low adjunct within the *vP*, where it does not serve as a topic of the structure (cf. Section 4.2 on lexical doubling). The syntax of adverbial subordination is clearly a complex issue and is beyond the scope of this paper, but the analysis proposed here seems empirically promising in multiple ways, a topic I hope to return to in a subsequent article.

41. The tree for a non-scalar CS should be identical to the one in (66), except that the QP binds an event variable in T. For CSs of Type B, the cartographic tree is also identical, except that the measure PP is in SpecFinP, checking the [quant] feature of Fin, and there is no quantifier-variable dependency. For considerations of space, I do not show these trees here. Notice, finally, that, for the sake of presentation, the Fin head of C2 and the QP in its specifier are both lexically

- (65) [_{C1} kull mā ti-ʔrā ʔaktar] [_{C2} (kull mā)
 every COMP IPFV-read.2SGM more every COMP
 ha-ti-fham ʔaktar]
 FUT-IPFV-understand.2SGM more
 ‘The more you read, the more you will understand.’



This cartographic tree has several empirical and conceptual advantages, which I highlight in turn.

First, it resolves the complementizer problem right away, since root and embedded Cs would appear in the Force position, whereas the adverbial clause complementizer *mā* is now treated as a Fin head. This resolves the word order problem arising with the adjunction analysis, and provides structural space that allows both Force and Fin to host lexical material.⁴²

represented in tree (66), even though I assume that their lexical content is the result of a process of phonological parallelism at PF, as argued for earlier in Section 4.2.

42. This cartographic proposal actually receives direct empirical support from Standard Arabic, where the main clause complementizer *ʔinna* can co-occur with *mā*, forming the compound complementizer *ʔinna-mā*.

Second, by encoding the [correlative] feature in Top, we are able to capture the syntactic dependency between both clauses, thereby resolving the technical problem pointed out earlier with the adjunction analysis (cf. footnote 31). Meanwhile, within each clause, unselective binding of a degree variable takes place, as in the adjunction analysis (see footnote 41 for notes on the derivation of non-scalar CSs and Type B CSs), giving rise to the scalar correlation interpretation.⁴³

Third, under the structure in (66), C1 (the subordinate clause) occupies SpecTopP, checking a feature of the Top head, and as such is an obligatory part of the structure, which allows us to avoid the optionality problem raised against the adjunction analysis by Abeillé and Borsley (2008).⁴⁴

Fourth, under the assumption that it is the [correlative] feature on Top that triggers the (optional) application of the phonological parallelism rule (cf. footnote 35), the absence of lexical doubling in ICSs follows (cf. example (60)), given that the subordinate clause in ICSs occurs within the *v*P, and not as a topic, in such contexts.

Fifth, this analysis has the potential of going beyond the syntactic properties of CSs to also address their information structure properties. While a full discussion of information structure is beyond the scope of this section, I provide suggestive evidence below showing that C1 typically expresses old information (hence, a topic), while C2 encodes new information (hence, a comment), along the lines proposed for Spanish CSs by Borgonovo and Valmala (2010).⁴⁵ Consider, for instance, the following exchange.

(67) *Speaker A:*

- a. Ahmad ʔayiz yā-kul fūl sudānī
 Ahmad want.PTCP.3SGM IPFV-eat.3SGM beans Sudanese
 ‘Ahmad wants to eat peanuts.’

Speaker B:

- b. māšī bass kull mā yā-kul fūl sudānī ʔaktar (kull
 Ok but every COMP IPFV-eat.3SGM beans Sudanese more every
 mā) ha-yi-ʔtaš ʔaktar
 COMP FUT-IPFV-get.thirsty.3SGM more
 ‘That’s fine, but the more peanuts he eats, the more thirsty he will get.’

43. Another potential empirical advantage of the cartographic analysis is that it can be extended to all other cases of high adjuncts, as pointed out in footnote 40.

44. Notice that, under the cartographic analysis, lack of DegP displacement in Type A of CSs still follows from relativized minimality, given that SpecFinP is filled by a QP. Thanks to an anonymous reviewer for raising this question.

45. Another potential empirical consequence of this cartographic approach is that it would allow us to explain empirical facts regarding the interaction between CSs and *wh*-questions in the language, a topic I hope to return to in future research.

Speaker C:

- c. *māšī bass kull mā ha-yi-ŧtaš ʔaktar (kull mā)
 Ok but every COMP FUT-IPFV-get.thirsty.3SGM more every COMP
 ha-ddī-l-u fūl sudānī ʔaʔall
 FUT-IPFV.give.1SG-to-him beans Sudanese less
 ‘#That’s fine, but the more thirsty he gets, the less peanuts I will give him.’

Speaker C’s response in (67c) is infelicitous because the ‘eating peanuts’ part occupies a focal position, even though it is part of the given information in Speaker A’s utterance. In addition, the ‘getting thirsty’ part sounds odd, because it sounds as if it were ‘old information’ in the exchange, even though it clearly is not. Speaker B’s response, by contrast, is felicitous, since the sequence of C1 and C2 within the CS provides a correlation between eating peanuts (given information) and getting thirsty (new information). The explanation of this contrast follows if the information structure of CSs is such that C1 hosts old information, whereas C2 encodes new information.

In addition, as Borgonovo and Valmala (2010) point out for Spanish data, topics can appear left-dislocated to either the embedded clause or the matrix clause, as illustrated by the positions of the DP topic *ʔil-muḥādra di* (‘this lecture’) in (68a, b).

- (68) a. ʔil-duktūr ʔāl ʔinn ʔil-muḥādra di kull ʔil-ṭalaba
 the-professor said.3SGM COMP the-lecture this all the-students
 lāzim yi-ḥdar-ū-hā
 obligated.PTCP IPFV-attend-3PL-it
 ‘The professor said that this lecture, all the students must attend it.’
 b. ʔil-muḥādra di ʔil-duktūr ʔāl ʔinn kull ʔil-ṭalaba
 the-lecture this the-professor said.3SGM COMP all the-students
 lāzim yi-ḥdar-ū-hā
 obligated.PTCP IPFV-attend-3PL-it
 ‘This lecture, the professor said that all the students must attend it.’

If C1 functions as a topic in CSs, we expect the same pattern of positioning to obtain, a prediction that is borne out by the data in (69).

- (69) a. ʔil-duktūr ʔāl ʔinn kull mā ti-lʔab riyāda ʔaktar
 the-doctor said.3SGM COMP every COMP IPFV-play.2SGM sports more
 (kull mā) šihhit-ak ha-t-kūn ʔaḥsan
 every COMP health-your.2SGM FUT-IPFV-be.3SGF better
 ‘The doctor said that the more you exercise, the better your health will be.’

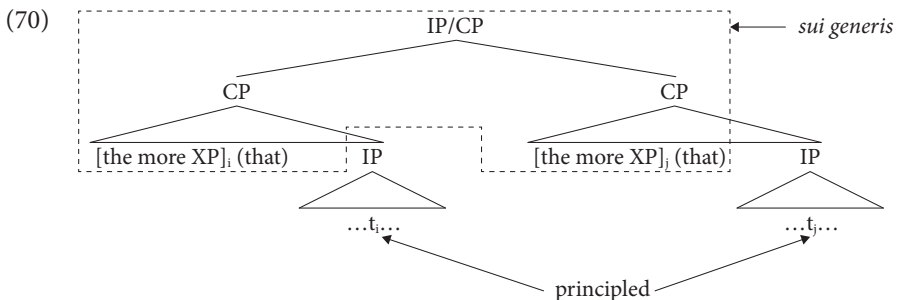
- b. kull mā ti-lʃab riyāda ʔaktar ʔil-duktūr ʔāl ʔinn
 every COMP IPFV-play.2SGM sports more the-doctor said.3SGM COMP
 (kull mā) ʃihhit-ak ha-t-kūn ʔahsan
 every COMP health-your.2SGM FUT-IPFV-be.3SGF better
 ‘The more you exercise, the doctor said that the better your health will be.’

It is reasonable then to conclude that in CSs, C1 functions as a topic, whereas C2 is a comment.

To sum up this section, I have shown that a cartographic approach to the micro-syntax of CSs overcomes the challenges facing an adjunction analysis, while still accounting for the grammatical properties of such structures. In addition, it has the potential of accounting for the properties associated with the information structure of CSs, particularly with regard to C1 encoding given information, as opposed to C2, which expresses new information.⁴⁶

6. Brief remarks on ‘constructions’

As pointed out in the earlier sections of this paper, Fillmore (1987), McCawley (1988), Culicover and Jackendoff (1999), and multiple works by Borsley and associates, argue that CSs should be analyzed as ‘constructions’ within the grammar, pointing out that they represent a challenge to syntactic analysis within the principles and parameters framework due to their grammatical idiosyncrasies. To use Culicover and Jackendoff’s term, CSs are *sui generis* constructions, because they challenge one of the key assumptions of syntactic analysis, namely, endocentricity, as shown by the tree in (70), which they assign to English CSs.



46. The information structure of CSs and other types of high adjunct clauses is indeed an area that is worthy of further investigation in Arabic dialects.

The analysis presented here for CSs in EA argues against the need for constructions. CSs in EA are indeed compositional. They exhibit subordination properties. The one ‘idiosyncratic’ feature of lexical doubling receives a principled account in terms of parallelism, a property that seems to hold of other structures in human languages at the syntax-phonology interface. The quantifier-variable dependencies in Type A follow from the presence of a QP. The lack of such a dependency in Type B follows from the presence of a measure PP. A more fine-grained approach to the cartography of the clausal left periphery accounts for a wider range of empirical facts of CSs in the language, while also providing insights into their information structure properties.

If correct, the analysis presented here provides evidence that CSs follow from the interaction between general principles of grammar (e.g., variable binding, parallelism, and cartography) and the features of the morphosyntactic heads involved (e.g., the quantifier *kull* and the functional head *mā* in EA). Their semantics should be read off the syntax, along the modified lines suggested by Bhatt (2009) for the semantic analysis of Beck (1997). As such, CSs are not *sui generis*. If anything, they provide evidence for a compositional theory of syntax where lexical and functional morphemes are the only building blocks for syntactic structure, and where constructions are strictly epiphenomenal.⁴⁷

7. Conclusion

In this paper, I have discussed two types of CSs in EA, referred to as Type A and Type B. On the macro-syntax level, I have provided multiple syntactic facts showing that both types involve subordination. While most of the grammatical properties of CSs can be accounted for under an adjunction analysis whereby the first

47. An anonymous reviewer points out that the proposed analysis does not necessarily provide an argument against construction-based approaches. This is not an inaccurate assessment, given that the goal of this paper is to show that the facts of CSs in EA are indeed derivable within a framework where heads are the only atoms in the syntax and where positing a construction for such structures in the grammar is unnecessary. I should point out, however, that construction-based approaches have to posit both heads and constructions as primitives of the grammar, whereas head-based approaches of the type adopted in this paper require only the presence of heads as atoms. So, if both approaches can equally account for the same set of facts in CSs, there is a conceptual argument to be made in favor of the latter over the former, at least to my mind. That said, a discussion of this particular issue is beyond the scope of this article. I, therefore, choose to refer the reader to the existing literature on both sides of the debate, such as Culicover and Jackendoff (1999), den Dikken (2005), Borsley (2011), and Taylor (2013), among several others.

clause is adjoined to the second, I have also argued that a cartographic approach, whereby the first clause is a topic and the second is a comment, is more empirically and theoretically adequate. I have shown that both types of CSs, while sharing the same syntactic skeleton, differ as a consequence of the type of XP that marks the structure at the left periphery: A QP in the case of Type A and a measure PP in the case of Type B. Absence of overt displacement of the comparative DegP in Type A follows trivially under the proposed analysis (SpecFinP being occupied by a QP), as does the occurrence of a SoC, an expected property of comparative DegPs in general. Type A involves a quantifier-variable dependency, which gives rise to either a scalar correlation in degree, or a non-scalar correlation between events. Type B, by contrast, is marked by a measure PP at the left periphery, is incompatible with a comparative DegP, and gives rise to a non-proportional correlation between fixed degrees. Finally, I have provided empirical evidence that the lexical doubling of the QP/PP and the complementizer *mā* is best understood as the result of a post-syntactic operation of phonological parallelism.

In sum, the data and analysis presented here add to the body of literature on CSs cross-linguistically, showing that such structures, while peculiar on the surface, do indeed follow from general assumptions of grammatical theory and they do not warrant being posited as ‘constructions’ in the grammar. As such, they support a theory of compositional syntax where ‘constructions’ are strictly derivatives, not primitives.

Acknowledgements

For insightful questions and comments, I would like to thank multiple members of the audiences at the *32nd Annual Symposium on Arabic Linguistics* at Arizona State University, the *49th Annual Conference on African Linguistics* at Michigan State University, the *36th West Coast Conference on Formal Linguistics* at University of California, Los Angeles, and the *Linguistics Unit at the Indian Institute of Technology Delhi, India*. I am most thankful to an anonymous reviewer for very detailed and insightful comments on the initial version of the manuscript that helped improve the article in multiple ways. I am especially indebted to Middlebury College for supporting my 2017–2018 academic leave, during which I was able to research the topic of this paper. I am thankful to Heather Taylor for first introducing me to comparative correlatives many years ago and for helpful conversations on the topic. Many thanks to Elly van Gelderen for her generous patience as the manuscript was written, revised, and finalized. Last but not least, I am very grateful to my Egyptian Arabic native speaker consultants for their generosity of time and gracious patience during the acceptability judgement tasks: Fatma Assef, Ahmad Ghazala, Nader Morkus, and Ahmad Soliman. Any errors herein remain solely my responsibility.

References

- Abeillé, A., & Borsley, R. D. (2008). Comparative correlatives and parameters. *Lingua*, 118, 1139–1157. <https://doi.org/10.1016/j.lingua.2008.02.001>
- Abu Ḥayyān (1256–1344). *Al-Baḥr Al-Muḥīṭ*. A 1993 edition. Beirut, Lebanon: Dar Al-Kutub Al-Ṣilmiya.
- Al-Khalil (718–786) & Sibawayh (ca. 765–796). *Al-Kitāb*. A 1988 edition. Cairo, Egypt: Al-Khangī Bookshop.
- Alqurashi, A. A., & Borsley, R. D. (2014). The Comparative correlative construction in Modern Standard Arabic. In S. Müller (Ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo* (pp. 6–26). Stanford, CA: CSLI Publications.
- Az-Zamakhṣari (1075–1144). *Al-Kaṣṣāf*. A 2009 edition. Beirut, Lebanon: Dar Al-Maʿrifa.
- Beck, S. (1997). On the semantics of comparative conditionals. *Linguistics and Philosophy*, 20, 229–271. <https://doi.org/10.1023/A:1005361901518>
- Bhatt, R. (2009). Comments on ‘Comparative correlatives in Greek: The syntax of *oso*’. In C. Halpert, J. Hartman, & D. Hill (Eds.), *MIT Working Papers in Linguistics 57: Proceedings of the 2007 Workshop on Greek Syntax and Semantics at MIT* (pp. 107–122). Cambridge, Mass.: MIT.
- Borgonovo, C., & Valmala, V. (2010). The syntax of Spanish comparative correlatives. In B. Hollebrandse, B. Kampers-Manhe, & R. Bok-Bennema (Eds.), *Romance Languages and Linguistic Theory 2008: Selected Papers from ‘Going Romance’ Groningen 2008* (pp. 17–35). Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.2.03bor>
- Borsley, R. D. (2003). On the Polish periphery: Comparative correlatives in Polish. In P. Banski & A. Przepiórkowski (Eds.), *GLIP-5: Proceedings of Fifth Generative Linguistics in Poland Conference* (pp. 15–28). Warsaw, Poland: Polish Academy of Science.
- Borsley, R. D. (2004). An Approach to English Comparative Correlatives. In S. Müller (Ed.) *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar* (pp. 70–92). Stanford, CA: CSLI Publications.
- Borsley, R. D. (2011). Constructions, functional heads, and comparative correlatives. In O. Bonami & P. C. Hofherr (Eds.), *Empirical Issues in Syntax and Semantics 8* (pp. 7–20). Available at: <http://www.cssp.cnrs.fr/eiss8>
- Chaves, R. (2012). On the grammar of extraction and coordination. *Natural Language and Linguistic Theory*, 30, 465–512. <https://doi.org/10.1007/s11049-011-9164-y>
- Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: The framework. In M. Roger, D. Michaels, & J. Uriagereka (Eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* (pp. 89–156). Cambridge, Mass.: MIT Press.
- Chomsky, N. (2001). Derivation by Phase. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Language* (pp. 1–52). Cambridge, Mass.: MIT Press.
- Culicover, P. W., & Jackendoff, R. (1999). The View from the periphery: The English comparative correlative. *Linguistic Inquiry*, 30, 543–571. <https://doi.org/10.1162/002438999554200>
- den Dikken, M. (2005). Comparative correlatives comparatively. *Linguistic Inquiry*, 36, 497–532. <https://doi.org/10.1162/002438905774464377>

- Dooley, S. (2014). The Swedish comparative correlative construction: *Ju... Desto...* and variations. In B. Lyngfelt & C. Wide (Eds.), *a special issue of Constructions*. Available at: <https://www.constructions.uni-osnabrueck.de/journal/vol-2014/>
- E, C. (2014). The Syntax of comparative correlatives in Mandarin Chinese. PhD Dissertation, University of Arizona.
- Fillmore, C. (1987). Varieties of conditional sentences. *ESCOL*, 3, 163–182.
- Heim, I. (1982). The semantics of definite and indefinite noun phrases. PhD Dissertation, University of Massachusetts, Amherst.
- Hsiao, S. (2003). On proportional correlative constructions in Chinese and Mongolian. *Journal of Taiwanese Languages and Literature*, 1, 239–267.
- Huber, S. (2007). Comparative correlatives – The case of German. In T. Crane, O. David, D. Fenton, H. J. Haynie, S. Katseff, R. Lee-Goldman, R. Rouvier, & D. Yu (Eds.), *Proceedings of the 33rd Annual Meeting of the Berkeley Linguistics Society* (pp. 220–231). <https://doi.org/10.3765/bls.v33i1.3529>
- Ibn Hišām (1309–1360). *Mughnī Al-Labīb*. A 1964 edition. Damascus, Syria: Dar Al-Fikr.
- Iwasaki, E., & Radford, A. (2009). Comparative correlatives in English: A minimalist-cartographic analysis. *Essex Research Reports in Linguistics*, Vol. 57.6. University of Essex, Colchester, UK. Available at: <http://repository.essex.ac.uk/96/>
- Jespersen, O. (1961). *A modern English grammar on historical principles*. London: George Allen and Unwin/Copenhagen: Munksgaard.
- Kapetangianni, K., & Taylor, H. L. (2009). Comparative correlatives in Greek: The syntax of *oso*. In C. Halpert, J. Hartman, & D. Hill (Eds.), *MIT Working Papers in Linguistics 57: Proceedings of the 2007 Workshop on Greek Syntax and Semantics at MIT* (pp. 91–105). Cambridge, Mass.: MIT.
- Leung, T. T-C. (2003). Comparative correlatives and parallel occurrence of elements. PhD screening paper, University of Southern California.
- Lewis, D. (1975). Adverbs of quantification. In E. Keenan (Ed.), *Formal Semantics of Natural Language* (pp. 3–15). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511897696.003>
- Lin, J. (2007). On the semantics of comparative correlatives in Mandarin Chinese. *Journal of Semantics*, 24, 169–213. <https://doi.org/10.1093/jos/ffm001>
- McCawley, J. D. (1988). The comparative conditional constructions in English, German and Chinese. In S. Axmaker, A. Jaissner, & H. Singmaster (Eds.), *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society* (pp. 176–187). <https://doi.org/10.3765/bls.v14i0.1791>
- Michaelis, L. (1994). A case of constructional polysemy in Latin. *Studies in Language*, 18, 45–70. <https://doi.org/10.1075/sl.18.1.o4mic>
- Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In E. J. Reuland & A. G. B. ter Meulen (Eds.), *The representation of (in)definiteness* (pp. 98–129). Cambridge, Mass.: MIT Press.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of Grammar* (281–337). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7
- Ross, J. R. (1967). Constraints on variables in syntax. PhD Dissertation, MIT.
- Soltan, U. (2009). Agree in the functional domain: Evidence from the morphosyntax of positive and negative imperatives in Standard Arabic. In A. Ojo & L. Moshi (Eds.), *Selected Proceedings of the 39th Annual Conference on African Linguistics* (pp. 86–99). Somerville, MA: Cascadilla Proceedings Project. Available at: <http://www.lingref.com/cpp/acal/39/index.html>

- Soltan, U. (2011). On strategies of question-formation and the grammatical status of the Q-particle *huwwa* in Egyptian Arabic wh-questions. In L. Friedman (Ed.), *University of Pennsylvania Working Papers in Linguistics Volume 17: Issue 1, Article 24*. Available at: <http://repository.upenn.edu/pwpl/vol17/iss1/24>
- Soltan, U. (2019). On Two Types of Correlation Structures in Egyptian Arabic. In R. Stockwell, M. O'Leary, Z. Xu, and Z. L. Zhou (Eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics* (pp. 316–325). Somerville, MA: Cascadilla Proceedings Project. Available at: <http://www.lingref.com/cpp/wccfl/36/index.html>
- Srivastav, V. (1991). The syntax and semantics of correlatives. *Natural Language and Linguistic Theory*, 9, 637–686. <https://doi.org/10.1007/BF00134752>
- Taylor, H. L. (2006). Out on Good Syntactic Behavior. General Paper, University of Maryland, College Park.
- Taylor, H. L. (2013). Grammar deconstructed: Constructions and the curious case of the comparative correlative. PhD Dissertation, University of Maryland, College Park.

The complementizer layer in Standard Arabic revisited

Salem Albuhayri and Hamid Ouali

King Saud University / University of Wisconsin-Milwaukee

This paper revisits three issues related to the complementizer layer (CP) in Standard Arabic. We reexamine them against the backdrop of Shlonsky's 2000 analysis, and put forward a new proposal couched in Rizzi's 1997 split-CP hypothesis. First, we examine the apparent distributional and interpretive differences between *?inna* and *?anna*, and subsequently argue that the former is a lexical verum operator that projects a VerumP in the middle of the split-CP field, whereas *?anna* is an indicative force head. Second, the current work presents another view on the elements analyzed as agreement clitics by Shlonsky (2000). We argue that they are not for agreement but rather are expletive or resumptive pronouns. Evidence that they are not agreement derives primarily from contexts where they appear in coordinated structures with overt DPs. We then investigate extraction patterns for questions and focus in matrix clauses as well as embedded clauses and propose that the preverbal subject DP in SVO is externally merged in SpecTopP in the split CP while it binds a null *pro* in SpecvP. This analysis captures the ban on extraction across the subject in SVO in that it shows that this DP is already higher in the structural hierarchy than the positions dedicated for focus and wh-questions.

Keywords: left periphery, verum, complementizer agreement

1. Introduction

The left periphery in Standard Arabic (SA henceforth) has been the subject of many studies in the literature (e.g., Bakir, 1980; Ouhalla, 1994, 1997; Shlonsky, 2000; Aoun et al., 2010). This work takes in particular Shlonsky (2000) as a background for the investigation of various relevant phenomena. Among the issues Shlonsky attempts to account for is the complementizer position and agreement. He discusses *?anna* and *?inna*. Before we review his account, a precise description of the distribution of these two complementizers is in order. The literature shows that both appear in

embedded contexts, but they diverge in that *ʔinna* can also appear in unembedded contexts. In embedded contexts, the choice of the complementizer is determined by the matrix verb (Mohammad, 2000: 19). Among the predicates under which *ʔanna* appear are what traditional grammarians refer to as “verbs of hearts”,¹ as in (1a), and “verbs of perception”, as in (1b). *ʔinna* can appear unembedded as in (2a), or embedded under *qaala* ‘say’ as in (2b) and under “verbs of hearts” provided that its accompanied by the focus particle *la* as in (2c).²

- (1) a. ʕalim-tu ʔanna tʕ-tʕaalib-a muʒtahid-un
learn.PERF-1SG that the-student-ACC assiduous-NOM
‘I learned that the student is assiduous.’
- b. samif-tu ʔanna ʕalijj-an ʔiztaaz-a l-ixtibaar-a
hear.PERF-1SG that Ali-ACC pass.PREF-3SG.MASC the-exam-ACC
‘I heard that Ali passed the exam.’
- (2) a. ʔinna tʕ-tʕaalib-a qaraʔ-a l-kitaab-a
verily the-student-ACC read.PERF-3SG.MASC the-book-ACC
‘Verily, the student read the book.’
- b. qult-tu ʔinna ʕalijj-an qaraʔ-a l-kitaab-a
say.PERF-1SG verily Ali-ACC read.PERF-3SG.MASC the-book-ACC
‘I said that, verily, Ali read the book.’
- c. ʔa-ʕlam-u ʔinna-ka la-ta-quul-u l-ħaqq-a
1SG-know.IMPERF-IND verily-you F-2SG-say.IMPERF-IND the-truth-ACC
‘I know that, verily, you are telling the truth.’

2. The left periphery in SA

2.1 Complementizer distribution

Cross-linguistic evidence shows that an array of elements appear in the left periphery, including subordinating and relative complementizers, *wh*-expressions, quantifiers, topics, scrambled arguments and focalized constituents. (3) is an example from English (Shlonsky, 2000: 326).

1. This is a set of verbs that relate to people’s mental states with regard to the veridicality of the proposition the verb embeds. They fall within the realm of what Fintel and Heim (2011: 19) call Propositional Attitudes, following Russell (1940) and Hintikka (1969).

2. All examples, including proper names, are transcribed using IPA and its associated diacritics as delineated in Ladefoged and Johnson (2015).

- (3) She told me *that, in her class, this book*, students would never read.

The ordering of these constituents is constrained, and therefore several works (e.g., Brody, 1990; Culicover, 1992; Müller and Sternefeld, 1993; Shlonsky, 1994) have all pointed out that these ordering restrictions cannot be captured by the assumption of a single CP (the CP projection was first proposed by Chomsky (1986) in his extension of the relational notions of X-bar theory from lexical to functional categories).

The same problem was encountered before with the inflectional and verbal layers, IP and VP, respectively, which are cartographically decomposed into sub-layers so that they can assimilate various elements specific to each. The VP layer was broken down into two VP projections (VP-shell), one embedding the other, the motivation being the need to account for multi-argument predicates such as double object constructions (Larson, 1988). Shortly afterwards, Pollock (1989) argued for a similar dissection of the inflectional domain in which two projections are proposed: the higher encodes tense, and the lower encodes subject agreement. Following this cartographic approach, Rizzi (1997) puts forward the ‘Split-CP’ hypothesis, in which (4) below is proposed as a universal structure of the left periphery, whose parametric variation is confined to the absence of some projections, but not the way they are ordered. Under this proposal, each of the constituents in (3) is assigned to a designated position based on whether it is a topic or focus.

- (4) ForceP > TopicP* > FocusP > TopicP* > FinP³

Rizzi argues that complementizers can figure as heads of ForceP or FinP in (4).

Adopting Rizzi’s proposal as well as the view that the presence of agreement in phi-features requires an agreement projection (AgrP) in the clausal hierarchy, Shlonsky (2000) proposes (5) as the structure of the left periphery in Standard Arabic. His proposal derives from the distributional permutations demonstrated by sentences like (6).⁴

- (5) ForceP > TopicP > AgrκP > κP > TopicP* > FocusP...⁵

3. This hierarchy is refined in Rizzi (2001, 2004) and new functional heads are incorporated including Mod⁰ and Int⁰.

4. Note that (*) means that the projection is recursive, whereas caps stand for focused constituents.

5. For Shlonsky, κP is the projection where the complementizer originates before it moves to AgrP to check agreement as will be delineated in section (3); this projection is basically motivated by the need to account for agreement, and that the complementizer collects it as it rolls up to Force⁰.

- (6) a. *samiŋ-tu* *ʔanna* *ʕaliŋ-an* *ʔiztaaz-a* *l-ixtibaar-a*
 hear.PERF-1SG that Ali-ACC pass.PREF-3SG.MASC >the-exam-ACC
 ‘I heard that Ali passed the exam.’
- b. *samiŋ-tu* *ʔanna* *ʕaliŋ-an* *l-ixtibaar-a* *ʔiztaaz-a*
 hear.PERF-1SG that Ali-ACC the-exam-ACC pass.PREF-3SG.MASC
 ‘I heard that [as for] Ali, THE EXAM, he passed.’
- c. *samiŋ-tu* *ʔanna* *l-ixtibaar-a_i* *ʕaliŋ-un* *ʔiztaaz-a-hu_i*
 hear.PERF-1SG that the-exam-ACC Ali-NOM pass.PREF-3SG.MASC-it
 ‘I heard that, [as for] the exam, Ali passed it.’
- d. **samiŋ-tu* *ʔanna* *l-ixtibaar-a* *ʔiztaaz-a* *ʕaliŋ-un*
 hear.PERF-1SG that the-exam-ACC pass.PREF-3SG.MASC Ali-NOM
 ‘I heard that THE EXAM Ali passed.’

Shlonsky argues that *ʔanna*, and *ʔinna*, by extension, starts out as head of κ P, then raises to $\text{Agr}\kappa$ P and eventually to ForceP. He associates *ʔanna* with a formal feature [+F], whose phonological exponent corresponds to the accusative case morphology on the DP that follows it as in (6). This feature is not checked when *ʔanna* is followed by a focused element as in (6d), resulting in ungrammaticality. $\text{Spec}\kappa$ P is the position of the DP with the [+F] that matches the feature on C.

Shlonsky’s analysis predicts that neither complementizer can be preceded by Topics. While this is borne out in the case of *ʔanna* as shown by the ungrammaticality of the sentences in (7), *ʔinna* does not comply with this prediction, as it may be preceded by a topicalized DP as shown in (8a) and (8b), in main and embedded contexts respectively, although in embedded clauses the sentence is degraded.

- (7) a. **samiŋ-tu* [*l-ixtibaar-u* *ʔanna* *ʕaliŋ-an* *ʔiztaaz-a-hu_i*]
 hear.PERF-1SG the-exam-NOM that Ali-ACC pass.PREF-3SG.MASC-it
 ‘I heard that [as for] the exam, Ali passed it.’
- b. **samiŋ-tu* [*ʕaliŋ_i-un* *ʔanna-hu_i* *ʔiztaaz-a* *l-ixtibaar-a*]
 hear.PERF-1SG Ali-NOM that-him pass.PREF-3SG.MASC the-exam-ACC
 ‘I heard that [as for] Ali, he passed the exam.’
- (8) a. *ʔar-risaalat_j-u* *ʔinna* *ʕaliŋ-an* *ʔaŋʔaa-haa_j*
 the-letter-NOM verily Ali-ACC give.PERF.3SG.MASC-it
li- muhammad-in
 to- Mohammed-GEN
 ‘[As for] the letter, verily, Ali gave it to Mohammed.’
- b. *ʔ qult-tu* [*ʕaliŋ_i-un* *ʔinna-hu_i* *qara?-a* *l-kitaab-a*]
 say.PERF-1SG Ali-NOM that-him read.PERF-3SG.MASC the-book-ACC
 ‘I said that, [as for] Ali, verily, he read the book.’

(7) and (8) above coupled with the sentences below call into question the assumption that *?anna* and *?inna* share the same distribution. (9) illustrates constructions where *?inna* co-occurs with elements that are either in Force or in positions that separate *?inna* from Force⁰.

- (9) a. ?a ?inna-ka la-ta-quul-u l-?aqq-a
 Q verily-you F-2SG-say.IMPERF-IND the-truth-ACC
 ‘Do you, verily, say the truth?’
 b. ?alaa ?inna l-?amr-a la-?aqq-un
 CI verily the-thing-ACC F-truth-NOM
 ‘Verily, the matter is true.’

In (9a), *?inna* appears with the polar question particle *?a*, which either occupies Force⁰ or Int⁰ in (10), and gives the clause the force of a question.⁶ In either case, it is conspicuous that *?inna* itself is not a Force head since it is lower than *?a*. Rizzi (2001: 289) positions IntP between the upper topic projections as illustrated in the hierarchy in (10).

- (10) FORCE (TOP*) INT (TOP*) FOC (TOP*) FIN IP

In (9b), it is also preceded by the clause initiator *?alaa* (CI), which supposedly occupies Force⁰, marking the clause as declarative.⁷ This leads to the conclusion that *?inna* is not a force-expressing element. Therefore, based on the data in sentences (7) through (9), two questions arise: first, since *?inna* does not express force, what is its function? Second, what positions do *?anna* and *?inna* assume in the left periphery?

6. A reviewer raised a question about *?a* and whether it can be preceded by topics; *?a* can be preceded by topics as shown below, an observation that accords with the position Rizzi allocates for interrogative heads in (10).

- (i) ?ar-risaalat_j-u ?a ?arsal-ta-haa_j
 the-letter-NOM Q give.PERF-2SG.MASC-it
 ‘As for the letter, did you send it?’

The point we are making is that regardless of whether *?a* is in IntP or ForceP, in a sentence like (9a), it is evidence that *?inna* itself is not in Force⁰. We are not as much concerned with committing ourselves to a specific view on where *?a* is as we are with the fact that its presence above *?inna* means that the latter itself is not in Force⁰.

7. This clause initiator is an interjection-like element, designated in traditional grammarians’ terms as “*?alaa ?al-istiftaahijja-tu*” (Omar et al., 1994). It is usually used to draw somebody’s attention to what is coming. In traditional grammar, its use is associated with emphasis on the proposition it embeds, and it does not appear in embedded clauses.

Based on Shlonsky (2000), they originate as heads of κP in (5), and then roll up the syntactic structure through head movement and join with Agr^0 in their way to $Force^0$. Agreement was a key factor in Shlonsky's analysis, as will be shown in section (3). With the new analysis of agreement we present there, Shlonsky's perspective is no longer warranted. Moreover, the evidence presented so far illustrates that there is a clear dichotomy between the two complementizers whose manifestation is the possibility of *?inna* appearing with another element, which either expresses force or at least shows that *?inna* does not reach $Force^0$. It follows from this that the assumption that *?anna* and *?inna* originate in the same position would have to shoulder the burden of coming up with a justification for why only *?anna* obligatorily moves to $Force^0$, whereas *?inna* moves there only optionally. In fact, there is a semantic difference between the two complementizers, alluded to in Shlonsky (2000), that might further justify an analysis that assumes a non-uniform base position. Shlonsky points out that *?inna* has the force of a strong affirmation or assertion translated by the English adverb *verily* or the French *certes* whereas *?anna* is neutral (p. 336). This offers a thread to reformulate the question on the functions of the two complementizers to be: how could these descriptive terms, *affirmation* and *emphasis*, be captured in a theory of the discourse layer of syntactic structure? To answer this question, let us take a short excursus into the notion of *verum*, which is dovetailed with *affirmation*. Gutzmann et al. (2017) indicate that this notion was first introduced by Höhle (1992) to designate cases where an accent is used to emphasize the truth of a proposition. (11) is an example from German.

- (11) Peter HAT den Hund getreten
 Peter has the dog kicked
 'Peter DID kick the dog' (Gutzmann et al., 2017: 4)

Several researchers give this phenomenon the designation "Verum Focus" (e.g., Höhle, 1992 (as cited in Gutzmann et al. (2017)), Lohnstein and Stommel, 2009; Krifka, 2008) and imply that it can only be realized by an accent on some expression in the sentence. However, Gutzmann et al. (2017) provide cross-linguistic evidence that the phenomenon is a case of a Verum Operator, which can either be realized by an accent or lexically. (12) shows examples from Spanish and Dutch, respectively.

- (12) a. Bien ha cantado la soprano
 indeed has sung the soprano
 'The soprano DID sing.'
 b. Ik heb het boek WEL gelezen
 I_{SG} have the book PRT read
 'I DID read the book.' (Gutzmann et al., 2017: 14)

In (12a), the verum operator is realized lexically by the clause-initial element *Bien* ‘indeed’, whereas in (12b) it is realized by the accented particle *WEL*.⁸ In SA, *?inna* usually figures in contexts where its use emphasizes the truth of the proposition it embeds; its use in an out of the blue context, where there is no contextual clue that the veracity of the proposition might be in dispute, is infelicitous. Therefore, following Gutzmann et al. (2017), we propose that *?inna* is a lexical verum operator whose function is to relate a proposition to a question under discussion (QUD), and is subject to the following felicity condition:

$$(13) \llbracket \text{VERUM} \rrbracket^c(p) = \surd, \text{ if } \{p, \neg p\} = \text{QUD}(c)^9 \quad (\text{Gutzmann et al., 2017: 9})$$

(13) means that a verum operator is only felicitous in a context where the veracity of the proposition given is at issue. The set $\{p, \neg p\}$, which simply includes a proposition and its negation, is the set of alternatives induced by the question ‘whether P?’ in the sense of alternative semantics developed in Rooth (1985, 1992) and Büring (1997). The contribution of the operator is to rule out the possibility of the proposition being false, i.e., it eliminates $\neg p$. This is how the notion of affirmation and emphasis associated with *?inna* is captured. What transpires from this approach to the function of *?inna* is that it is a lexicalization of verum. Therefore, we propose that the left periphery in SA houses a Verum Projection (VerumP) and that *?inna* is a lexicalization of the verum head. *?anna*, on the other hand, is assumed to originate in Force⁰. An analysis which assumes that *?anna* originates lower than Force⁰ would have to grapple with what the defining characteristics of this position are and what motivates *?anna* to always vacate it. Since *?anna* expresses indicative force and always appear clause-initially, it sounds more plausible to assume that it externally merges in Force.¹⁰

8. A reviewer questioned the position of verum crosslinguistically as it appears low in the structure in (12b) as opposed to its clause-initial position in (12a). Verum operates at the level of the proposition, but its correlated realization may appear anywhere in the clause, based on whether it is expressed lexically or by an accent, or by a combination of both. Lohnstein and Stommel (2009: 1) indicates that “the fronted finite verb can bear this accent (F-verum focus) or the complementizer (C-verum focus). In verb final clauses a similar effect can be achieved if an auxiliary bears the accent, but not a main verb.”

9. $\neg p$ denotes the negation of the proposition p .

10. Chomsky (1995, 1998) proposes that *Merge* should have preference over *Move* (*Merge over Move*) as an economy condition. Although this principle is primarily invoked to account for cases such as expletive-insertion to satisfy EPP and does not go without challenges (see Shima 2000 for a different view on *Merge over Move*), the basic principle can be extended to conceptually defend the view of *?anna* base-position in (15b). A derivation assuming that it originates lower and then moves obligatorily to Force⁰ would be less economical based on the lengthy path it

- (14) a. hind_k-un ?al-kitaab_j-u ?inna ?ahmad-a ?aft'aa-hu_j
 Hind-NOM the-book-NOM that Ahmed-ACC give.PERF.3SG.MASC-it
 la-haa_k
 to-her
 'Hind, the book, verily, Ahmed gave it to her.'
- b. ?ahmad_j-u hind_k-un ?inna-haa_k/(^{*}hu_j) ?aft'aat-hu_k
 Ahmed-NOM Hind-NOM that-her/(him) give.PERF.3SG.FEM-it
 l-kitaab-a
 the-book-ACC
 'Ahmed, Hind, verily, she gave him the book.'
- c. hind_k-un ?inna ?ahmad-a ?al-kitaab_j-u ?aft'aa-hu_j
 Hind-NOM that Ahmed-ACC the-book-NOM give.PERF.3SG.MASC-it
 la-haa_k
 to-her
 'Hind, the book, verily, Ahmed gave it to her.'

(14a) is a double object construction in which both objects are left-dislocated. Left dislocation targets SpecTopP, which means that *?inna* is preceded by two topic projections and followed by one if the preverbal subject is analyzed as a topic as well.¹¹ In (14b), the topic projection that follows *?inna* is shown to be recursive too. Therefore, we propose that the VerumP headed by *?inna* is situated between two recursive topic projections as in (15a), whereas *?anna* appears in Force⁰ as in (15b) since it cannot be preceded by any topic projection as shown above in (7).

- (15) a. ForceP>TopP*> VerumP *?inna* >TopP*>FocP>FinP>TP
 b. ForceP *?anna*>TopP*>FocP>FinP>TP

To summarize, as indicated, we propose that *?inna* can only appear in contexts where the question 'whether P?' is the question under discussion, and the question need not be stated explicitly; it might be implicit. Its function is to emphasize the truth of *p* over $\neg p$, the two alternatives induced by the polar question 'whether P?.'

takes as opposed to the one proposed here, especially when we take into account that (15b) is further tenable based on the reanalysis we present for agreement in Section 3, since agreement is the main catalyst for Shlonsky's argument for a lower base position for *?anna*. With the analysis we present for agreement along with the fact that *?anna* expresses force and always appears in Force⁰, Shlonsky's argument becomes economically unwarranted.

11. In the last section, we propose that the preverbal subject in SVO is a topic externally merged in SpecTopP.

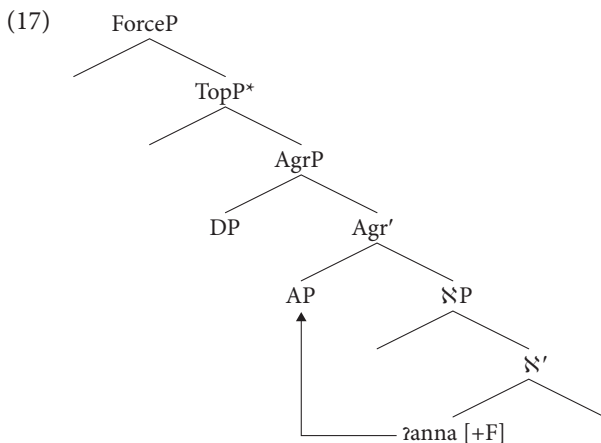
3. Complementizer agreement (CA)

3.1 Shlonsky's account

Shlonsky (2000: 334) illustrates that the formal feature [+F] on *ʔanna* can be satisfied by null nominal expressions whose presence is evident through agreement. In (16a), it is satisfied by a null referential pronoun, in (16b) by a null impersonal pronoun, and in (16c) by a subject wh-trace (glosses are maintained as they are in Shlonsky's work).

- (16) a. zaʔamtu ʔanna-hu kataba ʔal-risaalat-a
 (I) claimed that-[3MASC.SG] wrote the-letter-ACC
 'I claimed that he wrote the letter.'
- b. zaʔam-tu ʔanna-hu niima fi ʔal-sariir-i
 (I) claimed that that-[3MASC.SG] slept in the-bed-GEN
 'I claimed that it was slept in the bed.'
- c. man zaʔamta ʔanna-hu dʕaraba zayd-an
 who (you) claimed that that-[3MASC.SG] hit Zayd-ACC
 'Who did you claim that hit Zayd?' (Shlonsky, 2000: 334)

To explain the pattern in (16), Shlonsky assumes that the Comp System contains an AgrP above κP , to which κ^0 raises. SpecAgrP is an A-position, which is occupied by a nominal element (e.g., *pro*). The structure is given in (17).



For Shlonsky, the pronominal attached to *ʔanna* in every sentence in (16) is an agreement clitic that associates with *ʔanna* when it moves from κ^0 to Agr^0 . It marks agreement with a null nominal expression in SpecAgrP . His main argument for this account derives from the assumption that in all three cases the pronominal is not a topic by itself.

We argue that Shlonsky's account suffers from a number of theoretical and empirical shortcomings. First, the assumption of an agreement projection is theoretically undesirable in recent minimalist approaches. Chomsky (1995: 349) indicates that AgrPs have no semantic import whatsoever and are postulated for theory-internal reasons. Second, it is also theoretically undesirable to assume an A-position within the Comp System, as A-positions are usually associated with the thematic domain and case positions. An A-position in the left periphery neither has a thematic role, nor is it valued for case structurally; it gets its case either as the default case (nominative) or via valuation by a lexical category such as a complementizer or a preposition. Third, this analysis adheres to the view that agreement is checked in a *spec-head* configuration, which is abandoned in favor of the simple operation *Agree* under *c-command* (Chomsky, 2000, 2001). A major tenet of *Agree* is that *spec-head* configurations are not necessary for feature valuation.

What is more substantial than the theoretical issues above is that the assumed agreement-nature of the pronominals that appear on complementizers in these sentences is questionable. One piece of evidence against this comes from the observation that these pronouns can appear in Coordinate Structures conjoined with overt DPs. (18b) is an example¹²:

- (18) a. ʕalijj-un wa ʔahmad-u ʔʕnan-tu ʔanna-humaa
 Ali-NOM and Ahmed-NOM think.PERF-1SG that-them
 katab-aa d-dars-a
 write.PERF-3.DUAL the-lesson-ACC
 'Ali and Ahmed, I thought that they wrote the lesson.'
- b. ʕalijj-un ʔʕnan-tu ʔanna-[hu_i wa ʔahmad-a] katab-aa
 Ali-NOM think.PERF-1SG that-him and Ahmed-ACC write. PERF-3.DUAL
 d-dars-a
 the-lesson-ACC
 'Ali, I thought that he and Ahmed wrote the lesson.'

12. Similar examples are found in Ibn Maalik (1990: 629) where a clitic pronoun is conjoined with an overt DP.

(18b), in which the pronominal (-hu) is conjoined with *Ahmed*, poses a challenge to the view that these pronominals are agreement heads in AgrP, to which the complementizer raises and left-adjoins. If they are agreement elements, they would not have been expected to appear coordinated with other DPs. Compare (18) to (19) below (adapted from Soltan, 2007: 195).

- (19) ʔatay-tu *(ʔanaa) wa muḥammad-un
 come.PERF-1SG I and Mohammed-NOM
 ‘Mohammed and I came.’

In (19), the first person pronominal (-tu) attached to the verb is an agreement suffix and therefore is unconjoinable with the DP *Mohammed*.¹³ So, to gain some insight into what kind of elements these pronominals are and what function they assume in the structure, let us first review in the next section the range of elements that can follow the two complementizers in question.

3.2 What is it that follows ʔanna and ʔinna?

Various constituents can follow ʔanna and ʔinna to yield licit constructions. There is no strict requirement for them to be DPs.

- (20) a. ʔʕanan-tu ʔanna ʔaḥmad-a ʔar-risaalat_j-u
 think.PERF-1SG that Ahmed-ACC the-letter-NOM
 qaraʔ-a-haa_j
 read.PERF-3SG.MASC-it
 ‘I thought that [as for] Ahmed, the letter, he read (it).’
 b. ʔʕanan-tu ʔanna ʔar-risaalat_j-a ʔaḥmad-u
 think.PERF-1SG that the-letter-ACC Ahmed-NOM
 qaraʔ-a-haa_j
 read.PERF-3SG.MASC-it
 ‘I thought that [as for] the letter, Ahmed read (it).’
 c. ʕalim-tu ʔanna fii l-bayt-i razul-an
 know.PERF-1SG that in the-home-GEN man-ACC
 ‘I came to know that in the house is a man.’

13. In classical Arabic, the conjunction in (19) is also grammatical, in which case (-tu) is better treated as a pronominal clitic incorporated into the verb, as proposed in Fassi Fehri (1993).

- d. Ṣalim-tu ʔanna **raʒul-an** fii l-bayt-i
 know.PERF-1SG that man-ACC in the-home-GEN
 ‘I thought that a man is in the house.’
- e. Ṣalim-tu ʔanna **bi-l-ʔams-i** wasʕal-a
 know.PERF-1SG that in-the-yesterday-GEN arrive.PERF-3.DUAL
 ʔaḥmad-u
 Ahmed-NOM
 ‘I came to know that yesterday Ahmed arrived.’
- f. ʔʕanan-tu ʔanna l-**jawm-a** sa-ya-sʕil-u
 think.PERF-1SG that today-ACC will-3SG.MASC-arrive-IND
 ʔaḥmad-u
 Ahmed-NOM
 ‘I thought that today Ahmed would arrive.’
- g. Ṣalim-tu ʔanna **raʒul-an** tʕawiil-an fii l-bayt-i
 know.PERF-1SG that man-ACC tall-ACC in the-home-GEN
 ‘I came to know that a tall man is in the house.’

In (20a–b), the complementizer is followed by definite DPs, whereas in (20d–g), it is followed by indefinite ones. In (20c–f), it is followed by a prepositional phrase and an adverb. (20) is evidence that what comes after the complementizer can be a phrase of any type so long as it is not a verb, i.e., the complementizer cannot be followed by a VSO clause (Mohammad, 2000). In all these sentences, the complementizer does not inflect for agreement. What transpires from this is that the agreement analysis seems more structure-specific due to its dependency on the presence of a null *pro* in SpecAgrP. In other words, why does not a complementizer show agreement when there is an overt subject? (20) along with (18b) are an indication that the occurrence of pronominals in (16) might be due to the same reason as other DPs, and that their attachment to the complementizer is a PF cliticization that is ascribed to their phonological deficiency. So, the question that sequels is what is responsible for the distributional facts reflected in (20). In the next section, we take up this question, and then we provide an alternative analysis of the pronominals in (16).

3.3 Possible account

The presence of these diverse constituents after the complementizer can be ascribed to a requirement that the clausal semantic content be anchored to information structure (i.e., discourse layer) through a mechanism of feature checking (Franco, 2012; Sigurðsson, 2011, C/Edge-linking in Sigurðsson’s terms). In her account of stylistic fronting in Norwegian, Danish and Icelandic, which is viewed

as a fulfillment of clause anchoring to information structure, Franco associates two features with Fin^0 , namely [finiteness] and [definiteness].¹⁴ [Definiteness] can equally be checked by a [-definite] or [+definite] constituent, whereas finiteness can be checked by a temporal or locative adverb. Finiteness anchors the event time and location with the speech time and location, whereas definiteness anchors the clausal content to information structure, the common ground knowledge shared by all interlocutors. This proposal bears the seeds of an appropriate explanation, but cannot be extended to SA as it is for two reasons: First, it entails that *?anna* and *?inna* select FinP , which is not supported empirically. They both cannot appear directly on top of FinP following FocP as shown in (21a), nor can they directly be followed by FocP as in (21b). Rather, as proposed in (15), both appear higher than TopP . Second, one might argue that this account can be extended to SA if the base-position of both complementizers is taken to be Fin^0 and then they roll up the structure to Force^0 . This account, however, is theoretically untenable. It would have difficulty providing a conceptual justification for complementizer movement. Put differently, if both express force, what is the motivation for the assumption that they start out in Fin^0 ?

- (21) a. *kitaab-an ?inna t^c-t^caalib-a sa-ja-qra?-u
 book-ACC that the-student-ACC will-3SG.MASC-read-IND
 ‘A BOOK, verily, the student will read.’
 b. *?inna risaalat-an ?ahmad-u qara?-a
 that letter-ACC Ahmed-NOM read.PERF-3SG.MASC
 ‘Verily, THE LETTER, Ahmed read.’

Based on the hierarchies given in (15) and the facts in (20), the constituents following the complementizer in (20) must be higher than SpecFinP . This rules out the explanation provided by Franco’s proposal in its original formulation. In the next section, we develop a more elaborate account which has some basic commonalities with Franco’s.

14. This proposal has its conceptual roots in Haegeman (2006: 47) who, in her investigation of Romance clitic left dislocation as opposed to topicalization in English, concluded that Fin^0 might be responsible for the projection of the lower TopP in the split CP in Romance. Haegeman puts it as follows: “Based on the observation that control complements allow the (low) topicalized constituent and raising complements do not, I propose that the lower topic position depends on Fin , the lowest position of the left periphery. The analysis raises the question why the lower topic position is not available to English (and similar languages). One option is to directly associate the low topic position with the content of Fin . Possibly the rich mood system of Romance is encoded in Fin and contributes to the licensing of the lower topic position.”

3.4 Proposal

3.4.1 *Basic assumptions*

As indicated above, a reconsideration of the sentences in (16), for which Shlonsky designed his agreement account, reveals that the embedded clause in each is a null-subject clause, i.e., the subject argument is dropped. Argument drop is widespread across languages and is not confined to subjects since some languages, such as Pashto, Swedish and Chinese, allow object drop as well (Sigurðsson, 2011: 2f). This cross-linguistic structure has received several accounts, some of which were embedded in the Government and Binding framework (GB) (e.g., Rizzi, 1982), and some were embedded in the minimalist approach enriched with the utilization of cartography (e.g., Sigurðsson, 2010, 2011; Frascarelli, 2007; Sigurðsson and Maling, 2007, 2008).

A recurring theme in recent accounts is that argument drop is conditioned by the linkage of the dropped argument to discourse. In her investigation of null subjects in a set of languages with specific focus on Italian, Frascarelli (2007) argues that the interpretation of a thematic *pro* in the subject position is crucially dependent on the syntax and discourse properties of topic constituents. In this vein, Frascarelli proposes a tripartite classification of topics into *Aboutness Topics*, *Contrastive Topics* and *Familiar Topics*, which appear in the C-domain in this order.¹⁵ She then attributes argument linkage to discourse in the case of Italian third person null subjects to a [+aboutness] feature on the higher topic head. In an elaboration of the same line of argument, Sigurðsson (2011: 5) accommodates first and second person null subjects by his assumption of two other features, viz. [_A_A] and [_A_p], which are matched/valued by first and second person null subjects, respectively. He then states his generalization that “any definite argument, overt or silent, positively matches at least one C/Edge-linker in its local C-domain, where C/Edge-linkers include Top(ic) features and speech participant features (‘speaker’ and ‘hearer’).

Frascarelli (2007) argues that the constituent that satisfies the proposed [+aboutness] feature can either be a DP or a strong pronoun. Below is an illustrative example (p. 703).

- (22) [*il mio capo*]_i *come diceva Carlo [...]* *pro*_i *è un exreporter [...]* *pro*_i *è stato in giro per il mondo [...]* *pro*_i *mi ha preso in simpatia solo che siccome pro*_i *è mostruosamente lunatico, è capace che domani non gli sto più simpatica e pro*_i *mi sbatte fuori [...]* *comunque a parte questo pro*_i *mi diverte moltissimo – poi c’è M.F._k che*

15. The pursuit of whether these three types appear in SA in the same order is beyond the scope of this paper, but it suffices to mention that the definite nominative DPs that appear preverbally in SA are either subjects or clitic-left dislocated arguments, and they display flexibility in order (see Bakir, 1980; Shlonsky, 2000; Aoun et al., 2010).

è questo che appunto sta facendo tipo praticantato per poi andare a fare l'esame da giornalista/ fra un anno e mezzo quindi lui_k c'ha quanto meno la garanzia che pro_k può rimanere lì finché pro_k non farà l'esame cioè ehm lui_i poi gli deve fare scrivere le referenze...

'[my boss]_i as Carlo used to say [...] pro_i is a former reporter [...] pro_i has been all over the world [...] pro_i likes me, however, as pro_i is extremely moody, maybe tomorrow pro_i does not like me any longer and pro_i fires me [...] anyway, apart from this, pro_i is really funny – then there is M.F._k who is practicing for his exam as a journalist/ in one and a half years, so at least he_k has a guarantee that pro_k will stay there till pro_k has made the exam because he_i then must make/ write a report ...'

As can be seen in (22), the aboutness topic *my boss* is followed by a number of clauses with null co-indexed *pros* in their thematic domains. When the speech shifts to another topic, another individual with the name M.F, the immediate clause that follows has as an aboutness topic the strong pronoun *Lui* which bears the same index as M.F. This pronoun links the null *pros* that follow to their joint referent, namely M.F. Frascarelli (2007: 697) states that these topics qualify as clitic left dislocated constituents (CLLD), and hence are argued to be derived by base-generation in the C-domain. Frascarelli assumes that the relation between the aboutness topic, which is in an A-bar position, and the null *pro* in the thematic domain is established through Agree. Below is her schematization of how this relation holds (p. 718).

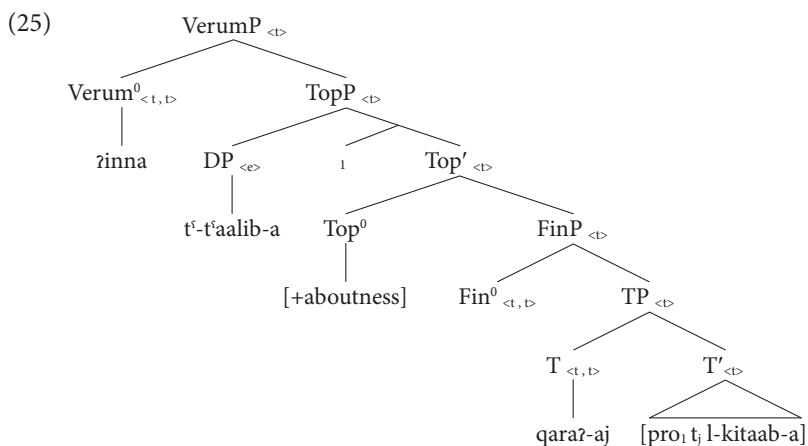
- (23) [ShiftP DP_[aPn] [Shift° [... [AgrSP [Agr° [vP pro_[aPn][VP]] ...]]]]]
 └──────────────────────────────────┬──────────────────────────────────┘
 AGREE
 [+aboutness] [φ-features]

Frascarelli's conceptualization of (23) is that the [φ-features] of *pro* are valued via Agree with the topic DP, and this is what is responsible for the identification of *pro*. In this spirit, she argues then that *pro* has the same function as clitic pronouns in topic constructions in that it serves as a pronoun, which resumes the topic (p. 693).

Sigurðsson (2010: 159ff) considers discourse-linking features as a form of EPP in the C-domain, and calls it the Filled Left Edge Effects (FLEE). Although EPP is no longer considered part of UG, the effects it was invoked to account for are still attested in languages, and the data above is an example. Formerly, when effects similar to the above were ascribed to EPP, some researchers (e.g., Nissenbaum, 1998; Sauerland, 1998; Butler, 2004) argue that this EPP has a semantic contribution in that it introduces a predicate abstract to bind some variable (e.g., a null *pro*, anaphoric pronoun, trace) internal to the thematic domain. Predicate abstraction is oftentimes concomitant with certain types of movement such as quantifier raising, but in these works, some forms of external merge are also taken to be accompanied by predicate abstraction.

Since the data we intend to account for is a mixture of third person subjects, as in (16a) and (16b), and expletives, as in (16c), we propose, following Frascarelli (2007), that Top^0 has a [+aboutness] feature. We add to this basic proposal the assumption that predicate abstraction is the semantic role this feature has in the structure. To see how this works, let us consider (24) below. The DP $t^s-t^s\text{aalib-a}$ ‘the student’ introduces a binder (λ -abstract) as shown in (25). The semantic denotation of the thematic domain yields (x read the unique z s.t. z is a book). The [+aboutness] on Top^0 introduces a binder that reopens the proposition to become $[\lambda x: x$ read the unique z s.t. z is a book], which allows it then to combine with the DP $t^s-t^s\text{aalib-a}$ ‘the student’ merged in SpecTopP .

- (24) $\text{?inna } t^s-t^s\text{aalib-a}$ [qaraʔ-a *pro* l-kitaab-a]
 that the-student-ACC read.PEFR3SG.MASC the-book-ACC
 ‘Verily, the student read the book.’



In addition to the assumption of a [+aboutness] feature on Top^0 , we maintain from Shlonsky the idea that *?anna* and *?inna* are associated with a formal feature [+F], which happens to be spelled out as accusative case when the following constituent is a DP. The reason to maintain this assumption is sentences like (20e) above where this feature is satisfied but with no overt morphology; there is no constituent in the domain of the complementizer that bears accusative case. With these assumptions in place, let us revisit the pronominals in (16) and provide an account of what they are and what their semantic contribution is.

3.4.2 Complementizer + pronominal

3.4.2.1 Null referential pronouns

The first case is that of referential pronouns as in (16a) repeated below.

- (26) zaḡamtu ḡanna-**hu** kataba *pro* ḡal-risaalat-a
 (I) claimed that-[3MASC.SG] wrote the-letter-ACC
 ‘I claimed that he wrote the letter.’

-hu satisfies the [+aboutness] on Top⁰ below *ḡanna*, and, as indicated, is responsible for the identification of *pro* through Agree. This [+aboutness] feature introduces a predicate abstract such that the end result of the semantic computation of (26) is as shown in (27):

- (27) *I claimed that* [λx : *x* wrote the letter] (1)

The index (1) in (27) denotes an individual whose value is determined by a contextual assignment function in the sense of Heim and Krazter (1998). An assignment function simply assigns to a variable a value drawn from a set whose members are part of the common ground shared by interlocutors. Possible values may include {Ahmed, Ali, Zayd...}. If the value given to (1) from the context is *Ahmed*, (26) would be true if and only if *I claimed that Ahmed wrote the letter*. Besides that observation that pronominals like (-hu) are conjoinable with other DPs as shown in (18), what the above discussion means is that the pronominal cannot be an agreement suffix as argued in Shlonsky (2000). Rather, it functions as a discourse linker in the sense of Frascarelli (2007) and Sigurðsson (2010, 2011), i.e., it links the null *pro* to an individual salient in the context of the utterance. The same argument can be run for (16c), repeated in (28).

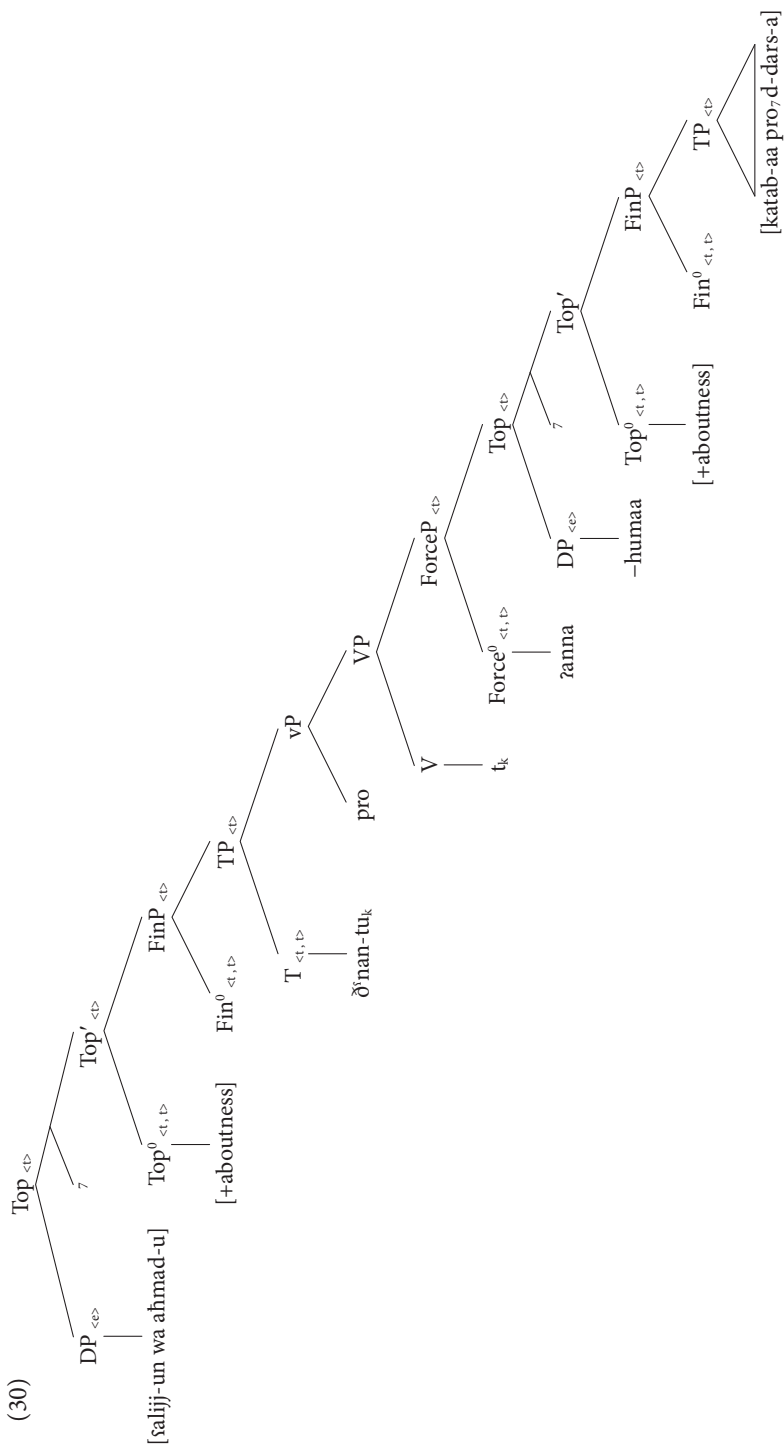
- (28) man zaḡamta ḡanna-hu dʿaraba *pro* zayd-an
 who (you) claimed that-[3MASC.SG] hit Zayd-ACC
 ‘Who did you claim that hit Zayd?’

The pronoun (-hu) identifies the null *pro* in the embedded thematic subject position, and links it to the *wh*-word in the matrix clause. This kind of analysis presupposes that *man* ‘who’ is externally merged in the matrix clause, i.e., it does not arrive there by movement (See Soltan (2007: 53) for a similar argument on using *man* to question an object in an SVO clause).

3.4.2.2 Anaphoric pronouns

Let us consider (18a), repeated in (29):

- (29) ḡalijj-un wa ḡahmad-u ḡʿnan-tu ḡanna-**humaa**
 Ali-NOM and Ahmed-NOM think.PERF-1SG.MASC that-them
 katab-aa d-dars-a
 write.PERF-3.DUAL the-lesson-ACC
 ‘Ali and Ahmed, I thought that they wrote the lesson.’



The satisfaction of [+aboutness] on the embedded Top⁰ introduces a predicate abstract such that the pronoun (-humaa) links the embedded subject *pro* to the matrix aboutness Topic ‘*Ali and Ahmed*’ which in turn satisfies [+aboutness] on the matrix Top⁰. This means that by the time the semantic component reaches in its computation to the matrix TopP, it yields the following statement:

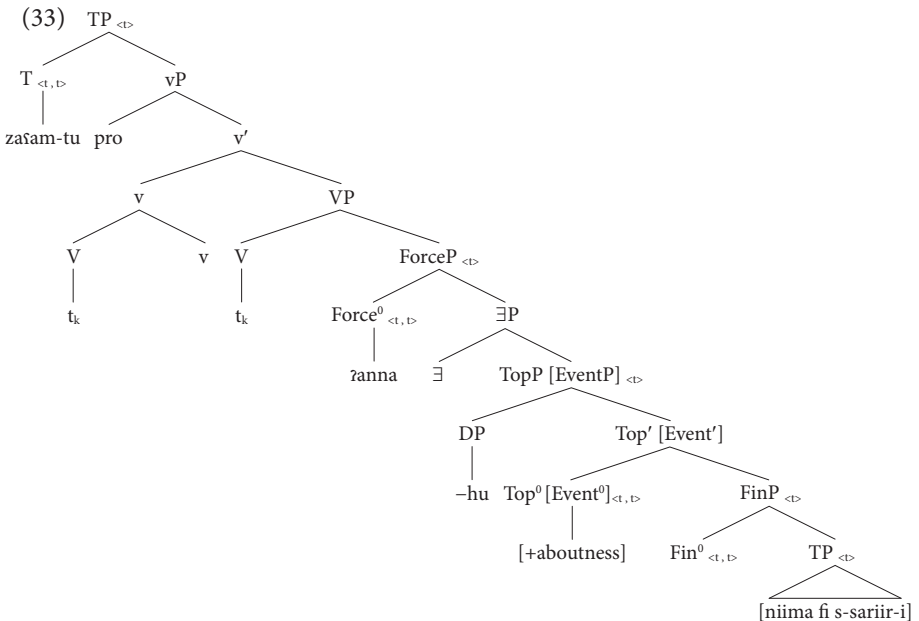
(31) I thought that [λx : x wrote the lesson] (*Ali & Ahmed*)

(31) shows that the pronominal in the embedded clause serves as a linker in a way that links the embedded subject to the matrix, as both are bound by the conjunction ‘*Ali and Ahmed*’.

3.4.2.3 Non-referential pronouns (expletives)

Expletives are among the elements whose presence has mostly been attributed to structural reasons. However, many analyses (e.g., Higginbotham, 1987; Stowell, 1991; Kratzer, 1995) argue that an expletive has a semantic import in that it serves the function of existential quantification over event or situation arguments. We follow them in this assumption. That is, when [+aboutness] is satisfied by an expletive as in (16b), repeated in (32), it introduces an existential quantifier that quantifies over events or situations.

(32) zaʃam-tu ʔanna-hu niima fi ʔal-sariir-i
 (I) claimed that-[3MASC.SG] slept in The-bed-GEN
 ‘I claimed that it was slept in the bed.’



The semantic computation of (32), in an approximation, amounts to: *I claimed that there exists an event or situation of sleeping in the bed.* The expletive anchors the utterance to an event argument.¹⁶

To conclude, we have argued in this section that the pronominals treated as agreement by Shlonsky are either variables, whose values are fixed referentially or anaphorically, or existential quantifiers over events in the case of expletives.¹⁷

16. A reviewer pointed out that this approach to expletives may be valid in this specific context but cannot naturally be extended to expletives such as “it” in sentences like (*it seems that...*). This may be so, but we would like to mention that a thorough discussion of expletives in all contexts is beyond the scope of the current work, and our attention is limited to those instances where an expletive is viewed as some form of complementizer agreement as argued in Shlonsky (2000). However, as cited in Butler (2004: 13), many researchers contend that expletives have the same semantics in all contexts, i.e., they are existential quantifiers. Butler points out that: “various people have convincingly propounded an alternative view of expletives that does take into account their semantic effects, where they relate not to the subject of vP, but rather its situation interpretation”.

17. We limited the discussion to DPs for the simple reason that they are the only type of constituents that are related to what Shlonsky calls complementizer agreement, which is one of the questions we are after. However, if an analysis of the locative and temporal adverbs in (20c), e and f is to be pursued, they can be analyzed in terms of Sigurðsson’s (2010) features of *Speech Time* and *Speech Location*. Sigurðsson (2010: 161f) proposes splitting Fin^0 into two separate head features which he terms as *Speech Time* and *Speech Location*, S_T and S_L respectively. These two features can be *negatively* matched, in Sigurðsson’s terms, by the expletive *there* or the time adverbial *then*, which dissociate the utterance from the *here* and *now* local to the speech event. Nevertheless, unlike Sigurðsson, an analysis along this line for SA would have to assume that these features are not located on Fin^0 , but are higher in the C-domain because these adverbs have to precede focus. Below is an example.

- (i) fii l-bayt-i qaabl-a fahd-un ʕalijj-an
 in the-home-GEN meet.PERF-3SG.MASC Fahad-NOM Ali-ACC
 ‘In the house, Fahd met Ali.’
- (ii) *ʕalijj-an fii l-bayt-i qaabl-a fahd-un
 Ali-ACC in the-home-GEN meet.PERF-3SG.MASC Fahad-NOM
 ‘ALI, in the house, Fahd met.’

The ungrammaticality of (ii) is due to the locative appearing below the focused phrase *Ali*. One might also argue that the locative in (i) is analyzable in terms of Sigurðsson’s *Speaker Feature* [Λ_A], which he assumes to be higher in the C-domain than *Speech Time* and *Speech Location*. This line of analysis is justified by the observation that the locative can have a speaker-oriented epistemic modality reading; the sentence under this reading would amount to saying “*given the knowledge that the speaker has, it was in the house that Fahd met Ali.*” Therefore, incorporating adverbs into the discussion would have to deal with all these observations and this would take us far afield, given the diverse semantic contributions of adverbs (see Cinque (1999) for an elaborate discussion of the diverse adverbial functions).

4. Extraction patterns

4.1 Ban on extraction across the preverbal DP in SVO

Fassi Fehri (1993) and Soltan (2007) point out that SVO structures do not allow extraction across the preverbal subject, neither for focus-preposing nor for wh-questions. Below are examples.

- (34) a. muḥammad-un qaraʔ-a kitaab-an
 Mohammed-NOM read.PERF-3SG.MASC book-ACC
 ‘Mohammed read a book.’
 b. *kitaab-an muḥammad-un qaraʔ-a
 book-ACC Mohammed-NOM read.PERF-3SG.MASC
 ‘A BOOK Mohammed read.’
 c. *maaḏaa muḥammad-un qaraʔ-a
 What Mohammed-NOM read.PERF-3SG.MASC
 ‘What did Mohammed read?’

The question is: how are focus-fronting and wh-questions derived in such constructions? Let us consider (35).

- (35) a. muḥammad-un kitaab-an qaraʔ-a
 Mohammed-NOM book-ACC read.PERF-3SG.MASC
 ‘Mohammed, a BOOK, he read’
 b. muḥammad-un maaḏaa qaraʔ-a
 Mohammed-NOM what read.PERF-3SG.MASC
 ‘Mohammed, what did he read?’

We take (34) and (35) as an indication that the preverbal DP is not base-generated in the thematic domain; rather, it is base-generated in an A-bar position in the left periphery. Soltan (2007) argues that this position is SpecTP, an A-bar position for him, and that it binds a null pro in SpecvP. The question is: how could Soltan’s analysis be extended to explain the patterns in (34) and (35)? Before we attempt to answer this question, let us expand our data by the examination of long-distance extractions from embedded clauses. Unlike root clauses, extraction across preverbal DPs from embedded clauses is unproblematic as shown in (36) below.

- (36) a. ḏʿanan-ta ʔanna muḥammad-an qaraʔ-a
 think.PERF-2SG.MASC that Mohammed-ACC read.PERF-3SG.MASC
 kitaab-an
 book-ACC
 ‘You though that Mohammed read a book.’

- b. maaḏaa ḏʿanan-ta ʔanna muḥammad-an
 what think.PERF-2SG.MASC that Mohammed-ACC
 qaraʔ-a
 read.PERF-3SG.MASC
 ‘What did you think that Mohammed read?’
- c. kitaab-an ḏʿanan-ta ʔanna muḥammad-an
 book-ACC think.PERF-2SG.MASC that Mohammed-ACC
 qaraʔ-a
 read.PERF-3SG.MASC
 ‘A BOOK, you thought Mohammed read.’

Extraction is problematic if it involves moving across a preverbal DP in the matrix clause as in (37).¹⁸

- (37) a. ʕalijj-un ḏʿann-a ʔanna fahd-an qaraʔ-a
 Ali-NOM think.PERF-3SG.MASC that Fahd-ACC read.PERF-3SG.MASC
 kitaab-an
 book-ACC
 ‘Ali thought that Fahd read a book.’
- b. *maaḏaa ʕalijj-un ḏʿann-a ʔanna fahd-an
 what Ali-NOM think.PERF-3SG.MASC that Fahd-ACC
 qaraʔ-a
 read.PERF-3SG.MASC
 ‘What did Ali think Fahd read?’
- c. *kitaab-an ʕalijj-un ḏʿann-a ʔanna fahd-an
 book-ACC Ali-NOM think.PERF-3SG.MASC that Fahd-ACC
 qaraʔ-a
 read.PERF-3SG.MASC
 ‘A BOOK Ali thought Mohammad read.’

4.2 Proposal

To account for the facts demonstrated in (34) through (37), we push Soltan’s 2007 analysis further and propose that preverbal DPs in SVO structures are base-generated in the left periphery in SpecTopP, rather than SpecTP. This position is higher than focus as shown in (15), and is assumed to be higher than wh-questions as well. These assumptions account for (34) and (35). It also follows from this proposal that the ban on extraction across preverbal DPs in root clauses is an inaccurate

18. Judgments on these sentences reflect the intuition of the authors and they correspond with the judgments reported in the literature for extraction across topics in simple clauses (e.g., Fassi Fehri, 1993; Soltan, 2007)

rendition of a fixed hierarchical order in the left periphery, namely that topics precede *wh*-questions and focus.¹⁹ The ban does not exist, otherwise (36), where the object is extracted all the way from its embedded position across the preverbal DP in its respective clause to the matrix clause, would remain unsolved mysteries. These long-distance extractions are assumed to proceed successive-cyclically through the edge of each phase up until the end position in the matrix clause. For example, the derivation of (36c) proceeds as shown in (38).

$$(38) \left[\text{FocP} [\text{kitaab-an}] \dots \text{Foc}^0 \left[\text{TP} \dots \text{T}^0 \left[\text{VP} \text{t}_1 \dots \nu^0 \left[\text{VP} \dots \text{V}^0 \left[\text{ForceP} \text{t}_1 \dots [\text{?anna}] \left[\text{TopP} \dots \left[\text{TP} \dots \text{T}^0 \left[\text{VP} \text{t}_1 \dots \nu^0 \left[\text{VP} \dots \text{V}^0 \text{t}_1 \right] \right] \right] \right] \right] \right] \right] \right] \right]$$

The movement in (38) does not violate subadjacency and is in line with the phase impenetrability condition (Chomsky, 2000). Extracted elements cannot remain in SpecForceP of embedded clauses, as this position is considered as a mere escape hatch.

5. Conclusion

We argued in this paper that, contrary to Shlonsky (2000), the two complementizers *?inna* and *?anna* assume two different positions in the left periphery in Standard Arabic, based on their distributional patterns as well as on their inherent semantics. *?inna* does not express Force but is rather a Lexical Verum Operator whose function is to emphasize the truth of a proposition (p) whenever the question ‘*whether p?*’

19. A reviewer pointed out that an assumption of a fixed hierarchy to explain the adjacency of focus to T is less explanatory than an explanation in terms of some derivational mechanisms, such as the traditional T-to-C movement. In fact, there are many ways to explain this adjacency without resorting to T-to-C movement. For instance, Miyagawa (2010: 12), based on a feature inheritance approach, argues that T inherits [focus] from C, hence the movement of focused constituents to SpecTP. Clearly, his proposal derives the adjacency requirement as a by-product of feature inheritance. What we have shown is that the so called pre-verbal subject in SVO is distributionally identical to CLLD in that they appear in an A-bar position, namely SpecTopP. This idea is not completely new to the literature as the same is argued for Italian by Frascarelli (2007). Besides the distributional facts presented in sentences (33) through (37), this proposal is further supported by the fact that a pre-verbal subject can alternate in position with a CLLD element unrestrictedly as shown below.

- (i) *?ar-risaalat-u l-walad-u katab-a-haa*
 the-letter-NOM the-boy-NOM write.PERF-3SG.MASC-it
 ‘The letter, the boy, he wrote it.’
- (ii) *?al-walad-u r-risaalat-u katab-a-haa*
 the-boy-NOM the-letter-NOM write.PERF-3SG.MASC-it
 ‘The boy, the letter, he wrote it.’

is the question under discussion. The two complementizers project in structures as follows:

- (39) ForceP > TopP* > VerumP ?*inna* > TopP* > FocP > FinP > TP
 ForceP ?*anna* > TopP* > FocP > FinP > TP

The fact that both complementizers cannot directly be followed by focus or verbs is put down to a number of features including a [+aboutness] feature on Top⁰, Speaker and Hearer features as well as Speech time and Speech Location features. These features require the presence of a constituent in the pre-FocP field. Definite and indefinite DPs as well as adverbials and PPs can satisfy these features. The pronominal clitics analyzed as agreement by Shlonsky (2000) are then taken to value the [+aboutness] feature on Top⁰, and hence are not for agreement. The semantic contribution of [+aboutness] is that it introduces a predicate abstract that binds a variable in the thematic domain. In the case of expletives, it introduces an existential quantifier which quantifies over events or situations. Finally, the ban on extraction across preverbal DPs in SVOs is reexamined, and two arguments are made. First, we argue that the preverbal DP in these structures is base-generated in SpecTopP in the left periphery, a position that precedes focus and wh-questions. Based on the positions of Focus and Wh-questions in matrix clauses and on the patterns of long-distance extractions, we argue that the ban on extraction across preverbal DPs does not exist.

References

- Aoun, J. E., Benmamoun, E., & Choueiri, L. (2010). *The syntax of Arabic*. New York: Cambridge University Press.
- Bakir, M. (1980). Aspects of clause structure in Arabic. (Doctoral Dissertation). Indiana University, Bloomington.
- Brody, M. (1990). Some remarks on the focus field in Hungarian. *UCL Working Papers in Linguistics*, 2, 201–225.
- Büring, D. (1997). *On the meaning of topic and focus*. London: Routledge.
- Butler, J. (2004). On having arguments and agreeing: Semantic EPP. *York Papers in Linguistics Series*, 2(1), 1–27.
- Chomsky, N. (1986). *Barriers*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1998). Some observations on economy in generative grammar. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky (Eds.), *Is the best good enough* (pp. 115–127). Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: The framework. In Martin, R., D. Michaels & J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in Honor of Howard Lasnik* (pp. 89–156). Cambridge, Massachusetts: MIT Press.

- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1–52). Cambridge, Massachusetts: MIT Press.
- Culicover, P. (1992). Topicalization, inversion, and complementizers in English. In D. Delfitto, M. Everaert, A. Evers & F. Stuurman (Eds.), *OTS working papers: Going Romance and beyond* (pp. 1–45). Utrecht: University of Utrecht.
- Fassi-Fehri, A. (1993). *Issues in the structure of Arabic clauses and word order*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-017-1986-5>
- Fintel, K. V., & Heim, I. (2011). *Lecture notes in intensional semantics*. Ms.: MIT.
- Franco, I. (2012). *Verbal stylistic fronting in Old Florentine*. Ms.: Leiden University.
- Frascarelli, M. (2007). Subjects, topics and the interpretation of referential pro. *Natural Language & Linguistic Theory*, 25(4), 691–734. <https://doi.org/10.1007/s11049-007-9025-x>
- Gutzmann, D., Hartmann, K., & Matthewson, L. (2017). Verum focus is verum, not focus.
- Haegeman, L. (2006). Argument fronting in English, Romance CLLD, and the left periphery. In R. Zanuttini, H. Campos, E. Herburger, & P. Portner (Eds.), *Crosslinguistic research in syntax and semantics: Negation, tense and clausal architecture* (pp. 27–52). Washington: Georgetown University Press.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Higginbotham, J. (1987). Indefiniteness and predication. In: E. Reuland & A. ter Meulen (Eds.), *The representation of (in)definiteness* (pp. 43–70). Cambridge, Massachusetts: MIT Press.
- Hintikka, J. (1969). Semantics for propositional attitudes. In J. W. Davis, D. J. Hockney & W. K. Wilson (Eds.), *Philosophical logic* (pp. 21–45). Dordrecht: Reidel. https://doi.org/10.1007/978-94-010-9614-0_2
- Höhle, T. N. (1992). Über Verum-Fokus im Deutschen. In J. Jacobs (Ed.), *Infonnationsstruktur und Grammatik* (pp. 112–141). Opladen: Westdeutscher Verlag. https://doi.org/10.1007/978-3-663-12176-3_5
- Ibn Maalik, M. (1990). *faarhu ʔtashiil* (volume 4). Beirut: Dar Al-kotob Al-ilmiyah.
- Kratzer, A. (1995). Stage and individual level predicates. In G. N. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 125–175). Chicago, IL: University of Chicago Press.
- Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55 (3–4), 243–276. <https://doi.org/10.1556/ALing.55.2008.3-4.2>
- Ladefoged, P., & Johnson, K. (2015). *A course in phonetics*. Connecticut, Stamford: Cengage Learning.
- Larson, R. K. (1988). On the double object construction. *Linguistic Inquiry*, 19 (3), 335–391.
- Lohnstein, H., & Stommel, H. (2009). Verum focus and phases. *Linguistic Analysis*, 35(1), 109–140.
- Miyagawa, S. (2010). *Why Agree? Why Move? Unifying agreement-based and discourse configurational languages*. Cambridge, Massachusetts: MIT Press.
- Mohammad, M. A. (2000). *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Philadelphia: John Benjamins Publishing. <https://doi.org/10.1075/cilt.181>
- Müller, G., & Sternefeld, W. (1993). Improper movement and unambiguous binding. *Linguistic Inquiry*, 24(3), 461–507.
- Nissenbaum, J. (1998). Movement and derived predicates: Evidence from parasitic gaps. *MIT Working Papers in Linguistics*, 25, 247–295.
- Omar, M., Zahran, M., & Abdullatif, M. (1994). *Al-Nahw-u Al-Asasijj-u ‘the basic syntax’* (4th Edition). Kuwait: Dar-u Al-Ssalaasil-i.

- Ouhalla, J. (1994). Verb movement and word order in Arabic. In D. Lightfoot & N. Hornstein (Eds.), *Verb movement* (pp. 41–72). Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511627705.004>
- Ouhalla, J. (1997). Remarks on focus in Standard Arabic. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 9–46.
- Pollock, J. Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, 20 (3), 365–424.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of grammar* (pp. 281–337). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7
- Rizzi, L. (2001). On the position “Int(errogative)” in the left periphery of the clause. In G. Cinque, L. Renzi, and G. Salvi (Eds.), *Current studies in Italian syntax* (pp. 287–296). Amsterdam: Elsevier.
- Rizzi, L. (2004). Locality and left periphery. In A. Belletti (Ed.), *Structures and beyond* (pp. 223–251). Oxford: Oxford University Press.
- Rooth, M. (1985). Association with focus. (Doctoral dissertation). University of Massachusetts, Amherst.
- Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, 1(1), 75–116.
<https://doi.org/10.1007/BF02342617>
- Russell, B. (1940). *An inquiry into meaning and truth*. London: George Allen and Unwin LTD.
- Sauerland, U. (1998). Two structures for English restrictive relative clauses: The meaning of chains. (Doctoral dissertation). MIT, Cambridge.
- Shima, E. (2000). A preference for Move over Merge. *Linguistic Inquiry*, 31(2), 375–385.
<https://doi.org/10.1162/ling.2000.31.2.375>
- Shlonsky, U. (1994). Agreement in comp. *The Linguistic Review*, 11(3–4), 351–376.
- Shlonsky, U. (2000). Remarks on the complementizer layer of Standard Arabic. In J. Lecarme, J. Lowenstamm & U. Shlonsky (Eds.), *Research in Afroasiatic grammar: Papers from the third conference on Afroasiatic languages* (pp. 325–343). Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.202.15shl>
- Sigurðsson, H. Á. (2010). On EPP effects. *Studia Linguistica*, 64(2), 159–189.
- Sigurðsson, H. Á. (2011). Conditions on argument drop. *Linguistic Inquiry*, 42(2), 267–304.
https://doi.org/10.1162/LING_a_00042
- Sigurðsson, H. Á., & Maling, J. (2007). On null arguments. In M. C. Picci, & A. Pona (Eds.), *Proceedings of the XXXII Incontro di Grammatica Generativa* (pp. 167–180). Firenze: Edizioni dell’Orso.
- Sigurðsson, H. Á., & Maling, J. (2008). Argument drop and the empty left edge condition (ELEC). *Working Papers in Scandinavian Syntax*, 81, 1–27.
- Soltan, U. (2007). On agree and postcyclic merge in syntactic derivations: First conjunct agreement in Standard Arabic. In E. Benmamoun (Ed.) *Perspectives on Arabic Linguistics XIX* (pp. 191–213). Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.289.15sol>
- Soltan, U. (2007). On formal feature licensing in minimalism: Aspects of Standard Arabic morphosyntax. (Doctoral dissertation). University of Maryland, College Park.
- Stowell, T. (1991). The alignment of arguments in adjective phrases. *Syntax and Semantics*, 25, 105–135.

Sluicing and sprouting in Jordanian Arabic

Juman Al Bukhari

University of North Georgia

This paper investigates elliptical constructions sluicing and sprouting in Jordanian Arabic (JA), and describes the structure of the sluice clause, which depends on the type of *wh*-expressions on the one hand, and whether the preposition at the sluice is pied-piped on the other hand. The central research question is how such constructions are derived: via ordinary sluicing or via pseudosluicing. Cases where a preposition is pied-piped with the *wh*-word in the remnant require an ordinary sluicing analysis; pied-piped PPs are not otherwise found in *wh*-clefts in JA. However, in cases where a preposition is not pied-piped, the only possible underlying derivation is pseudosluicing, with a *wh*-cleft structure. I also discuss sprouting, in which the *wh*-word has no corresponding overt correlate in the antecedent clause.

Keywords: sluicing, pseudo-sluicing sprouting, preposition stranding generalization, Jordanian Arabic

1. Introduction

Sluicing was first discussed (and named) by Ross (1969). An example is shown in (1).

- (1) Jack bought something, but I do not know what [~~Jack bought~~].
(Merchant 2001, p. 1)

Sluicing is widespread cross-linguistically. The interrogative phrase in the second clause is interpreted as if it were part of an embedded constituent question, identical in form to the first clause but with the *wh*-phrase in place of the corresponding phrase in the first clause (*something*, in (1)).

There have been two general approaches to sluicing in the literature. The first approach is the nonstructural approach, which posits no syntactic structure in the ellipsis site (Ginzburg and Sag, 2000; Culicover and Jackendoff, 2005). The second approach is the structural approach, which assumes a fully articulated but unpronounced syntactic structure in the ellipsis site (Merchant, 2001). In this paper, I will adopt the second approach.

There are few studies of Arabic elliptical constructions in the literature, and as such the topic is relatively unsettled. Notable exceptions include Algryani's (2010, 2011a, 2011b, 2013) work on P-stranding, sluicing, VP ellipsis, and stripping in Libyan Arabic (LA), and Leung's (2014a, b) research on sluicing in Emirati Arabic (EA). The two elliptical constructions I will investigate in this paper are sluicing, which I have described above, and sprouting, which is a sub-type of sluicing.

Sprouting differs from sluicing in that the *wh*-word finds no corresponding overt antecedent in the antecedent clause; an example, along with its intended interpretation, is shown in (2).

- (2) a. John was eating, but I am not sure what.
 b. John was eating, but I am not sure what [~~he was eating~~].

JA exhibits both sluicing and sprouting, as shown in (3) and (4):

- (3) *ḥasan ftara: ?ifi:_i bas ma: b-a-ʃraf fu:_i*
 Hasan bought.3ms something.ms but NEG Asp-1s-know what
 [~~*ḥasan ftara:*~~]
 Hasan bought.3ms
 'Hasan bought something, but I do not know what [~~Hasan bought~~].'
- (4) *ḥasan ftara: bas ma: b-a-ʃraf fu: [~~*ḥasan ftara:*~~]*
 Hasan bought.3ms but NEG Asp-1s-know what Hasan bought.3ms
 'Hasan bought, but I do not know what [~~Hasan bought~~].'

In order to investigate the properties of sluicing and sprouting in JA, we will need to attend closely to the behavior of different kinds of *wh*-phrases. *Wh*-arguments, *wh*-adjuncts, *wh*-PPs, and *which*-NPs do not all participate in the same range of *wh*-movement constructions in JA. Unlike English, JA does not permit preposition stranding, a fact that will be important for the analysis of examples where the correspondent of the *wh*-phrase is the object of a preposition. JA also exhibits question formation strategies (*wh*-fronting and *wh*-clefting) different from those found in other Arabic dialects.¹

First, I would like to introduce some terminology that has been conventionally used in the literature on sluicing. I will refer to the *wh*-interrogative that survives the ellipsis process as the "remnant"; in (3), this is *fu*: 'what'. Following Toosarvandani (2008), I call the portion of the sluiced clause that undergoes ellipsis the "target"; in (3), this is *ḥasan ftara* 'Hasan bought'. Both the remnant and the target make the "sluice." The remnant corresponds to an overt constituent in the antecedent

1. It would be interesting to see how Arabic dialects that do not exhibit the *wh*-fronting strategy (e.g. Egyptian Arabic) behave under sluicing.

clause called the “correlate”; in (3), this is *?ifi*: ‘something’. When the remnant lacks an overt correlate, as in (4), we have the sub-type of sluicing known as sprouting.

It is crucial to distinguish between two possible ways of deriving sluicing: ordinary sluicing and pseudo-sluicing. The former is formed by a fully-fledged CP, as in (1), while the latter is formed by a *clefted question* (Merchant, 1998: p. 91), as in (5), where the pseudo-slucose arises from the deletion of the subject *it* and the copula *was*, unlike true sluicing which arises from TP-ellipsis.²

- (5) John bought something but I do not know what [~~it-was~~].

Merchant (1998) makes this proposal for Japanese following Whitman & Yi (1996) for Korean and Adams (2004) for Mandarin Chinese, which Van Craenenbroeck (2010) called a pro-analysis.

Since *wh*-fronting and *wh*-cleft play a role in the underlying derivation of the sluice, it is necessary to discuss both *wh*-fronting and *wh*-clefting in JA in the next section. In Section 2, I will give a brief background on the types of questions in JA. Then, I will discuss the issue of sluicing in JA in Section 3 and sprouting in JA in Section 4. Finally, I conclude with a summary of findings on sluicing and sprouting.

2. *Wh*-questions in JA

Wh-questions in JA can be derived in two ways: via *wh*-fronting and via *wh*-clefting (Wahba, 1984; Shlonsky, 1997; Aoun et al., 2010). *Wh*-fronting leaves a gap behind, while *wh*-clefting involves the use of a resumptive pronoun; the two strategies are shown in (6).

- (6) a. Gap strategy (*wh*-fronting)
fu:_i *stare*:-*t* *t*_i *mba*:*reh*?
 What bought-2ms yesterday
 ‘What did you buy yesterday?’
- b. Resumptive strategy (*wh*-cleft)
fu:_i *huww* *illi* *stare*:-*t-o*_i *mba*:*reh*?
 What 3ms.COP that bought-2ms-it yesterday?
 ‘What was it that you bought yesterday?’ Cleft Structure (Eid, 1983)

2. The example in (5) is not a plausible analysis for English, but rather a schematic demonstration of a pseudo-sluicing derivation. Pseudo-sluicing is found in, e.g., Japanese, where null-subject *wh*-clefts like (i) can serve as the underlying source for pseudo-sluicing.

- (i) *Dareka-ga sono hon-o yon-da ga, watashi-wa dare data ka wakaranai.*
 someone-NOM that book-ACC read-past but, I-top who was Q know.not
 ‘Someone read that book, but I don’t know who it was.’ (Merchant, 1998, p. 91)

The two types of *wh*-question derivation allow different classes of *wh*-phrase to participate, and this plays a role in determining the underlying derivation of the sluice. The types of phrases that can undergo *wh*-fronting in JA include simple *wh*-expressions like *fu*: ‘what’ and *mi:n* ‘who’, as in (6), along with *wh*-PPs, *wh*-adjuncts, and *which*-NPs, as in (7).

- (7) a. *bi-ʔajj za:mʕa daras ʕumar*
 in-which university studied.3ms Omar
 ‘At which university did Omar study?’
 b. *ki:f xallas^ʕ-t ir-risa:le*
 how finished-2ms the-dissertation
 ‘how did you finish the dissertation?’
 c. *ʔajj taxass^ʕos^ʕ daras ʕumar*
 which major studied.3ms Omar
 ‘which major did Omar study?’

On the other hand, *wh*-clefts in JA allow only *wh*-arguments (bare *wh*-words and *which*-NPs), as in (8) and (9), respectively. The remaining *wh*-expression types – pied-piped *wh*-PPs and *wh*-adjuncts, as in (9) and (10) – are ungrammatical in *wh*-clefts. Note that *wh*-PPs are acceptable in *wh*-clefts when the preposition is stranded and followed by a resumptive pronoun, as in (9b).

- (8) a. *fu:_i huwwe illi ʔaxad-o_i ʕumar*
 what 3ms.COP that took.3ms-it Omar
 ‘What is it that Omar took?’
 b. *mi:n_i huwwe illi ʔaxad-o_i ʕumar*
 who 3ms.COP that took.3ms-him Omar
 ‘Who is it that Omar picked up?’
 (9) a. **bi-ʔajj za:mʕa hijje illi laʔe:-t ʕumar*
 in-which university 3fs.COP that found-2ms Omar
 ‘In what university is it that you found Omar?’
 b. *ʔajj za:mʕa hijje illi laʔe:-t ʕumar fi-*(ha)*
 which university 3fs.COP that found-2ms Omar in-it
 ‘what university is it that you studied linguistics at?’
 (10) a. **ki:f hijje illi xallas^ʕ-t ir-risa:le*
 how 3fs.COP that finished-2ms the-dissertation
 ‘how is it that you finished the dissertation?’
 b. **we:n huwwe illi ʔaxad-o*
 where 3ms.COP that took.3ms-him
 ‘where did he take him/it?’

Wh-fronting thus permits a superset of the *wh*-phrase types found in *wh*-clefts in JA. Having introduced these basic facts about JA questions, I will discuss sluicing in the next section.

3. Sluicing in JA

Ordinary sluicing is derived via *wh*-fronting and pseudo-sluicing via *wh*-clefting. Given the fact that these two constructions permit different sets of *wh*-expressions in JA, as discussed above, we can use the distribution of permissible *wh*-expressions to probe the underlying sources of JA sluicing constructions.

Recall from Example (3), repeated in (11), that the underlying sluice site can be analyzed as an instance of *wh*-fronting. This involves *wh*-movement followed by deletion of the target, i.e. IP deletion (Merchant, 2001).

- (11) *ḥasan ftara: ?ifti:₁ bas ma: b-a-šraf fu:ᵢ*
 Hasan bought.3ms something.ms but NEG Asp-1s-know what
 [~~*ḥasan ftara:*~~]
 Hasan bought.3ms
 ‘Hasan bought something, but I do not know what [~~Hasan bought~~].’

Such examples are also compatible with a *wh*-clefting analysis, as sketched in (12). As discussed above, *wh*-clefting is possible with bare *wh*-words in argument position, such as *fu*: ‘what’ and *mi:n* ‘who’. A *wh*-clefting analysis requires us to assume that the pronominal copula, *huwwe*, undergoes deletion as part of the target.

- (12) *ḥasan ftara: ?ifti:₁ bas ma: b-a-šraf fu:ᵢ* [~~*huwwe*~~
 Hasan bought.3ms something.ms but NEG Asp-1s-know what 3ms.COP
~~*illi ḥasan ftara:-a*~~]
 that Hasan bought.3ms-it
 ‘Hasan bought something, but I do not know what it is [~~that Hasan bought~~].’

Since *wh*-expressions of the sort seen in these examples can undergo *wh*-fronting and also occur in *wh*-clefts, such examples do not allow us to determine which derivation is the appropriate one (or whether both derivations are available). In some cases, the copular pronoun *huwwe* surfaces as part of the remnant along with the *wh*-word, as in (13). For such examples, *wh*-clefting (i.e. pseudo-sluicing) is the only possible underlying derivation.

- (13) *ḥasan ftara: ?ifti:₁ bas ma: b-a-šraf fu:ᵢ huwwe*
 Hasan bought.3ms something.ms but NEG Asp-1s-know what 3ms.COP
 [~~*illi ḥasan ftara:-a*~~]
 that Hasan bought.3ms-it
 ‘Hasan bought something, but I do not know what it is [~~that Hasan bought~~].’

This means there are two sources for the IP deletion. Also, Rodrigues et al. (2009) propose that there are two sources for IP deletion in Brazilian Portuguese (BP) and Spanish: sluicing and pseudosluicing (clefting) which I will discuss shortly.

In order to determine whether one or the other type of derivation can be ruled out, we must test examples with *wh*-expressions that fail to occur independently in the relevant underlying construction. Leung (2014a) argues that the underlying source for EA examples like (14) must be *wh*-fronting, since *wh*-NPs like *ʔajj no:ʃ* ‘which kind’ can be used in *wh*-fronting but not in *wh*-clefting in EA. This is confirmed by the ungrammaticality of **ʔajj no:ʃ huwwe* ‘which kind COP’ in EA in (14b).

- (14) a. *ʔaj kitab ʃtər-et ʔms?*
 which book bought-2SM yesterday
 ‘Which book did you buy yesterday?’
 b. *John ʔəʃrəb xamər bəs maa ʃərʃ [ʔajj nooʃ (*hu)]*
 John drink alcohol but not 1.know which kind 3SM
 ‘John drinks alcohol, but I don’t know which kind.’

(Leung, 2014a, p. 334–5)

However, JA differs from EA in this respect: *ʔajj*-NP ‘which-NP’ can occur in both *wh*-fronting and *wh*-clefting, as in (15).

- (15) a. *ħasan ʃtara: sajj:ra, bas ma: b-a-ʃraf ʔajj no:ʃ*
 Hasan bought.3ms car, but NEG Asp-1s-know which kind
 ‘Hasan bought a car, but I do not know which brand.’
 b. *ħasan ʃtara: sajj:ra bas ma: b-a-ʃraf ʔajj no:ʃ hijje*
 Hasan bought.3ms car, but NEG Asp-1s-know which kind 3fs.COP
 ‘Hasan bought a car, but I do not know which brand.’

Unlike in EA, the fronted *which*-NP *ʔajj no:ʃ* ‘which-brand’ in JA may be the result of either sluicing or pseudo-sluicing. Such examples leave us with the same analytical indeterminacy as (11) and (12).

However, there are JA cases that allow only *wh*-fronting or *wh*-clefting. Since *wh*-clefting is impossible with *wh*-adjuncts and *wh*-PPs where the preposition is pied-piped, as shown in (9), the only possible underlying derivation of the sluice is *wh*-fronting, not *wh*-clefting. Thus it must be analyzed as ordinary sluicing, as illustrated in (16) and (17).

- (16) *Wh*-adjunct
ʃumar itʃʔasal la-sabab bas ma: b-a-ʃraf le:ʃ [ʃumar
 Omar called.3ms for-reason but NEG Asp-1s-know why Omar
*itʃʔasal / *ħuwwe illi itʃʔasal]*
 called.3ms / 3ms.COP that called.3ms
 ‘Omar called for a reason, but I do not know why [Omar called/ *COP that he called.]’

- (17)
- Wh*
- PP with a pied-piped preposition

ʕumar ʔaka: maʕ ʔada:, bas ma: b-a-ʕraf maʕ mi:n [*ʕumar*
Omar talked.3ms with someone but NEG Asp-1s-know with who Omar
*ʔaka:/ *huwwe illi ʔaka:]*

talked.3ms/ 3ms.COP that talked.3ms

‘Omar talked to someone, but I do not know with who [Omar talked / *it is that he talked].’

JA sluicing behaves differently when the preposition is stranded. Since regular questions with a stranded preposition allow either *wh*-fronting or *wh*-clefting (18), one would expect to see some discrepancy in examples of sluicing.

- (18) *ʔajj za:mʕa daras-t liɲwistiks fi-ha*
 which university studied-2ms. linguistics in-it.fis
 ‘which university is it that you studied linguistics at?’

Merchant (2001) proposed the Preposition Stranding Generalization (PSG), which captures the parallelism between sluicing and *wh*-questions. The PSG states that if a language allows preposition stranding (P-stranding) in sluicing constructions, then it must allow P-stranding under regular *wh*-movement, as illustrated in (19).

- (19) Preposition Stranding Generalization

A Language *L* will allow preposition stranding under sluicing iff *L*
 allows preposition stranding under regular *wh*-movement.

(Merchant, 2001, p. 92)

Merchant (2001) surveyed twenty languages that confirm the PSG, including English (20), which allows P-stranding. English allows a preposition to strand in regular questions, and thus it allows the preposition to strand in the sluice site, as shown in (20). This is not the case in French, as in (21), nor in other languages that do not allow P-stranding.

- (20) English

- a. Who did Peter talk to?
- b. To whom did Peter talk?
- c. Peter talked to someone, but I do not know who [~~Peter talked to~~]

- (21) French

- a. **Qui est-ce qu' elle l'a offert à?*
 who Q she it-has offered to
 ‘whom has she offered it to?’
- b. *À qui l'a-t-elle offert?*
 to whom it-has-she offered
 ‘To whom has she offered it?’

- c. Anne l'a offert à quelqu'un mais je ne sais pas *(à) qui
 Anne it-has offered to someone but I NEG know NEG to whom
 'Anne has offered it to someone, but I don't know (to) whom.'

(Merchant, 2001, p. 98)

In other words, the possibility of P-stranding under *wh*-movement predicts the possibility of a preposition's remaining in-situ in sluicing when its *wh*-complement moves out of the target IP. On the other hand, Leung (2014b) suggested treating P-stranding as a PF-condition for Emirati Arabic.

The availability of pseudo-sluicing, with *wh*-clefting as the underlying source, complicates the picture. Some languages, like Mandarin Chinese (Wang, 2006) and Serbo-Croatian (Stjepanović, 2008), seem to violate the PSG on the surface, but utilize strategies that in fact salvage the PSG. These strategies include (i) resumption and (ii) preposition loss at PF. For instance, although Mandarin Chinese seems to violate the PSG in examples like (22) where Mandarin Chinese show P-omission pattern, Wang (2006) proposes that such examples involve a resumptive pronoun underlyingly. With a resumptive pronoun as its complement, the "stranded" preposition is no longer really stranded. On this view, Wang (2006) proposes that Mandarin Chinese confirms the PSG because apparently stranded prepositions in fact always involve a resumptive pronoun, following *wh*-movement in *wh*-questions and sluicing, as shown (23a, b).

- (22) a. *(shi) [na-ge ren]_i Lisi gen *t*_i zai shuohua?
 FOC.COP which-CL person Lisi with PROG talk
 'which one is Lisi talking with?'
 b. Lisi gen mou-ge ren quwan dan
 wo bu zhidao shi gen shei.
 Lisi with certain-CL person go-play but
 I NEG know FOC/COP with who
 'Lisi has a trip with a certain person, but I do not know who.'
- (23) a. [na-ge ren]_i Lisi hen zihuan ta-*t*_i?
 which-CL person Lisi very like him
 'which person does Lisi like (him) very much?'
 b. Keshi wo bu zahidao na-ge ren_i <TP Lisi gen ta-*t*_i qu
 with him go see movies but I NEG know which-CL
 kan dianying>
 person Lisi
 'but I don't know which person (did) Lisi go to the movies with him.'
- (Wang, 2006, pp. 248–9)

Stjepanović (2008), on the other hand, proposes preposition loss at PF as a strategy to rescue PSG violation in Serbo-Croatian.³ On the contrary, Sato (2011) argues that Indonesian sluicing violates the PSG. He examined Indonesian P-stranding under sluicing and adopted “repair by ellipsis” analysis (Ross, 1969) to explain why Indonesian⁴ allows P-stranding under sluicing although it is a non P-stranding language. He supported his claim by showing that same repair effect is available in pseudogapping that involves P-stranding in Indonesian. The details of these two analyses are not relevant to the current study, so I will not delve into them here.

Like BP, examined first by Almeida & Yoshida (2007),⁵ JA is a non P-stranding language, as shown in (24a). In other words, a preposition cannot be stranded in regular *wh*-questions in JA. All else being equal, the PSG leads us to expect that P-stranding should be unavailable under sluicing. Yet P-stranding under sluicing is allowed, as in (24b).

- (24) a. **mi:n ĥaka: ĥumar maʕ*⁶
 who talked.3ms Omar with
 ‘who did Omar talk with?’
 b. *ĥumar ĥaka: maʕ ĥada:, bas ma: b-a-ʕraf mi:n*
 Omar talked.3ms with someone but NEG Asp-1s-know who
 ‘Omar talked to someone, but I do not know who.’

Examining the data in (24), it is tempting to argue that JA is another language that violates the PSG. However, we cannot conclude that JA falsifies the PSG simply by looking at the data in (24). Taking into consideration (i) the underlying source of

3. Serbo-Croatian does not allow P-stranding, as shown in (i) and (ii).

- (i) *Sa kim je Ana govorila?*
 with whom.INST is Ana spoken
 ‘Who did Ana speak with?’ (Stjepanović, 2008, p. 180)
 (ii) **Kim je govorila Ana sa?*
 whom.INST is spoken Ana with

4. Like Serbo-Croatian, Indonesian does not allow – stranding, as shown in (i) and (ii).

- (i) *Dengan siapa kamu berdansa?*
 with who you dance
 ‘With whom did you dance?’
 (ii) **Siapa yang kamu berdansa dengan?*
 Who that you dance with
 ‘whom did you dance with?’ (Sato, 2011, p. 341)

5. Almeida & Yoshida (2007) proposed that BP challenges the PSG and BP exhibits the same features as English sluicing, not pseudosluicing.

6. The grammatical question in JA is, “*maʕ mi:n ĥaka ĥumar?*” by pied-piping the preposition with the question word *mi:n* ‘who’.

the sluice, (ii) the behavior of prepositions in JA questions, and (iii) the properties of JA sluicing, we may examine the underlying derivation of the sluice in (24). In cases where a preposition is pied-piped, as in examples like (17), the only possible underlying derivation of the sluice is *wh*-fronting; thus, these are ordinary sluicing cases. Conversely, when the preposition is not pied-piped, as in (24), the preposition is left in the sluice site and the only possible underlying derivation is *wh*-clefting,⁷ as in (25), in which *-o* ‘-him’ is a resumptive pronoun linked to the *wh*-word *mi:n* ‘who’. Rodrigues et al. (2009) conclude that the IP deletion that involves P-stranding is a deletion of a clefted IP out of which *wh*-word has moved. Adopting their argument, cases of P-stranding are only grammatical under sluicing from a source different from regular interrogative.⁸ It is a source that consists of a cleft structure whose IP is deleted (Rodrigues et al., 2009; van Craenenbroeck, 2010). This means that when clefting is not possible, P-stranding becomes implausible.

- (25) *ʕumar ʕaka: maʕ ʕada:, bas ma: b-a-ʕraf mi:n_i [ʕuwwe*
 Omar talked.3ms with someone but NEG Asp-1s-know who [3ms.COP
illi ʕaka: maʕ-o_j]
 that talked.3ms with-him]
 ‘Omar talked to someone, but I do not know who.’

The moved *wh*-word must be linked to a position or an element in the elided material inside the sluice site, which is the resumptive pronoun *-o* ‘-him’ in (25). Consequently, I argue that the data in (25) is consistent with a *wh*-clefting (i.e. pseudo-sluicing) analysis of examples like (24). Resumption is a strategy that salvages the PSG in JA, as argued for Mandarin by Wang (2006). The resumptive pronoun will always appear in tandem with the relative pronoun *illi* ‘that’ and the copular pronoun *ʕuwwe* ‘him’ in the sluice site.

Evidence for the presence of the resumptive pronoun along with the preposition under sluicing comes from the Class II resumption strategy (Aoun et al., 2010), a variation of the resumption strategy shown in (26). In this strategy, the fronted *wh*-constituent is related to a resumptive pronoun in its base position, and the *wh*-word immediately precedes the complementizer *illi* ‘that’.⁹ This question-formation strategy is found in many Arabic dialects, including JA.

7. *Wh*-fronting is not available under the intended meaning.

8. Rodrigues et al. (2009), a response to Almeida and Yoshida (2007), proposes that BP does not violate the PSG.

9. Example (26) is adapted from Aoun et al. (2010), where the clause-initial *wh*-constituent immediately precedes the definite relative-clause complementizer (*ya*)*lli* ‘that’, as in (i). This corresponds to *illi* ‘that’ in the examples shown from JA.

(26) Class II Resumptive strategy

fu:_i illi ftare:t-o_i mba:reħ?

What that bought-2ms-it yesterday

‘What was it that you bought yesterday?’

(Aoun et al., 2010)

Other constructions in which a resumptive pronoun is active is found in relative clauses (27) and embedded questions (28).

(27) *ha:da d-dakto:r illi ħak-e:t maʕ-*(o) b-l-za:mħa*

this the-professor who talked-2ms with-him at-the-university

‘This is the doctor that you talked with at the university.’

(28) *b-t-iʕraf mi:n (huwwe) illi ħake:-t maʕ-*(o) b-l-za:mħa*

Asp.2s.know who 3ms.COP that talked-2ms with-him at-the-university

‘Do you know who was it that you talked with at the university?’

As shown in this section, sluicing and pseudo-sluicing both exist in JA, yet in some cases only a *wh*-fronting, ordinary sluicing derivation is available, such as with *wh*-adjuncts and *wh*-PPs (with a pied-piped preposition). On the other hand, *wh*-arguments such as *fu:* ‘what’ and *mi:n* ‘who’, which allow both *wh*-fronting and *wh*-clefting, pose a challenge, because it is not necessarily clear whether particular examples containing them involve sluicing or pseudo-sluicing. In the next section, I discuss sprouting in JA.

4. Sprouting

As a subtype of sluicing, sprouting is expected to behave in a similar manner to sluicing. Recall that in sprouting, the *wh*-remnant does not have an overt correlate in the antecedent clause. I will begin by discussing the licensing conditions for JA sprouting, and then discuss discrepancies between sluicing and sprouting in JA. Sprouting is exemplified in (4), repeated in (29) for convenience.

(29) *ħasan ftara: bas ma: b-a-ʕraf fu:*

Hasan bought.3ms but NEG Asp-1s-know what

‘Hasan bought, but I do not know what.’

(i) Lebanese Arabic

miin (ya)lli s'əft-o b-l-maTʕam?

who that saw.2ms-him in-the-restaurant ‘

Who is it that you saw in the restaurant?’

(Aoun et al., 2010, p. 128)

In this case, a direct object (*fu*: ‘what’) is sprouted. As in the sluicing cases discussed earlier, it seems that the underlying derivation could be either *wh*-fronting or *wh*-clefting, as shown in the non-elliptical examples in (30).

- (30) a. *ḥasan ftara: bas ma: b-a-ṣraf fu: [ḥasan ftara:]*
 Hasan bought.3ms but NEG Asp-1s-know what Hasan bought.3ms
 ‘Hasan bought, but I do not know what Hasan bought.’
 b. *ḥasan ftara: bas ma: b-a-ṣraf fu: huwwe illi*
 Hasan bought.3ms but NEG Asp-1s-know what 3ms.COP that
 (*ḥasan*) *ftara:-a*
 Hasan bought.3ms-it
 ‘Hasan bought, but I do not know what it is that Hasan bought.’

In addition, sprouting is licensed when an entire PP is pied-piped, as in (31).

- (31) *ḥasan ḥaka: bas ma: b-a-ṣraf maṣ mi:n*
 Hasan talked.3ms but NEG Asp-1s-know with who
 ‘Hasan talked, but I do not know with who.’

However, one difference between sluicing and sprouting in JA is that indirect objects can be sluicing remnants but cannot be sprouted, as in (32).

- (32) a. **ḥasan ʔaṣṭʔa: hadijje, bas ma: b-a-ṣraf mi:n*
 Hasan gave.3ms gift but NEG Asp-1s-know who
 ‘Hasan gave a gift, but I do not know who.’
 b. *ḥasan ʔaṣṭʔa: ḥada: hadijje, bas ma: b-a-ṣraf mi:n*
 Hasan gave.3ms someone gift but NEG Asp-1s-know who
 ‘Hasan gave someone a gift, but I do not know who.’

Another difference is that JA cannot sprout a subject, whereas sluicing is grammatical with an overt indefinite correspondent, as in (33).

- (33) a. **l-waḍʔiṣ b-j-tʔʔallab l-ḥudu:r la-l-mi:da:n, bas ma:*
 the-situation Asp-3s-require the-heading to-the-square, but not
ḥaddad mi:n.
 specified.3ms who
 ‘the situation requires to head to the square, but it did not specify who.’
 b. *l-waḍʔiṣ b-j-tʔʔallab ḥudu:r ḥada la-l-mi:da:n, bas ma:*
 the-situation Asp-3s-require heading someone to-the-square, but NEG
ḥaddad mi:n.
 specified.3ms who
 ‘the situation requires someone to head to the square, but it did not specify who.’

The example in (33a) shows that, like English, JA does not exhibit subject sprouting. The subject in JA cannot be sprouted because the *wh*-word has no corresponding overt correlate in the antecedent clause.

A third difference is that JA sprouting does not allow the object of a preposition to sprout (34a), a restriction known as Chung's Generalization (Chung, 2005). Sluicing does permit this (as also seen earlier), as shown in (34b).

- (34) a. **ḥasan ra: ḥ ʕa-l-zim bas ma: b-a-ʕraf mi:n*
 Hasan went.3ms to-the-gym but NEG Asp-1p-know who
 'Hasan went to the gym, but I do not know who.'
 (intended interpretation: with whom)
- b. *ḥasan ra: ḥ ʕ-al-zim maʕ ḥada: bas ma:*
 Hasan went.3ms to-the-gym with someone but NEG
ba-ʕraf mi:n
 Asp-1psknow who
 'Hasan went to the gym with someone, but I do not know who.'

Chung (2005) observes that even in P-stranding languages like English, the object of a stranded preposition cannot be sprouted, as shown in (35).

- (35) a. *They are jealous, but it is not clear who.
 b. They are jealous, but it is not clear of who/who of.
 c. They are jealous of someone, but it is not clear who.
 (Chung et al., 2011, p. 10)

This raises the issue of non P-stranding languages like JA, in which the preposition remains in-situ along with the resumptive pronoun in the underlying structure. For *wh*-PPs in JA, sluicing permits the two options seen earlier: the preposition either pied-pipes as in (36a) or strands as in (36b).¹⁰

10. Syntactically, this can be explained by optional [wh] feature percolation (Chomsky, 1998), the ability of the [+wh] feature of the interrogative element DP to percolate onto its dominating PP, which requires the preposition to pied-pipe along with the *wh*-word to the specifier of the matrix CP. On the other hand, when the preposition strands or remains in-situ as in English, the [+wh] feature does not percolate, and thus only the DP has the [+wh] feature, which forces only the *wh*-word to move to the specifier of the matrix CP. In a non P-stranding language, e.g. Arabic, [+wh] percolates to PP, which causes the entire PP to pied-pipe. In JA sluicing, the preposition can remain in-situ with the presence of a resumptive pronominal item, as I have shown above as a strategy to salvage the PSG in JA.

- (36) a. *ʕumar ʔaka: maʕ ʔada:, bas ma: b-a-ʕraf maʕ mi:n_i*
 Omar talked.3ms with someone, but NEG Asp-1s-know with who
 [*ʕumar ʔaka-t_i*]
 Omar talked.3ms
 ‘Omar talked with someone, but I do not know with who [Omar talked]
- b. *ʕumar ʔaka: maʕ ʔada:, bas ma: b-a-ʕraf mi:n_i*
 Omar talked.3ms with someone, but NEG Asp-1s-know who
 [*huwwe illi ʕumar ʔaka: (maʕ-o)-t_i*]
 3ms.COP that Omar talked.3ms with-him
 ‘Omar talked with someone, but I do not know who [is it that Omar talked to him]

Therefore, a question that arises here is whether sprouting in JA allows P-stranding in violation of Chung’s Generalization. As (37) shows, it does not.

- (37) **ʕumar ʔaka:, bas ma: b-a-ʕraf mi:n_i*
 Omar talked.3ms but NEG Asp-1s-know who
 ‘Omar talked, but I do not know who.’

Examining the underlying derivation, the relevant underlying structures are ungrammatical both with *wh*-fronting and with *wh*-clefting, as shown in the non-elliptical examples in (38).

- (38) a. **ʕumar ʔaka:, bas ma: b-a-ʕraf mi:n* [*ʕumar*
 Omar talked.3ms but NEG Asp-1s-know who Omar
ʔaka: maʕ-o]
 talked.3ms with-him
 ‘Omar talked, but I do not know who Omar talked with.’
- b. **ʕumar ʔaka:, bas ma: b-a-ʕraf mi:n_i* [*huwwe illi ʕumar*
 Omar talked.3ms but NEG Asp-1s-know who 3ms.COP that Omar
ʔaka: maʕ-o]
 talked.3ms with-him
 ‘Omar talked, but I do not know who it is that Omar talked with.’

It is presumably allowed to have either *wh*-fronting or *wh*-clefting in the underlying derivation independently. Sprouting a PP is permitted, but only when the entire PP is pied-piped out of the ellipsis site. Sprouting is not allowed where only the object of the preposition is fronted and the preposition itself remains in the ellipsis site (followed by a resumptive pronoun or not). Chung’s Generalization rules out both the *wh*-fronting and *wh*-clefting options in such cases.

As such, the preposition has to pied-pipe in order to create a grammatical sprouting construction, as in (39). The only available underlying derivation is *wh*-fronting of the *wh*-PP, a derivation corresponding to what we find in ordinary sluicing.

- (39) *ʕumar ʔaka:, bas ma: b-a-ʕraf maʕ mi:n* [*ʕumar ʔaka:ʔ*]
 Omar talked.3ms but NEG Asp-1s-know with who Omar talked.3ms
 ‘Omar talked, but I do not know with whom Omar talked.’

Returning to examples with *wh*-arguments such as *ʕu*: ‘what’ and *mi:n* ‘who’ in sprouting as in (29), recall that *ʕu*: ‘what’ and *mi:n* ‘who’ allow both *wh*-fronting and *wh*-clefting in regular constituent questions and in sluicing underlyingly, as shown above in (11) and (12). Sprouting in JA allows *wh*-fronting and also *wh*-clefting with the copular pronoun *uwe* underlyingly, as in (30).

However, the size of the remnant and the target differ under sprouting. Although Example (30) is grammatical, it is not possible to elide the complement of the copular pronoun and leave *uwe* as a remnant along with the *wh*-word. This shows that sprouting in JA (with implicit antecedent correlate) allows only *wh*-fronting.

Leung (2014a) shows that the EA examples in (40) are grammatical with the use of either *wh*-construction in the second clause when there is no overt correlate in the antecedent clause, and the same is true in JA. *Wh*-fronting is demonstrated in (40a); and *wh*-clefting is demonstrated in (40b).

- (40) a. *John jsuug, bas ma ʕraf [ʕuu jsuug]*
 John 3_{SM}.drive but not I.know what 3_{SM}.drive
 ‘John drives, but I don’t know what he drives.’
 b. *John jsuug, bas ma ʕraf [ʕuu (hu) Elli jsuug-ah]*
 John 3_{SM}.drive but not I.know what 3_{SM} that 3_{SM}.drive-3_{SM}.
 ‘John drives, but I don’t know what it is that he drives (it).’

(Leung, 2014a, p. 337)

Leung (2014a) also shows that the use of the *wh*-sluice *ʕu*: ‘what’ is grammatical in EA constructions without an overt correlate, while the *wh*-pseudo-sluicing indicated by *ʕuu hu* ‘what it is’ is not allowed, as indicated in (41).

- (41) *John jsuug>___, bas ma ʕraf [ʕuu (*hu)]*
 John 3_{SM}.drive but not I.know what 3_{SM}
 ‘John drives, but I don’t know what it is that he drives (it).’

(Leung, 2014a, p. 338)

Leung concludes that *wh*-pseudo-sluicing (*wh*-clefting) in EA is not allowed in sprouting, or if it is a *wh*-adjunct or a *wh*-phrase (*wh*-PP or *which*-NP). Only *wh*-fronting is a plausible sluicing source, and Leung (2014b) has solved this puzzle by proposing that the numeration of the antecedent clause has an impact on the sluice source. However, I will argue that the equivalent example to (41) in JA indicates that *wh*-clefting may not be analyzed as the underlying source when a PP is sprouted, which explains the impossibility of the copular pronoun *uwe* to surface as a remnant.

The evidence comes from the fact that the preposition has to pied-pipe in order to create a grammatical sprouting construction where a PP is sprouted, shown in (38) and (39). As explained previously, in JA and EA, *wh*-clefting is not allowed with a *wh*-PP where the preposition is fronted along with the *wh*-word. This prevents us from having a *wh*-cleft target in (39), thereby excluding *wh*-clefting as a possible underlying source for sprouting. The only available underlying derivation is *wh*-fronting of the *wh*-PP, which means that the copular pronoun *uwe* is not underlyingly present in the target in the first place.

A case where only *wh*-cleft is allowed in JA, solves this puzzle and supports Leung's proposal. In EA, there is one case that allows only *wh*-clefts (Leung, 2014b), that does not exist in JA (42), that is exhaustivity.

- (42) *kti:r na:s ?izu: ʕ-l-ħafle, bas ma: baʕraaf [* min: l-kul]*
 a lot people came.3ms to-the-party, but NEG Asp-1s-know [who the-all]
 'A lot of people came to the party, but I do not know them all.'

The grammatical equivalent example in JA includes only *l-kul* 'the-all' to the exclusion of the *wh*-word *mi:n* 'who'. This means that the result is not a possible sluicing. Leung's analysis cannot be extended and thus a clearer data that allows only *wh*-clefting is required, which remains a challenge in this paper.

5. Conclusion

In summary, distinguishing sluicing from pseudo-sluicing depends on the type of *wh*-word present and which *wh*-construction strategy each *wh*-word allows. Sluicing and pseudo-sluicing both exist in JA, yet in some cases only a *wh*-fronting, ordinary sluicing derivation is available, such as with *wh*-adjuncts and *wh*-PPs (with a pied-piped preposition). On the other hand, *wh*-arguments such as *ʕu:* 'what' and *mi:n* 'who', which allow both *wh*-fronting and *wh*-clefting, pose a challenge, because it is not necessarily clear whether particular examples containing them involve sluicing or pseudo-sluicing.

Apparent violations of the PSG in JA can be analyzed as cases of resumption, thus salvaging the PSG. Resumption is very productive in JA, occurring also in relative clauses and embedded questions.

Sprouting in JA exhibits behavior parallel to sluicing as regards the types of *wh*-expression that participate. However, JA sprouting is subject to Chung's Generalization, which bars sprouting of objects of a preposition; and subjects and indirect objects cannot be sprouted, in contrast to what we find in JA sluicing.

For the copular pronoun *huuwe* to be acceptable as part of the remnant, the overt correlate must be present as in (13), or there must be an implicit correlate

along with *huwwe* in (30). In sprouting cases, *huwwe* never appears as part of the remnant. For cases with *wh*-PPs, this can be attributed to the unavailability of the *wh*-clefting structure as an underlying source for the ellipsis. For sprouting cases with bare *wh*-word remnants, Leung (2014b) proposes that it is the numeration of the antecedent clause that plays a role in the numeration of the sluice. This can also be supported with cases where only *wh*-cleft is allowed, such as the acceptability of exhaustivity in sprouting examples in EA, that JA does not accept. Finally, *Wh*-adjuncts and *wh*-PPs cannot occur in *wh*-clefts in JA sprouting. This implies that a pseudo-sluicing derivation is unavailable in such cases, a fact that probably explains the impossibility of *huwwe* following the *wh*-PP.

References

- Algryani, A. (2010). "Preposition Stranding in Libyan Arabic Sluicing." *Newcastle working papers in Linguistics*, 16, 1–22.
- Algryani, A. (2011a). "VP ellipsis in Libyan Arabic." *Newcastle working papers in Linguistics*, 17, 1–22.
- Algryani, A. (2011b). "Sluicing in Libyan Arabic." *Newcastle working papers in Linguistics*, 18, 1–23.
- Algryani, A. (2013). On the Syntax of Stripping in Libyan Arabic. *International Journal of Linguistics*, 5(5), 156–173. <https://doi.org/10.5296/ijl.v5i5.4456>
- Almeida, Diogo A. de A. & Yoshida, M. (2007). A problem for the Preposition Stranding Generalization. *Linguistics Inquiry*: 28(2), 349–362. <https://doi.org/10.1162/ling.2007.38.2.349>
- Aoun, J., Benmamoun, E. & Choueiri, L. (2010). *The syntax of Arabic*. New York: Cambridge University Press.
- Chomsky, N. (1998). *Minimalist Inquiries: the Framework*. Ms., MIT.
- Chung, S. (2005). 'Sluicing and the Lexicon: The point of no return. In R. T. Cover, & Y. Kim (Eds.), *BLS 31: General Session and Parasession on Prosodic Variation and Change* (pp. 73–91). Berkeley, CA: Berkeley Linguistics.
- Chung, S., Ladusaw, W., & McCloskey, J. (2011). Sluicing (:): Between structure and inference. In R. Gutiérrez-Bravo, L. Mikkelsen, & E. Potsdam (Eds.), *Representing Language: Essays in Honor of Justin Aissen* (pp. 31–50). Santa Cruz, CA: Linguistics Research Center.
- Craenenbroeck, J. V. (2010). *The syntax of ellipsis: Evidence from Dutch dialects*. Oxford: Oxford University Press.
- Culicover, P., & Jackendoff, R. (2005). *Simpler Syntax*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199271092.001.0001>
- Eid, M. (1983). The Copula Function of Pronouns. *Lingua* 59(2–3), 197–207. [https://doi.org/10.1016/0024-3841\(83\)90063-3](https://doi.org/10.1016/0024-3841(83)90063-3)
- Ginzburg, J., & Sag, I. (2000). *Interrogative investigations: The form, meaning, and use of English interrogatives*. Stanford, California CSLI Publications.
- Leung, T. (2014a). The preposition stranding generalization and conditions on sluicing: Evidence from Emirati Arabic. *Linguistic Inquiry*, 45(2), 332–340. https://doi.org/10.1162/LING_a_00158

- Leung, T. (2014b). Modes of interrogatives entail modes of sluicing-evidence from Emirati Arabic. In S. Farawaneh & H. Ouali (Eds.), *Perspectives on Arabic Linguistics XXIV–XXV* (pp. 207–228). Amsterdam: John Benjamins.
- Merchant, J. (1998). ‘Pseudosluicing’: Elliptical clefts in Japanese and English. In A. Alexiadou, et al. (Eds.), *ZAS working paper in linguistics* (volume 10) (pp. 88–112). Berlin: Zentrum fur Allgemeine Sprachwissenschaft.
- Merchant, J. (2001). *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Rodrigues, C., & Nevins, A., & Vicente, L. (2009). Cleaving the interactions between sluicing and P-stranding. In D. Torck & W. Leo Wetzels (Eds.), *Roman Languages and Linguistic Theory* (pp. 175–198). Amsterdam: John Benjamins.
- Ross, J. (1969). Guess who? In R. Binnick, A. Davidson, G. Green, and J. Morgan (Eds.), *Proceedings of the 5th Chicago Linguistic Society (CLS5)* (pp. 252–286). Chicago, IL.
- Sato, Y. (2011). P-stranding under sluicing and repair by ellipsis: Why is Indonesian (not) special? *Journal of East Asian Linguistics*, 20(4), 339–382. <https://doi.org/10.1007/s10831-011-9082-3>
- Stjepanović, S. (2008). P-stranding under sluicing in a non P-stranding language? *Linguistic Inquiry*, 39(1), 179–190. <https://doi.org/10.1162/ling.2008.39.1.179>
- Shlonsky, U. (1997). *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford: Oxford University Press.
- Toosarvandani, M. (2008). Wh-movement and the syntax of sluicing. *Journal of Linguistics*, 44(3), 677–722. <https://doi.org/10.1017/S0022226708005367>
- Wahba, W. (1984). Wh-constructions in Egyptian Arabic. Doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.
- Wang, C. (2006). Sluicing and resumption. In E. Elfner, & M. Walkow (Eds.), *Proceedings of the 37th Annual meeting of the North-East Linguistic Society* (pp. 239–252). Amherst: GLSA publications.

PART III

Clinical linguistics

Clinical linguistic research in the study of Arabic diglossia

Reem Khamis-Dakwar
Adelphi University

Clinical linguistics was first defined as the “application of the theories, methods, and findings of linguistics (including phonetics) to the study of those clinical situations where disorders of language are diagnosed and treated” (Crystal, 1981: 1). Based on a Question and Answer presentation at the International Clinical Phonetics and Linguistics Association, Crystal called for more focused cross-linguistic comparative studies in clinical linguistics (Crystal, 2013). Although clinical applications of linguistics are emphasized in such studies, this kind of research may also add to theoretical inquiries in linguistics and the field of Communication Sciences and Disorders (CSD). This paper highlights the potential contributions of Arabic clinical linguistic studies to CSD, through a description of two studies on diglossia and its interactions with assessment and intervention services provided by speech-language pathologists (SLPs) in different CSD populations. The paper provides a brief overview of the SLP profession and its growth in the Arabic-speaking world, followed by descriptions of two preliminary clinical linguistic studies of Arabic diglossia: one focused on aphasia recovery in diglossic speech communities, and the other on diglossic codeswitching in Arabic-speaking individuals with Autism Spectrum Disorders. These studies share a broad theoretical framework, and both exemplify the relevance of linguistic theory and methods for investigations of communication disorders in diglossia. Of particular interest is the promise of systematic cross-dialectal investigations of communication disorders in line with Crystal’s (2013) call for expansion of deliberately designed comparative cross-linguistic studies in clinical linguistics.

Keywords: clinical linguistics, Arabic, speech-language pathology, communication sciences and disorders, diglossia, aphasia, autism-spectrum disorder

1. Communication Sciences and Disorders in the Arabic-speaking world

Communication disorders is a broad term referring to any impairment in the ability to receive, send, process, or comprehend concepts through verbal, nonverbal, and graphic symbol systems (ASHA, 1993). Speech-Language-Pathologists (SLPs), trained in the field of Communication Sciences and Disorders (CSD), are clinical professionals who diagnose and treat communication and swallowing disorders including speech, language, social communication, cognitive-communication, and feeding disorders in children and adults. They work in a range of settings including schools, hospitals, private clinics, home care, universities, rehabilitation centers, nursing homes, and research laboratories. In Arabic-speaking communities, there is an increased interest and awareness of speech-language pathology. However, this interest is limited by a shortage of trained and certified Arabic-speaking SLPs and a dearth of academic programs and resources.

1.1 Shortage of Arabic-speaking SLPs

There are varying levels of SLP shortages and expertise within the diverse Arabic-speaking countries due to differences in the numbers of training programs and the history of the profession in each country. There is no available comprehensive documentation of the shortage of Arabic-speaking SLPs in Arab countries, and much of what we know is based on partial information provided within introductory sections of clinical articles, graduation reports of academic programs, outdated reports of professional organizations, and anecdotes of professionals in the field. None of the Arab countries come close to the American national average of 51.1 certified speech-language pathologists for every 100,000 residents. For example, Kotby (2014) noted that there are 253 registered members of the Egyptian Society for phoniatricians and logopedists/SLPs serving a population of 95.6 million, which is less than 0.5 (0.26) for every 100,000 people. Similarly, only 42 SLPs are reported to work in Kuwait (counting both Kuwaiti and non-Kuwaiti clinicians; (Al-Khaledi, Lincoln, McCabe, Packman, & Alshati, 2009)) serving an estimated population of 4 million – that is, 1.05 speech pathologists for every 100,000 Kuwaitis. See Table 1 for available information on the reported number of speech pathologists per capita in the different Arab countries.

1.2 Paucity of academic programs and resources

Whereas some Arab countries like Morocco and Dubai have newly established speech-language-pathology programs, other Arab countries like Egypt and Saudi Arabia have had established programs since the early 1970s. An examination of these programs reveals that some offer general training in rehabilitation science (e.g. Dubai) while others offer specified speech-language pathology training (e.g. Saudi Arabia). Most noticeably, and in comparison to the United States in which a Master's degree is required for certification, most speech-language pathology programs in Arab countries require either a diploma or a Bachelor's degree for certification (as shown in Table 1). While in the U.S., there are over 300 Master's programs in speech-language pathology and 76 doctor of philosophy (PhD) speech-language pathology programs, there is documentation of only twenty-seven academic programs in Arab countries, mostly at the undergraduate level, and one doctoral program within a medical program in Egypt. The program in Egypt offers a Bachelor's degree in medicine and a Master's degree in phoniatics, with the option to pursue a doctoral degree in the profession (Mostafa & Ahmed, 2017).

This paucity of training programs has an even greater impact considering that these different programs serve diverse Arabic-speaking communities with distinct cultural practices and spoken dialects. These cultural and linguistic differences make it difficult to generalize resources and knowledge from one university or Arab country, like Jordan, to be used in another Arabic-speaking community like Morocco. Geographic proximity may ease generalization of knowledge and resources between Arabic-speaking communities, for example the use of Jordanian Arabic resources in Palestine, but this has not been fully realized in practice.

Table 1. Number of reported speech pathology training programs in different Arab countries

| Name of country | Number of reported academic training programs | Masters and/or doctoral level training reported | Number of speech pathologists reported | Source |
|-----------------|---|---|--|--|
| Morocco | 1 public university (B.A.) in Rabat, and 5 private diploma training programs across the country | No | NA | Anecdotal report by vice president of the SLP association in the north of Morocco (Zainab Lahboussi) |

(continued)

Table 1. (continued)

| Name of country | Number of reported academic training programs | Masters and/or doctoral level training reported | Number of speech pathologists reported | Source |
|-----------------------------|---|---|--|---|
| Saudi Arabia | 3 B.S. programs in audiology and speech language pathology <ul style="list-style-type: none"> – King Saud: Dual degree program in speech and hearing rehabilitation (founded 1967) – B.S. and M.A. programs in speech-language pathology within school of Education and Health sciences, Dar Al-Hekmah University (founded 2009) – Princess Nourah Bint Abdulrahman: separate programs in audiology and speech pathology (female students) | Yes | Total reported graduates from Dar Al-Hekmah to date: 85 | Alanazi (2017) |
| Kuwait | B.S. program in communication sciences and disorders, Kuwait University (first graduating class: 2008) | No | 42 Kuwaiti and non-Kuwaiti SLPs working in Kuwait | Abdalla & Al-Saddah (n.d.) |
| Qatar | B.S. program in rehabilitation treatment | No | NA | Anecdotal report of a speech pathologist in Qatar |
| United Arab Emirates | No speech language pathology programs | No | NA | Wu (2015) |
| Lebanon | 2 B.A. and 2. M.A. programs | Yes | NA | Ismail (2015) |
| Libya | 1 B.A. program | No | NA | Ismail (2015) |
| Egypt | <ul style="list-style-type: none"> – 1 B.A., Ain Shams University (established 1975) – 1 M.A. in medicine, phoniatics, Ain Shams University (established 1976) | Yes | 253 members reported at the Egyptian Society for Phoniatics and Logopedics 148 M.A., 73 doctoral degree | Kotby (2014) |

Table 1. (continued)

| Name of country | Number of reported academic training programs | Masters and/or doctoral level training reported | Number of speech pathologists reported | Source |
|----------------------|---|---|---|--|
| Israel/ Palestine | – 5 B.A. (Israel)* – 1 M.A. (Israel) – 1 B.A. (Palestine) | Yes | Only 21 of the 1,185 speech therapists in Israel were Palestinian Arabs in 2000 (Human Rights Watch, 2001). | Personal communication with Palestinian speech pathologists |
| Jordan | – 3 B.A. – 2 M.A. | Yes | NA | Personal communication with Dr. Laila Qanawati, from AlAhliya Amman University |
| Syria | 1 M.A. | Yes | NA | Ismail (2015) |
| Tunisia | 1 Higher diploma | No | NA | Ismail (2015) |

* Twenty percent of citizens of Israel are indigenous Palestinian who were able to stay in Israel after the 1948 war. Members of this minority have access to Israeli universities.

In summary, a shortage of academic programs in speech-language pathology in Arab countries and Arabic-speaking communities, coupled with the lack of graduate-level training for SLPs, impacts the capacity and quality of speech-language pathology services. Furthermore, this limits the development of the profession by restricting opportunities for specialized training, active professional organizations, scholarly clinical investigations, and development of clinical resources.

2. Clinical linguistic studies of Arabic diglossia

A number of distinctive features of Arabic render it a fruitful area of study for both language-specific and universal theoretical and clinical studies in linguistics. For example, the study of early language acquisition milestones across the different dialects can provide valuable insights into the interplay between universal principles of language acquisition and language-specific characteristics (for review, see Khamis-Dakwar & Froud, in press (a)). The diglossic nature of Arabic (Ferguson, 1959) also provides unique opportunities for study. The term diglossia refers to the

co-existence of two language varieties, one acquired naturally and used for daily communication while the other is perceived to be more prestigious by its community members, learned formally and used for formal communication, reading and writing. This sociolinguistic phenomenon is evident in several Western and Eastern speech communities around the world. There are critical differences between the learning mechanisms that underpin the two diglossic varieties in Arabic. One, the spoken dialect, is acquired naturalistically through incidental exposure, involving implicit learning mechanisms. The other, the Modern Standard Arabic (MSA), is learned formally in the course of literacy development, meaning that it is more reliant on explicit learning mechanisms. There are also inherent differences in the communicative contexts in which each variety is used. MSA is generally used in formal contexts and for reading and writing, while spoken dialects are used for daily communication. These critical differences interact with the course of language acquisition and processing in typical and atypical language development (for review, see Khamis-Dakwar & Froud, 2019 (b)). Bearing these differences in mind, and their applications for our understanding of language acquisition mechanisms, below I present on two clinical linguistics studies in Arabic diglossia.

In this section, I present two preliminary clinical linguistic studies focused on the clinical implications of Arabic diglossia, with the intention of exemplifying the power of expanded clinical linguistic inquiries in Arabic. I posit that these initial clinical linguistic investigations of diglossia in aphasia and in Autism Spectrum Disorders (ASD) illustrate the great promise and impact of such studies in Arabic for informing specific theoretical linguistic queries in the field as well as enhancing culturally and linguistically responsive clinical services for Arabic-speaking individuals with aphasia or ASD. From these examples, I hope to spark the interest of Arab linguists in proactively extending their work to clinical investigations, when appropriate, across the different dialects of Arabic. Such systematic expansions of clinical linguistic studies in Arabic have the potential to build an evidence-base for speech-language pathology practices, enhance the quality of speech-language pathology services provided for Arabic-speaking patients, develop awareness of the needs of individuals with communication impairments, and inform theoretical questions in the fields of linguistics and CSD.

2.1 Aphasia recovery and Arabic diglossia: Adaptation of the Bilingual Aphasia Test to Palestinian Arabic and Modern Standard Arabic

2.1.1 *Clinical and theoretical background*

Aphasia is a language impairment caused by brain damage such as stroke or traumatic brain injury. Aphasia impacts receptive and/or expressive language in spoken and/or written modalities. There are many different kinds of aphasia, typically

broadly subdivided into fluent and non-fluent presentations, depending on the primary impact on spoken output. Bilingual or multilingual speakers with aphasia may exhibit parallel or non-parallel recovery patterns in the languages spoken prior to the brain damage (Kohnert, 2013). Parallel recovery patterns refer to similar types and degrees of impairment across domains and modalities in both or all languages, relative to the degree of pre-morbid proficiency. In non-parallel patterns of recovery, one or more language may exhibit better recovery relative to the other languages spoken, again dependent on the speaker's premorbid linguistic profile (for review, see Kohnert, 2013).

The nature of non-parallel recovery was informed by Paradis' (1994, 2004) application of the declarative/procedural model of memory to differentiate between two types of linguistic knowledge: implicit and explicit. On this view, implicit linguistic knowledge is sustained and learned by procedural memory systems through incidental learning and automatic use. In contrast, explicit linguistic knowledge is dependent on declarative memory mechanisms, and is acquired through conscious learning and practice. These two memory systems are neurofunctionally and anatomically distinct. Processing implicit linguistic knowledge, such as knowledge of grammar, involves circuits extending to parts of the right cerebellum, left neostriatum, and the frontal basal ganglia as well as portions of the perisylvian cortex, whereas explicit linguistic knowledge involves the hippocampal system, mesial temporal lobes, parahippocampal gyri, and anterior cingulate cortex (Paradis, 2008). According to Paradis (1994, 2004), reported differences in the recovery of one language versus the other may depend on the inherent differences between these two types of linguistic knowledge and their distinct associations with different memory mechanisms. Accordingly, since aphasia affects implicit linguistic knowledge, individuals with aphasia may have greater access to explicit linguistic knowledge as a compensatory mechanism. Aphasia impairment impacts patients' access to all implicit linguistic knowledge, be it in the first language or a second language. Cases of severe damage may result in a total lack of access to implicit knowledge. In these cases, the individual with aphasia may rely on declarative knowledge as a compensatory mechanism, resulting in observations of relatively spared performance in a second, explicitly-learned, language.

Understanding the underlying nature of the various patterns of sparing and impairment requires effective evaluation of the multiple languages affected in bilingual or multilingual individuals with aphasia. This is a difficult task. Language testing in bilingual/multilingual and culturally diverse individuals with aphasia has most often involved separate assessments for each language. Often these assessments are translated versions of an original that was developed for English. Several limitations of this practice have been identified including cultural bias, psychometric deficiency, and lack of homogeneity of the bilingual/multilingual population to whom

the translated assessment is being applied (for review, see Kohnert, 2013). The *Bilingual Aphasia Test* (BAT) is a battery of subtests available in several languages and language pairs that enables valid comparison between patients' performance across languages. The BAT is a culturally non-biased language test for the assessment of aphasia in speakers of more than one language.

Development of the BAT requires the clinical application of linguistic theory in order to set criteria for equivalence across languages in the multiple adaptations of the test (Paradis, 1987). For example, several linguistic criteria were applied to the construction of the syntactic comprehension task in the BAT, including evaluation of both standard canonical sentences (such as *the girl read a book*, representing the unmarked SVO order for English) and non-canonical sentences in the target language (e.g., complex structures such as passives and relative clauses in English). Other relevant structural and syntactic-semantic properties are also evaluated (such as presence of animate pronouns that do not exist in English but are present in Arabic; the presence of analytic and synthetic possessives; masculine and feminine markers on adjectives, etc.). Similarly, contrastive analysis of morphological processes in different languages underlies the design of the semantic category task. In this task patients are asked to identify an antonym for a target word within the same grammatical category, resulting in the use of specific word forms (such as derived adverbs or adjectives). For example, in the English BAT, item 168 within this task, the patient is given the following prompt "now, for these, the choices are very similar, but only one is the opposite of the word you'll hear, ready?" and then is presented with the word *ugly* and the following three opposite options to choose as an answer: beautiful, beauty, beautifully) (see Paradis, 1987). These options can (and should) be modified to reflect language-specific constraints. For example, in the English version of the BAT, the answer choices include adverb, noun, and adjective; however since adverbs are usually produced as a preposition followed by the noun in dialectal Arabic, the Palestinian Arabic BAT offers past tense as a third alternative instead (Khamis-Dakwar, Ahmar, Farah, & Froud, 2018). Contrastive linguistic analysis is also foundational to the translation and grammaticality judgment tasks included in the BAT, which require patients to translate between languages and perform grammaticality judgments for sentences that would adhere to distinctive grammatical features of one language when compared to another. For example, for the BAT developed for the language pair MSA-Palestinian (Khamis-Dakwar, Ahmar, Farah, & Froud, 2018), the following contrastive features were identified: dual formation, interrogative sentence formation, and negation with one, two or three contrastive features (for a review of the adaptation of BAT to Palestinian Spoken Arabic and Modern Standard Arabic see Khamis-Dakwar et al., 2018).

2.1.2 *Potential clinical and theoretical contributions*

The adaptation of the BAT to the diglossic situation that exists between Palestinian and Modern Standard Arabic demonstrates the value of the clinical application of linguistic analysis to the study of communication disorders. To highlight the theoretical significance of such endeavors, I will describe the clinical and linguistic impact of this approach. This work began with the development of the BAT in the two language varieties spoken by Palestinian Arabic speakers, Palestinian Spoken Arabic and Modern Standard Arabic, to enable comprehensive clinical assessment of all language systems used by an Arabic-speaking individual with aphasia (Khamis-Dakwar, Ahmar, Farah, & Froud, 2018). This initiative was pioneering, since pre-existing Arabic language adaptations of the BAT were focused only on examining linguistic abilities in a specific Arabic dialect (such as Jordanian or Tunisian) and adaptations of language pairs focused on pairing one Arabic variety within a “traditional” conceptualization of linguistic system variations (e.g., Arabic-Armenian, Arabic-English, Arabic-French, Arabic-Somali, Arabic-Swahili),¹ rather than examining the relationships between spoken and Standard varieties of Arabic. All these adaptations are needed for comprehensive clinical assessment of Arabic-speaking individuals with aphasia, and linguistic knowledge and analysis is instrumental for the development and adaptation of these assessments.

However, by examining the recovery patterns of the two language varieties in Arabic diglossia, distinctively correlated with implicit versus explicit linguistic knowledge, it is possible to test Paradis’ (1994, 2004) assertion of greater accessibility to explicit rather than implicit linguistic knowledge in aphasia. Based on the declarative/procedural model as it applies to first and second language development (Paradis, 1994), we could predict cases of non-parallel recovery of MSA compared to spoken dialect in aphasia, reflecting the implicit nature of spoken dialect acquisition and the explicit learning of MSA. Current work is undertaking the evaluation of Arabic-speaking individuals with aphasia to investigate this hypothesis – an endeavor dependent on the adaptation of the BAT to the two language varieties of Arabic (e.g. Khamis-Dakwar et al., 2018). The findings of this ongoing research will contribute to enhancing the quality of language assessment of under-examined language varieties in diglossic speech communities. If the findings adhere to Paradis’s model, clinical intervention techniques may evolve based on the reported evidence. Moreover, theoretical discussions related to language recovery in bilingual and diglossic speech communities would be enhanced by the

1. All adaptations of BAT in different languages and cultural are accessible online via McGill University website <https://www.mcgill.ca/linguistics/research/bat>

availability of direct empirical evidence, potentially catalyzing additional investigations in other diglossic speech communities (e.g., speakers of African American English, Swiss German, or Cypriot Greek / Modern Greek) based on the distinctive characteristics of each specific diglossic situation with respect to language use and learning. Such theoretically based clinical linguistic inquiries would bring depth to our understanding of the interplay between learning mechanisms associated with distinct language varieties, language processing, linguistic representation in multilinguals, and language recovery in diglossic speech communities.

2.2 Diglossic codeswitching in Autism Spectrum Disorders (ASD)

2.2.1 *Clinical and theoretical background*

Autism Spectrum Disorders are defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM) 5th edition as neurodevelopmental disorders identified by persistent deficits in social communication and restricted interests and repetitive behaviors, with or without concomitant communication difficulties (American Psychiatric Association, 2013). These deficits can be identified through the manifestation of listed symptoms in each domain by history or by current behavioral presentations.

The literature on bilingualism and ASD is sparse. Few sources specifically address whether bilingual language exposure may affect the communicative and linguistic development of children with ASD. Some studies have performed comparative investigations of the social and linguistic abilities of children with ASD raised in bilingual versus monolingual environments (for review see Lund, Kohlmeier, & Durán, 2017). This line of study addresses a common perception that bilingual exposure could negatively impact the communicative-linguistic skills of children with ASD. Recent studies present empirical evidence against this notion, demonstrating that bilingual children with ASD have the ability to acquire and use two languages (Peterson, Marinova-Todd, & Mirenda 2012; Hambly & Fombonne, 2012; Lund, Kohlmeier, & Durán, 2017) and do not have greater delays in receptive-expressive language development than those raised in monolingual settings (Lund et al., 2017). It has also been noted that bilingual exposure does not impact executive functions in bilingual children with ASD (Li, Oi, Gondo, & Matsui, 2017).

Several benefits of bilingualism for children with ASD have been noted. For example, Gonzalez-Barrero & Nadig (2017) showed that bilingual children with ASD performed better at following directions than matched monolingual children with ASD. Peterson, Marinove-Todd, and Mirenda (2012) found that bilingual children with ASD had larger vocabulary size than monolingual peers. Indeed, there is initial evidence to suggest that bilingualism might even support communicative function in ASD, from an investigation of the use of codeswitching for

pragmatic functions by a child with ASD raised in a Chinese/English bilingual environment (Yu, 2016). Yu (2016) employed Conversational Analysis to analyze instances of Chinese-English codeswitching from video samples of the child interacting with their family (15 total hours of video data). The results showed that the child's codeswitching systematically served different pragmatic functions including clarification and emphasis, changes in footing, and commenting. These findings suggest that the purposeful use of codeswitching in bilinguals with ASD may also be evident in diglossic exposure, since the two language varieties in diglossia exist in a complementary functional distribution.

There has been progress in the identification of children with ASD, provision of services, and the establishment of a few research centers for ASD in the Arab world. However, there are no reports on the use of MSA, spoken Arabic or other languages within a typical linguistic profile of Arabic-speaking children with ASD. Anecdotally, parents and SLPs from different Arabic-speaking communities have reported that children with ASD exhibit frequent use of MSA and/or atypical development of MSA prior to spoken Arabic. Such reports have been dismissed as cases of echolalia, since MSA exposure typically includes cartoons on TV, and some children with ASD may exhibit pure delayed or immediate echoes of heard speech in social interactions.

However, an alternative account of such observations may be possible, since it could be the case that some children with ASD experience enhanced accessibility of MSA in comparison to their spoken dialects. This could be due to MSA's unique social status as a formal language, typically learned through schooling, which therefore relies on explicit rather than implicit learning mechanisms, therefore bypassing ASD's core deficits in social communication and interaction – similarly to the possible cases of non-parallel recovery of MSA in diglossic aphasia discussed above. To date, however, there have been no reported empirical investigations into the use of MSA and diglossic codeswitching in children with ASD in Arabic or other diglossic communities.

Here I present a case study of diglossic codeswitching in a 6-year old Palestinian child with high functioning ASD (Khamis-Dakwar, Asaad, & Salam, 2018), not only to exemplify the clinical impact of such studies and their use of linguistic analysis to answer a clinically-based question, but also to highlight their potential theoretical contribution. This pilot study used Conversational Analysis to examine MSA use during seven routine conversations collected and transcribed during this child's interactions with parents and siblings at home (breakfast, play), professionals in therapy (storytelling, guessing games, and show and tell), and self-play (symbolic play). All samples were transcribed using Jefferson transcription conventions and instances of MSA use and diglossic codeswitching were marked for later analysis of the pragmatic functions served.

The analysis revealed that of 279 conversational units, there were 47 occurrences of codeswitching from the spoken dialect to MSA. Moreover, two samples, recorded while playing house with a sibling and individual play with animal toys, indicated that MSA was the preferred language of communication (see Table 2). In addition, turn-taking, adjacency and repair analyses revealed that instances of sole MSA productions were not echolalic for this child, but were used mostly for self-talk and to indicate changes in participant roles. Codeswitching to MSA was also used to communicate change of footing, clarifications, and emphasis (see examples in Table 3).

Overall, these findings suggest that children with ASD, like other members of diglossic speech communities, may develop competence and use for all systems accessible to them, but may use these distinct systems for different functions – for example, switching to MSA to express specific communicative intentions or goals. The fluency levels of MSA productions for the child in this study are greater than expected from a child with typical development at the same age. This suggests a need to further investigate MSA acquisition in children with ASD in comparison to children with typical development matched for age, background, and exposure to different linguistic varieties.

Similar to the previously-presented work which began with the development of the Bilingual Aphasia test for the two language varieties in Arabic diglossia, this study also exemplifies the use of linguistic approaches, in this case Conversational Analysis, to analyze clinical data from a diglossic situation in a way that can inform SLP assessment. This linguistic approach is well established in clinical practice in speech-language-pathology and has been promoted in working with individuals with aphasia (Whitworth, 2003) as well as children with ASD (e.g. Solomon, Heritage, Yin, & Maynard, 2016; O'Reilly, Lester, & Muskett, 2016; Sterponi,

Table 2. MSA usage in seven transcribed samples produced by a 6-year old Arabic-speaking child with high functioning ASD

| Sampled session | Total conversational units | Total utterances | Number of MSA utterances |
|--|----------------------------|------------------|--------------------------|
| Board game with mother | 108 | 56 | 7 |
| Lunch with family | 110 | 43 | 10 |
| Category lotto game with clinician | 377 | 132 | 21 |
| Clinical session with occupational therapy | 5 | 5 | 0 |
| Story sequence activity with clinician | 5 | 5 | 0 |
| Symbolic play (alone) | 8 | 8 | 8 |
| Playing house with sibling | 55 | 30 | 29 |

Table 3. Examples of goal directed diglossic code-switching from sampled social-communication interactions of Arabic-speaking 6-year old child with high functioning ASD. OT=Occupational Therapist. Italicized bolded words are cases of codeswitching to MSA

| Emphasis | Excerpt 1 |
|--|--|
| | CHILD: (.) ?essa dori: (grabs dice and rolls) wa:hde fi:f ?andi wa:hde zaher] ‘Now it’s my turn (grabs dice and rolls) one I don’t have one pink’ |
| | MOTHER: [↑ki:f fi:f]?andak zaher? Fi: ↑ (.) ?etalla? mne:h (.) How you say you don’t have pink? You have. Look closely |
| | CHILD: fi: ?andi ↑ <i>fakat’ wa:hida</i> I have only one |
| | Excerpt 4 |
| | CHILD: (picks a card) <i>la: eθna:n faqat</i> no, only two. |
| | OT: le? Ho:n ?el ?axdar bye?bah ?el ?azra? kama::n (.)>> mef kti::r b?i::d X ki::f bedna nemsek ?el karet? ?a::ywan (..) fed fed ?al mal?at bi hado:l el ?osb- ?a::ywa fa:ter ‘why? Here is the green it looks like the blue one also. It’s not very far. How are we going to hold the card? Right! Tighter tighter on the clothes pin with these two fingers. Right! Good (boy)/ |
| Changes in Footing/ Commenting on one’s own state | MOTHER: =>Keef ka:n elyo:m ya S bel madrase?<= ‘how was your day at school?’ |
| | CHILD: ↑ <i>La:ysa ?elhamdullella</i> (1.9) ((pinches his lips as to seal them)) ↓ <i>?asna:ni:</i> <i>tu?limuni:</i> ‘Not thankful to Allah (God), my teeth are hurting me’ |
| Clarification | OT: xamse (.) stanna fwa:y ?anu: ?osba? ?ulna:? (.) Msek bel?asabe? mni:h ‘Five. Wait a little, which finger did we say? Hold (it) with both of your fingers well’ |
| | CHILD: <i>iθ ↑ na::n ↑ lakemana</i> bedna khamasa ‘Two! But we need five’ |

de Kirby, & Shankey, 2015; Sterponi & Shankey, 2014; Stribling, Rae, & Dickerson, 2006). Similarly, sample transcriptions were conducted using a linguistically-based method, the Jefferson transcription conventions. Our use of these linguistic methods mirror Yu’s (2016) methodology in analyzing Chinese-English codeswitching, and such approaches have been found useful for numerous investigations in the field. Information about the child’s social-communicative skills that would not otherwise have been available becomes readily apparent when these well-established methods are applied to diglossic situations.

2.2.2 *Potential clinical and theoretical contribution*

Khamis-Dakwar, Asaad, & Salam (2018) is the first reported case study on diglossic codeswitching in ASD, revealing that Arabic-speaking children with ASD may exhibit goal-directed diglossic codeswitching for a variety of communicative functions. No instances of echolalic MSA use were observed in the productions of this particular child. This preliminary investigation raises the possibility that children with ASD may access MSA preferentially compared to typically developing children, a finding that suggests potential enhanced accessibility to MSA versus spoken Arabic in the course of language acquisition in some children with ASD raised in diglossic communities. These differences may be supported by the differential learning mechanisms and contexts of use associated with the two language varieties in Arabic; possibly, these differences result in different markedness rankings for diglossic codeswitching in children with typical language development compared to children with ASD in their course of language development. New research exploring codeswitching in bilingual children with ASD can be informed by examinations of codeswitching in diglossia, again contributing to the theoretical discussion about the nature and functions of codeswitches in ASD.

These findings, though preliminary, suggest a clinical need to assess the use of MSA by children with ASD from Arabic-speaking communities. Such knowledge of language use in Arabic-speaking children with ASD may support the development of interventions that capitalize on their strengths in communication. I argue that such studies are necessary to position SLP services more effectively with respect to the cultural and linguistic differences that characterize diglossic speech communities.

To give a personal example, even though I have worked as an SLP for more than 20 years, and spent at least 6 of those years working with children with ASD in Nazareth, I did not recognize MSA use and codeswitching as a phenomenon worthy of my attention as a clinician. Similarly, speech-language pathologists in the Arab world typically would not attempt to examine the specific uses of MSA in ASD, and would automatically tend to index such instances as cases of echolalia. The evidence available and accessible to SLPs serving Arabic-speaking clients remains Anglo-centric in its framework, and SLPs serving Arabic-speaking clients are developing clinical assessments and interventions based on partial knowledge, without access to training in linguistic theoretical constructs and methods. So long as this remains the case, potential alternative explanations about the use of MSA in various communication disorders, in light of the differential learning mechanisms engaged and the markedness of codeswitching in diglossic language development, will remain unexplored.

3. Summary & discussions

Communication Sciences and Disorders is a growing scholarly and clinical field in Arabic-speaking communities around the world. High quality speech therapy services are dependent on making evidence-based clinical decisions regarding assessment of and intervention for individuals with communication disorders. Evidence-Based Practice (EBP) is defined as "...the conscientious, explicit, and judicious use of *current best evidence* in making decisions about the care of individual patients...[by] integrating the best available evidence, clinical expertise, and patients' values and preferences" (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996, as cited in ASHA, 2004, p. 1: emphasis mine). EBP has been endorsed as a guiding principle by the American Speech-Language and Hearing Association (ASHA, 2004). The implementation of EBP in working with Arabic-speaking clients with communication disorders in the Arab world and the diaspora relies on the integration of authentic high quality research evidence in clinical decision-making. Such research would not only be beneficial for enhancing the professional speech-language pathology practice, but also has the potential to inform theoretical queries in Communication Sciences and Disorders as well as linguistic fields of study.

The two examples presented in this paper exemplify the need for further systematic research utilizing rigorous clinical linguistic methods. Indeed, numerous theoretical propositions have benefited from clinical linguistic studies in English and other languages including markedness relationships, the Universal Grammar hypothesis, Optimality Theory, verb movement and finiteness, syllabic structure, and others (for an overview of such studies see Fava, 2002). Inclusion of the Arabic language, given its distinctive features, in these investigations may contribute to these lines of inquiry. In addition, theoretical Arabic linguistic studies may benefit from expanding their focus to include clinical studies such as those presented in this volume. For example, the distinct property of canonical emphasis in Moroccan Arabic, proposed by Aaron Freeman, or the constraints on Imala vowel-shifting suggested by Noor Abo Mokh and Stuart Davis, can be investigated and expanded through examinations of phonological development in Arabic-speaking children with and without phonological impairments. Such endeavors will bring depth to our understanding of the nature of specific communication disorders and highlight the contribution of Arabic to general linguistic inquiries as well as to Communication Sciences and Disorders.

Lastly, current clinical linguistic studies recognize the need to shift from serving mainly as a platform for clinical practice with linguistic knowledge and tools, towards comparative linguistic studies of disorders across languages. Such endeavors should move towards defining the nature of language disorders as expressed by the following quote from Crystal (2013):

We seem to be still at an anecdotal stage. I have seen a few comparative illustrations of disorders across languages – showing side by side data from, say, French and English fluent aphasia. I have used this approach myself ... But the examples are typically short and selective, and rarely do we see the author illustrating from more than two languages. The obvious way forward is to construct a methodology in which clinical linguists from a range of language backgrounds use the same descriptive procedure to process samples of data relating to particular disorders from their individual languages. We can step back and, in a metalinguistic way, compare and evaluate the outcomes.
(Crystal, 2013, p. 237)

Given the dialectal variations in Arabic, clinical Arabic linguistic studies can provide insights into the possible presentations of a specific disorder and the analysis of dialect-specific features associated with that disorder. Comprehensive investigations would bring theoretical and clinical experts in the field to recognize the value of bidialectal/diglossic studies in theoretical linguistics and CSD. Interprofessional, theoretically based, and clinically oriented collaboration between linguists and speech pathologists is a worthwhile effort to advance knowledge across fields.

References

- Alanazi, A. A. (2017). Audiology and speech-language pathology practice in Saudi Arabia. *International Journal of Health Sciences: A Scientific Publication by Qassim University*, 11 (5), 43–55.
- Al-Khaledi, M., Lincoln, M., McCabe, P., Packman, A., & Alshati, T. (2009). The attitudes, knowledge and beliefs of Arab parents in Kuwait about stuttering. *Journal of Fluency Disorders*, 43 (1), 44–59. <https://doi.org/10.1016/j.jfludis.2009.02.003>
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>
- American Speech-Language-Hearing Association (1993). *Definitions of communication disorders and variations*. Available from www.asha.org/policy
- American Speech-Language-Hearing Association (2004). *Evidence-Based Practice in Communication Disorders: An Introduction [Technical Report]*. Available from www.asha.org/policy
- American Speech-Language-Hearing Association (2016). *Scope of Practice in Speech Language Pathology [Scope of Practice]*. Available from www.asha.org/policy
- Barnes, S. & Ferguson, A. (2013). Conversation analysis applied to disordered speech and language. In Müller, & Ball, M. (Eds.), *Research Methods in Clinical Linguistics and Phonetics: A practical Guide* (pp. 126–145). Hoboken, NJ: Blackwell Publishing.
- Crystal, D. (1981). Aspects of clinical linguistic theory and practice. *Proceedings of the 1980 Annual Conference of the Australian Association of Speech and Hearing*, 1–25.
- Crystal, D. (2013). Clinical linguistics: conversational reflections. *Clinical Linguistics & Phonetics*, 27 (4), 236–243. <https://doi.org/10.3109/02699206.2012.726312>
- Fava, E. (2002). *Clinical linguistics: Theory and application in speech pathology and therapy*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.227>

- Ferguson, C. (1959). Diglossia. *Word*, 15, 325–340. <https://doi.org/10.1080/00437956.1959.11659702>
- Froud, K. & Khamis-Dakwar, R. (2012). Mismatch negativity responses in childhood apraxia of speech. *American Journal of Speech-Language Pathology*, 21(4), 302–312. [https://doi.org/10.1044/1058-0360\(2012/11-0003\)](https://doi.org/10.1044/1058-0360(2012/11-0003))
- Froud, K. & Khamis-Dakwar, R. (2018). Neurophysiological investigations in studies of Arabic linguistics: The case of Arabic diglossia. In Benmamoun, E. & Bassioune, R. (Eds.), *The Routledge Handbook of Arabic Linguistics* (pp. 306–319). Hoboken, NJ: Routledge Publishing.
- Gonzalez-Barrero, A. M. and Nadig, A. S. (2017). Can bilingualism mitigate set-shifting difficulties in children with autism spectrum disorders. *Child Development*, 90(4), 1043–1060.
- Hambly, C. & Fombonne, E. (2012). The impact of bilingual environments on language development in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42 (7), 1342–1352. <https://doi.org/10.1007/s10803-011-1365-z>
- Human Rights Watch (2001). *Second class: Discrimination against Palestinian Arab children in Israel's schools*. NY: Human Rights Watch.
- Ismail, H. (2015). *Challenges facing speech-language pathologists in the Arab world*. Seminar presentation at the American Speech-Language-Hearing Association Convention, Denver, CO.
- Khamis-Dakwar, R. & Froud, K. (in press). The study of Arabic language acquisition. To appear in Ryding, K. & Wilmsen, D. (Eds.), *The Cambridge Handbook of Arabic Linguistics*. Cambridge, UK: Cambridge University Press.
- Khamis-Dakwar, R. & Froud, K. (2019). Diglossia and language development. In Al-Wer, E. & Horesh, U. (Eds.). *The Routledge handbook of Arabic Sociolinguistics* (pp. 300–314). Oxford, UK: Blackwell Routledge.
- Khamis-Dakwar, R., Asaad, H., & Salam, I. (2018). *An exploratory study of MSA use in Arabic-speaking children with Autism Spectrum Disorders*. Research presentation at the Forum for Arabic Linguistics, School of Oriental and African Studies, U.K.
- Khamis-Dakwar, R., Ahmar, M., Farah, R., & Froud, K. (2018). Adaptation of the Bilingual Aphasia Test (BAT) to Arabic diglossia: Clinical and theoretical implications. *Journal of Neurolinguistics*, 47, 131–144. <https://doi.org/10.1016/j.jneuroling.2018.04.013>
- Khamis-Dakwar, R. (2012). Translational display of neurophysiological investigations in communication sciences and disorders. In Goldfarb, R. (Ed.), *Translational SLP/A: Essays in honor of Dr. Sadanand Singh* (pp. 267–274). San Diego, CA: Plural Publishing, Inc.
- Kotby, N. (2014). *News from different nations: Egypt*. Retrieved from <https://www.phoniatrics.eu/uep/news/news-from-different-nations-a-m/egypt/index.html>
- Kohnert, K. (2013). *Language disorders in bilingual children and adults* (2nd ed.). San Diego, CA: Plural Publishing.
- Li, H., Oi, M., Gondo, K., & Matsui, T. (2017). How does being bilingual influence children with autism in the aspect of executive functions and communication competence. *Journal of Brain Science*, 47, 21–49.
- Lund, E. M., Kohlmeier, T., & Durán, L. K. (2017). Comparative language development in bilingual and monolingual children with Autism Spectrum Disorder: A systematic review. *Journal of Early Intervention*, 39 (2), 106–124. <https://doi.org/10.1177/1053815117690871>
- Mostafa, E., & Ahmed, M. E. (2017). Public awareness of delayed language development in Upper Egypt. *The Egyptian Journal of Otolaryngology*, 43 (1), 94–102.
- O'Reilly, M., Lester, J. N., & Muskett, T. (2016). Discourse/Conversation analysis and Autism Spectrum Disorder. *Journal of Autism Development and Disorders*, 46 (2), 355–359. <https://doi.org/10.1007/s10803-015-2665-5>
- Paradis, M. (1987). *The Assessment of Bilingual Aphasia*. New York: Psychology Press.

- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 393–419). London: Academic Press.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (v.18). John Benjamins Publishing. <https://doi.org/10.1075/sibil.18>
- Paradis, M. (2008). Language and communication disorders in multilinguals. In Stemmer, B. & Whitaker, H. A. (Eds.), *Handbook of the Neuroscience of Language* (pp. 341–349). Oxford: Elsevier Press. <https://doi.org/10.1016/B978-008045352-1.00033-1>
- Petersen, J. M., Marinova-Todd, S. H., & Mirenda, P. (2011). Brief Report: An exploratory study of lexical skills in bilingual children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 42(7), 1499–1503. <https://doi.org/10.1007/s10803-011-1366-y>
- Sterponi, L., de Kirby, & Shankey, J. (2015). Rethinking language in autism. *Autism*, 19(5), 517–526. <https://doi.org/10.1177/1362361314537125>
- Sterponi, L., & Shankey, J. (2014). Rethinking echolalia: Repetition as interactional resource in the communication of a child with autism. *Journal of Child Language*, 41(2), 275–304. <https://doi.org/10.1017/S0305000912000682>
- Stribling, P., Rae, J., Dickerson, P. (2006). To forms of spoken repetitions in a girl with autism. *International Journal of Language and Communication Disorders*, 42 (4), 427–444. <https://doi.org/10.1080/13682820601183659>
- Solomon, O., Heritage, J., Yin, L., & Maynard, D. W. (2016). What brings him here today?: Medical problem presentation involving children with Autism Spectrum Disorders and typically developing children. *Journal of Autism and Developmental Disorders*, 46, 378–393. <https://doi.org/10.1007/s10803-015-2550-2>
- Wu, S. (2015, November 3). International speech pathologist series: Dubai [Blog post]. Retrieved from <https://speechisbeautiful.com/2015/11/international-speech-pathologist-series-dubai/>
- Whitworth, A. (2003). The application of conversation analysis (CA) to the management of aphasia. *Publié dans Revue Tranel (Travaux neuchâtelois de linguistique)*, 38/39, 63–76.
- Yu, B. (2016). Codeswitching as a communicative resource within routine, bilingual family interactions for a child on Autism Spectrum. *Perspectives of the ASHA Special Interest Group* 14, 1(1), 17–28. <https://doi.org/10.1044/persp1.SIG14.17>

Index

A

- adjunction 87–88, 93, 96–97, 99–100, 102–103
 - adjuncts 84, 96, 98, 100, 136, 138, 140, 145, 150–151
 - high adjuncts 98, 100
 - agreement 3–5, 7, 109, 111, 114–115, 117–120, 122, 125, 128, 132–134
 - Algerian Darija 55, 57
 - alternatives 115–116
 - antepenultimate 11–15, 17–18, 23, 26, 31
 - aphasia 6, 155, 160–163, 165–166, 170–172
 - argument 3, 5, 110–111, 122, 128, 133–134, 136–139, 144–145, 150
 - assimilation 41–42, 44, 74
 - Autism Spectrum Disorders 155, 160, 164, 171–172
- ## C
- cartographic/y 2, 5, 71, 73–74, 87, 91, 95–100, 102–3, 104, 106, 111, 122
 - clitic 4, 109, 118–119, 121–123
 - codeswitching 6, 164–168
 - Communication Sciences and Disorders 155–156, 158, 169, 171
 - comparative 71–74, 76–81, 83–84, 88–93, 95–96, 104–107, 152
 - conditional 72, 81, 91, 106
 - correlative 72–73, 78, 84, 105–107
 - standard of comparison 78, 90, 93
 - complementizer 5, 7, 74, 76, 79, 82, 85, 87, 89–91, 94, 96–99, 104, 109–112, 115, 117–121, 124, 128, 134, 144

Construction Morphology

- 46–48, 52
- conversational analysis 165–166
- copula 6, 137, 139, 151
- correlation 71–74, 76–81, 87, 90–93, 96, 100–101, 104, 107
 - correlative *see* correlation
- cycle 6–7

D

- DegP 74, 77–81, 88–93, 95–96, 100, 104
- diglossia 5, 60, 66–67, 155, 159–160, 163, 165–166, 168, 171
- dorsal fricative 4–5, 55–57, 61–66

E

- Egyptian Arabic 3, 5, 71, 73, 104, 107, 136, 152
 - Cairene 6, 53, 55, 57–59, 61–66, 152, 156, 158, 171
 - see also* Egyptian Arabic
- extrametricality 12–13, 16, 18–19, 21, 23, 26, 31

F

- Faifi (Saudi Arabia) 5, 55, 57–59, 63–65
- feature geometry 4, 35, 56
- feminine suffix 15, 35–39, 42, 46–47, 49–51, 56
- floating mora 20
- foot 4, 11–12, 16–19, 21–31
 - antepenultimate 18
 - degenerate 12, 15, 17–19, 21, 23, 25, 31

G

- geminate 12–15, 28–29, 31
- grammaticalization 7
- guttural 41–42, 51–53, 55–56, 65, 67

I

- 'imāla 35–51, 53
- imperative 74, 82, 86, 94
- in situ 77, 79, 90–91, 95
- information structure 71, 74, 98, 100–103, 120–121, 133

L

- Left Periphery 2, 73–74, 76–77, 79, 81, 88–91, 97–98, 103–104, 106, 109–111, 115, 118, 121, 129–134
- lexical doubling 77, 85, 93–96, 98, 100, 103–104
- loanwords 5, 35–37, 48–52

M

- Maltese 90, 93, 96
- measure PP 71, 73, 91–93, 96, 98, 103–104
- mora 13, 18, 20, 26, 32
 - augmentation 20, 32
- Moroccan Darija 55, 57
- morphosemantic class 35, 46

N

- nearest neighbor distance 63, 65
- negation 3, 6, 74, 106, 115, 133, 162
- nonexhaustivity 16
- non-proportional correlation 80, 93, 96, 104
- non-scalar correlation 79, 92–93, 96, 104

O

- object 3, 111, 116, 122, 125, 131, 136, 146–148, 150
- Optimality 4, 11, 13, 23, 32–33, 47, 169

P

- Palestinian Arabic 4, 7, 35–37, 50–53, 66, 133, 160, 162–163
 peripherality 15–16
 persistent footing 15, 17–18, 21, 25
 pre-antepenultimate 11–15, 17–18, 23, 26, 31
 Preposition Stranding
 Generalization 135, 141, 151
 Priority Clause 15, 17, 21, 25–26
 prominence 4, 11–12, 22–23, 28–31
 pronoun 3–4, 109, 117–118, 122–123, 125, 127, 137–139, 142, 144–145, 147–150
 copular 139, 144, 149–150
 resumptive 137–138, 142, 144–145
 proportional correlation 74, 77, 80, 93, 96, 104
 pseudo-sllicing 135, 137, 139–140, 142, 144–145, 149–151

Q

- quantificational 74, 76, 91
 quantifier 71, 73, 76, 82, 90, 92, 96, 98, 103–104, 123, 127, 132
 raising 123
 variable 92, 96, 98, 103–104
 quantity 4, 11–14, 31–32, 79
 QUD 115

R

- root complementizer 97

S

- scalar correlation 77, 79, 91–93, 96, 100, 104
 sensitivity 11–12, 31, 91
 Sibawayhi 35, 53
 sllicing 5, 135–152
 Speech Language Pathology 5, 158, 170
 sprouting 5, 135–137, 145–151
 stress 4, 11–33, 49–50
 antepenultimate 11–12, 15, 18, 23
 subject 2–3, 5, 82, 109, 116–117, 120, 122, 125, 127–129, 137, 146–147
 subjunctive 82–83
 subordination 71, 73, 81–85, 87, 93, 95–96, 98, 103
 suppletion 46
 syllable 11–19, 21–22, 24–26, 29–32
 antepenultimate 12–13
 syncope 12–13, 15–17, 19–21, 23, 27–28, 30–31, 33
 Syrian (Aleppo) 5–6, 36, 53, 55, 57–59, 63–65

T

- tag questions 82, 85–87, 90
 topic 80, 84, 94, 98, 100–102, 104, 111, 113, 116, 121–123, 127, 132, 136
 Triangle-Area Palestinian 55, 57
 trochee 4, 11, 22, 24

U

- ultrasound 5, 55, 57, 60–61, 66–67
 unselective binding 91, 93, 100, 106

V

- Veracity 115
 Verum 5, 109, 114–115, 131, 133
 vowel
 raising 4, 35–42, 44–49, 51–52
 shortening 20, 46–47, 50

S

- stress 4, 11–33, 49–50
 antepenultimate 11–13, 15, 18, 23

V

- VSO 2–3, 120

W

- Weak Local Parsing 4, 11–12, 15–16, 18–19, 23, 26, 31
 wh-clefting 136–137, 139–142, 144–146, 148–151
 wh-fronting 136–141, 144–146, 148–150
 word order 2, 7, 74, 82–83, 95–96, 99, 133–134, 152

This volume presents a collection of seven peer-reviewed articles on Arabic phonetics, phonology, syntax, semantics, and applied linguistics. The authors address stress assignment, the phenomenon of 'imāla, the place of articulation of the dorsal fricative, the structure of correlatives, the CP layer, sluicing and sprouting, and clinical linguistics. They do so by using data from Standard Arabic, and from Egyptian, Jordanian, Palestinian, and Saudi Arabian varieties of Arabic. The book will be of interest to linguists working in descriptive and theoretical areas of Arabic linguistics.

ISBN 978 90 272 0759 3



9 789027 207593

John Benjamins Publishing Company