
The 34th Annual Symposium on Arabic Linguistics
February 28 - March 1, 2020



Arabic Linguistics Society
رابطة اللسانيات العربية



COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
**School of Middle Eastern
& North African Studies**

المؤتمر السنوي الرابع والثلاثون لللسانيات العربية

University of Arizona
February 28 – March 1 , 2020
The Santa Rita Room and the Rincon Room
Student Union Memorial Center (SUMC)
University of Arizona

Conference Website: <https://arabiclinguistics.sbs.arizona.edu/>

Keynote Speakers



Rania Habib
*Associate Professor, Syracuse
University, USA*



Mohamed Embarki
*Professor, Université de
Franche-Comté, Besançon,
France*



Samira Farwaneh
*Associate Professor,
University of Arizona,
USA*



Ahmad Alqassas
*Assistant Professor,
Georgetown University,
USA*

TABLE OF CONTENTS

Table of Contents.....	2
Welcome.....	3
Acknowledgments.....	4
Getting Around.....	7
Conference Schedule	8
Abstracts: [Note: All abstracts are ordered according to their sequence in the schedule].....	12

WELCOME

Dear Participants,

We are very happy to welcome you to the 34th Annual Symposium on Arabic Linguistics (ASAL34) at the University of Arizona. This is the second time the conference is held in Tucson, and we hope it continues to open the door to future Arabic Linguistics events. We are grateful to our sponsors for enabling us to bring ASAL34 to the University of Arizona, and to our respective department chairs, administrative staff and students for helping us to make it happen.

We are fortunate to have four plenary keynote speakers this year: Dr. Mohamed Embarki, Professor of Language Sciences at the University of Franche-Comté, Besançon, France.; Dr. Rania Habib, Associate Professor of Linguistics and Arabic at Syracuse University; Dr. Samira Farwaneh, Associate Professor of Arabic Linguistics at the University of Arizona; and Dr. Ahmad Alqassas, Assistant Professor of Arabic at Georgetown University.

We wish to extend special thanks to our abstract reviewers for helping us navigate the large number of abstracts we had received and select submissions of the highest quality. In keeping with ASAL tradition, there are no parallel sessions. ASAL34 has eleven single sessions featuring 30 paper presentations and seven poster presentations. Please note that the prize for best student abstracts will be recognized at the Arabic Linguistics Society's Business Lunch Meeting on Saturday, February 29, 2020.

ASAL34 Banquet Dinner/Reception will be held in one of the most iconic rooms at the University of Arizona: *The Arizona Room*. Occupying the former Redington Restaurant space (3rd fl., SUMC), Arizona Room combines an elegant, relaxed setting with some of the best views on campus. You will enjoy the new buffet-style dining menu with an added emphasis on table service. We hope that you will be able to join us. The ASAL34 Banquet Dinner/Reception is free.

For important information about the conference, please visit the website:

<https://arabiclinguistics.sbs.arizona.edu/>

If there is anything else, we can do to make your stay comfortable please don't hesitate to let us know by emailing us at mazaz@email.arizona.edu OR asal34arizona@gmail.com.

Finally, we wish to thank the Executive Director of the Arabic Linguistics Society, Dr. Hamid Ouali, for answering our questions and sharing his extensive experience with us, and we look forward to sharing ours with ASAL35 organizers.

Enjoy the conference and the city of Tucson, the heart of the Southwest of the United States!

ASAL34 Organizers

Event Organizer

Mahmoud Aḡaḡ

School of Middle Eastern and North African Studies (MENAS)

Organizing Committee Members (ordered alphabetically by first name)

Anne Betteridge (MENAS & CMES)

Julie Ellison-Speight (CMES)

Samira Farwaneh (MENAS)

Sonia Shiri (MENAS)

ACKNOWLEDGMENTS

Special thanks to the following sponsors at the University of Arizona for their generous support (ordered alphabetically by their official name):

Adel Gamal (Emeritus Professor of Arabic Literature)

American Institute for Maghrib Studies

Arizona Arabic Flagship Program

Center for Educational Resources in Culture, Language and Literacy (CERCLL)

Center for Middle Eastern Studies

College of Humanities

College of Social and Behavioral Sciences

Department of East Asian Studies

Department of French and Italian

Department of Linguistics

Department of History

Department of Spanish and Portuguese

Interdisciplinary Ph.D. Program in Second Language Acquisition and Teaching

School of Middle Eastern and North African Studies

Adel S. Gamal



COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
Center for Middle Eastern Studies

College of Humanities

College of Social & Behavioral Sciences

COLLEGE OF HUMANITIES
East Asian Studies

COLLEGE OF HUMANITIES
French Studies

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
Linguistics

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
History

COLLEGE OF HUMANITIES
Spanish & Portuguese

Second Language Acquisition & Teaching

Graduate Interdisciplinary Program
COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Middle Eastern & North African Studies

Special thanks to our abstract referees (ordered alphabetically by first name):

- | | |
|------------------------|--|
| 1. Abbas Benmamoun | <i>Duke University</i> |
| 2. Abdel-Khalig Ali | <i>University of Toronto</i> |
| 3. Abdulkafi Albirini | <i>Utah State University</i> |
| 4. Adam Ussishkin | <i>University of Arizona</i> |
| 5. Ahmad Alqassas | <i>Georgetown University</i> |
| 6. Alexander Magidow | <i>University of Rhode Island</i> |
| 7. Amel Khalfaoui | <i>University of Oklahoma</i> |
| 8. Atiqa Hachimi | <i>University of Toronto</i> |
| 9. Christopher Lucas | <i>SOAS University of London</i> |
| 10. David Wilmsen | <i>American University of Sharjah</i> |
| 11. Ellen Broselow | <i>Stony Brook University</i> |
| 12. Islam Youssef | <i>University of South-Eastern Norway</i> |
| 13. John Eisele | <i>College of William and Mary</i> |
| 14. Keith Walters | <i>Portland State University</i> |
| 15. Kristen Brustad | <i>University of Texas at Austin</i> |
| 16. Lotfi Sayahi | <i>University of Albany, SUNY</i> |
| 17. Mohammad Alhawary | <i>University of Michigan</i> |
| 18. Mohammad Mohammad | <i>University of Texas at Austin</i> |
| 19. Mustafa Mughazy | <i>Western Michigan University</i> |
| 20. Nancy Hall | <i>California State University Long Beach</i> |
| 21. Peter Hallman | <i>Austrian Research Institute for Artificial Intelligence</i> |
| 22. Reem Khamis-Dakwar | <i>Adelphi University</i> |
| 23. Samantha Wray | <i>New York University Abu Dhabi</i> |
| 24. Stuart Davis | <i>Indiana University</i> |
| 25. Uri Horesh | <i>University of Essex</i> |
| 26. Usama Soltan | <i>Middlebury College</i> |
| 27. Youssef Haddad | <i>University of Florida</i> |
| 28. Zafer Lababidi | <i>Florida State University</i> |

Special thanks to our ASAL34 session chairs (ordered alphabetically by first name)

- | | |
|----------------------|-----------------------|
| 1. Adam Benkato | 6. Maris Camilleri |
| 2. Ahmad Alqassas | 7. Mohamed Embarki |
| 3. Alexander Magidow | 8. Peter Hallman |
| 4. Amel Khalfaoui | 9. Reem Khamis-Dakwar |
| 5. Hamid Ouali | 10. Samantha Wray |
| | 11. Stuart Davis |

Special thanks to our keynote speakers' introductions (ordered alphabetically by first name)

1. Peter Hallman (Introducing Dr. Embarki)
2. Reem Khamis-Dakwar (Introducing Dr. Habib)
3. Stuart Davies (Introducing Dr. Farwaneh)
4. Usama Soltan (Introducing Dr. Alqassas)

Special thanks to the following for their contributions and support:

College of Social and Behavioral Sciences (SBS)

- Dean, John Paul (JP) Jones III
- SBS Associate Deans for Faculty Affairs and Inclusion, Monica Casper
- SBS Associate Dean for Research and Graduate Studies, Jane Zavisca
- SBS Research Institute Director, Elizabeth Stahmer
- SBS Business Office Senior Manager, Kristine Hogeboom
- SBS Director of Information Technologies (website development), Edward Xia and Zaragoza Saenz

School of Middle Eastern and North African Studies (MENAS)

- Director and Professor, Benjamin Fortna
- MENAS Coordinator, Sergio Cañez
- MENAS Administrative Assistant, Randa Abdu

MENAS Graduate Students volunteers (ordered alphabetically by first name)

- Aaron Graybill, PhD Student
- El-Sayed Sabry, PhD Student
- Essa Alfaifi, PhD Student
- Renee Spellman, PhD Student
- Rose Hattab, PhD (President of MENAS Undergraduate Student Association)
- Ryan Blair, PhD Student
- Scott Jones, MA Student (President of MENAS Graduate Student Association)

MENAS Undergraduate Volunteers (ordered alphabetically by first name)

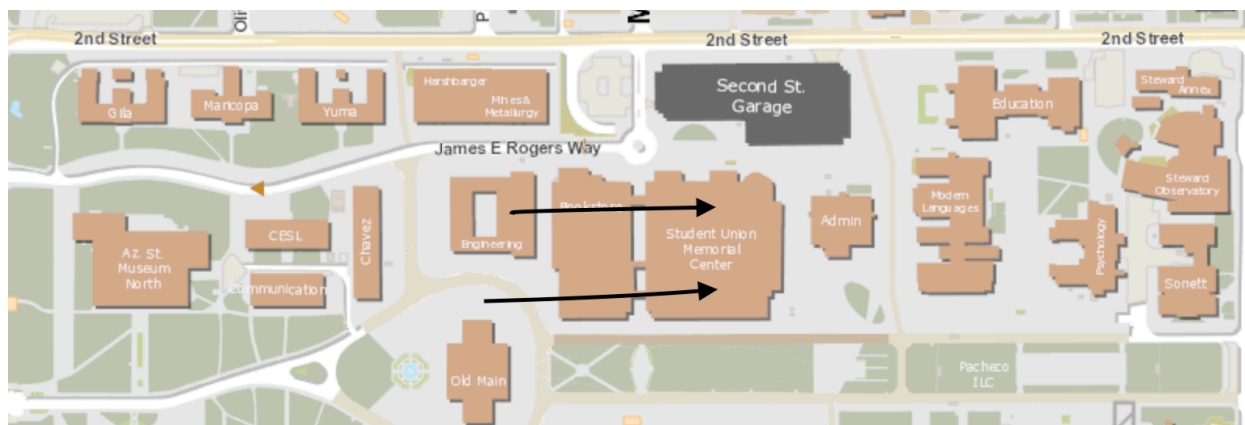
- Bethany Wepler, Flagship Student
- Evelyn Brezenski
- Madeliene Rodish, Flagship Student
- Marissa Haskell (Co-President of MENAS Undergraduate Student Association)
- Jennifer Castro

This booklet has been prepared by Mahmoud Azaz

GETTING AROUND

Conference venue

ASAL34 is held in **Student Union Memorial Center(SUMC), University of Arizona in the Santa Rita Room and the Rincon Room**. Both rooms are located in the **third floor of the Union**. The street address is 1303 E. University Blvd, Tucson, AZ 85719. To know the map of the University of Arizona, please visit this website: <https://map.arizona.edu/>



Eating around the Student Union Memorial Center

There are plenty of restaurants and places to eat around the conference venue. For all dietary restrictions (Halal, Kosher, Vegetarian, Vegan, Gluten free, etc.) please visit these links: <https://www.union.arizona.edu/dining/> and <https://www.union.arizona.edu/dining/sumc/?q=sumc&p=> to view all nearby eateries.

Venue of the banquet/reception on Saturday at 6:15pm

ASAL34 Banquet Dinner/Reception will be held in one of the most iconic rooms at the University of Arizona: **The Arizona Room (Student Union, 3rd floor, room # 305)**. To get from the conference venue to the Arizona Room, see the map below. If you are coming from elsewhere, here is the complete address is 1303 E. University Blvd Tucson, AZ 85719. Once there, there will be signs to guide you to the the Arizona Room on the 3rd floor of the Union.



<https://www.union.arizona.edu/dining/sumc/arizonaroom>

CONFERENCE SCHEDULE

Friday, February 28 (The Santa Rita Room)

8:00-8:45 **Registration and Coffee**
 8:45-9:00 **Opening Remarks**
 John Paul Jones III, Dean of College of Social and Behavioral Sciences

SYNTAX I

Chair: Dr. Stuart Davis

9:00-9:30 **Usama Soltan (Middlebury College)**
 Deriving Null Objects in Egyptian Arabic: Evidence for Argument Ellipsis
 9:30-10:00 **Hamid Ouali (University of Wisconsin-Milwaukee)**
 The Correlation Between Verb Movement and Tense in Arabic

10:00-11:00 *Keynote*

Introduction by Dr. Usama Soltan

Ahmad Alqassas (Georgetown University)
 The Syntax of Polarity and Coordination in Arabic

11:00-11:15 Coffee Break

PHONOLOGY I

Chair: Dr. Samantha Wray

11:15-11:45 **Abdel-Khalig Ali (University of Toronto)**
 On the Realization of the Glottal Stop in Sudanese Arabic
 11:45-12:15 **Abdullah Alfaifi & Stuart Davis (Indiana University)**
 Some Properties of the m-Definite Article in Upper Faifi Arabic
 12:15-12:45 **Ahmed Alnuqaydan (University of Utah)**
 Optionality in Avoidance of CCC Clusters in Qassimi Arabic

12:45-2:00 Lunch (on your own)

DIGLOSSIA, CORPUS, & DISCOURSE

Chair: Dr. Amel Khalfaoui

2:00-2:30 **Khaled Elghamry & Attia Youseif (Indiana University)**
 ArabiLight: A Corpus-Based Model for Spoken Modern Standard Arabic
 2:30-3:00 **Essa Alfaifi (University of Arizona and Jazan University)**
 An Emergent Grammar Model for Diglossic Variations between Arabic and Faifi
 3:00-3:30 **Karen McNeil (Georgetown University)**
 When the Leak Becomes a Flood: Destabilizing Diglossia in Tunisia
 3:30-4:00 **Khaled Elghamry (Indiana University)**
 Detecting and Measuring Hate Speech in Arabic Social Media

4:00-4:15 Coffee Break

Keynote

4:15-5:15 Samira Farwanah (University of Arizona)

Introduction by Dr. Stuart Davis

Do Languages Really “Undress”? Modern Arabic Varieties Between Simplicity and Complexity

SYNTAX II

Chair: Dr. Ahmad Alqassas

5:15-5:45 **Maris Camilleri & Louisa Sadler (University of Essex)**

5:45-6:15 Trajectories of Change: The Formation of Lexical Raising Predicates and Subj-to-Subj Raising Constructions in Vernacular Arabic
Basem Al-Raba'a (KIMEP University)
 Is Morphological Case a Feature of Individual Nominal Elements? Evidence from Standard Arabic

Saturday, February 29 (The Rincon Room)

7:30-8:30 Registration and Coffee

SOCIOLINGUISTICS

Chair: Dr. Adam Benkato

8:30-9:00 **Ekab Al-shawashreh (Yarmouk University) & Marwan Jarrah (University of Jordan)**
 Variation in Verbal Negation in Jordanian Arabic

9:00-9:30 **Vladimir Kulikov, Najlla Al-Hajri, & Buthaina Al-Kuwari (Qatar University)**
 Generational Changes in VOT in Qatari Arabic

9:30-10:00 **Sarah Schwartz (University of Texas at Austin)**
 Linguistic Globalization and New Urbanism: Mapping a Single Indexical Field of [tʃ] in French, Cairene Arabic, and Casablancon Arabic

10:00-11:00 *Keynote*
Rania Habib (Syracuse University)
Introduction by Dr. Reem Khamis-Dakwar
 Variation in Syrian Arabic

11:00-11:15 Special Guest & Coffee Break

SYNTAX III

Chair: Dr. Hamid Ouali

11:15-11:45 **Peter Hallman (Austrian Research Institute for Artificial Intelligence)**
 “ʔadd” as an Equative Degree Quantifier in Levantine Arabic

11:45-12:15 **Maris Camilleri (University of Essex)**
 “Ps” in the Varied Perfect Expressing Constructions across Arabic

12:15-12:45 **Marwan Jarrah (University of Jordan)**
 Agree Blocks Movement: Evidence from Arabic

**12:45-2:00 Catered Lunch and Business Meeting
 (The Rincon Room)**

Business Meeting Chair: Dr. Hamid Ouali

2:00-3:30 PLENARY POSTER SESSION

1. **Othman Alshehri (University of Arizona)**
 Blocking Effects of Nominalization on Arabic Passives: A Case of a Syntax-Semantics Conspiracy
2. **Letizia Cerqueglini (Tel Aviv University)**
 Conditional Clauses in Modern Standard Arabic and Jisr az-Zarqā Arabic: A Semantic Comparison
3. **Elsayed Issa (University of Arizona) & Aitor Arronte (University of Hawaii at Manoa)**
 Summarization of Moroccan Arabic News Texts Using Probabilistic Topic Modeling for Arabic L2 Micro-Learning Tasks
4. **Amal El-Haimeur (University of Kansas)**
 The Acquisition of Verbal-Derivational Patterns by Moroccan HS in France
5. **Hussein Al-Bataineh & Saleem Abdelhady (Memorial University of Newfoundland)**

6. The Morphosyntax of Exclamatives in Arabic: Against a Clausal Cartographic Analysis
Abdessamad M'barki (University of Arizona) & Ahmed Berrissoul (University of Mohamed V)
7. The Semantic Structure of the Spatial Arabic Prepositions, *fī* and *‘ala*
Saja Albuarabi (University of Wisconsin-Milwaukee)
The Syntactic Properties of Negative Concord in Iraqi Arabic

ACQUISITION

- 3:30-4:00 **Reem Khamis-Dakwar, Iman Salam, Amanda Nagler (Adelphi University) & Karen Froud (Columbia University)**
Arabic and English Plural Formation in Typically Developing Heritage Speakers of Arabic: A Longitudinal Study
- 4:00-4:30 **Iman Albadar (University of Delaware)**
Are Arabic Listeners 'Stress Deaf' to their Own L2 Pronunciation?
- 4:30-5:00 **Anna Gavarró & Iman el Hadeff Hammad (Universitat Autònoma de Barcelona)**
The Acquisition of Word Order in Early Moroccan Arabic
- 5:00-5:15 **Coffee Break**
- 5:15-6:15 ***Keynote***
Mohamed Embarki (Université de Franche-Comté, France)
Introduction by Dr. Peter Hallman
The Dynamic of the Sound System from Proto-Arabic to the Modern Dialects: Between Retention and Innovation
- 6-15-9:00 Banquet/Reception**
Arizona Room, Student Union Memorial Center, Third Floor
1303 E. University Blvd Tucson, AZ 85719

Sunday, March 1 (The Rincon Room)

- 8:45-9:00 **Coffee**

HISTORICAL LINGUISTICS

Chair: Dr. Mohamed Embarki

- 9:00-9:30 **Adam Benkato (University of California-Berkeley)**
Multiple Origins, Convergence, and the Earliest Maghribi Varieties: The Case of Šibilant-Merging
- 9:30-10:00 **Alexander Magidow (University of Rhode Island)**
Subgrouping Arabic Dialects via Historical Developments in Closed-Class Morphemes

MORPHOLOGY & PSYCHOLINGUISTICS

Chair Dr. Reem Khamis-Dakwar

- 10:00-10:30 **Laura Faircloth (University of Texas at Austin)**
Models of Root Co-Occurrence Restrictions and Their Implications for Arabic Phonology
- 10:30-11:00 **Samantha Wray (New York University Abu Dhabi)**
Morphosyntactic Derivation of the Form VIII Verb: Stuck in the Middle Again
- 11:00-11:30 **Ali Idrissi, Eiman Mustafawi, Tariq Khwaileh (Qatar University) & John E. Drury (Jiangsu Normal University)**
The Possible Root of Consonantal Root Effects in Arabic: Evidence from Neurophysiological Data

PHONOLOGY II

Chair Dr. Maris Camilleri

- 11:30-12:00 **Michael Grosvald, Noora Al-Ansari, & Ali Idrissi (Qatar University)**
Audiovisual Integration in Spoken Arabic: The Role of Lexicality and Syllabic Affiliation
- 12:00-12:30 **Saja Albuarabi & Park Hanyong (University of Wisconsin -Milwaukee)**
Production and Perception of Consonant Clusters in Nonwords by Iraqi and Najdi Speakers
- 12:30 -1:30 Lunch (on your own) and Coffee**

SYNTAX IV

Chair Dr. Alexander Magidow

- 1:30-2:00 **Saleem Abdelhady, Phil Branigan, & Hussien Albataineh (Memorial University of Newfoundland)**
Decomposing Arabic Complementizers
- 2:00-2:30 **Mohamed Jlassi (Sohar University)**
The Syntax of Confirmationals in Omani Arabic
- 2:30-3:00 **Mohammed Abuhaib (Imam Muhammad ibn Saud Islamic University & University of Wisconsin-Milwaukee)**
Notes on Najdi Arabic Scope Taking
- 3:00– 3:15 Closing Remarks**
Professor Benjamin Fortna, Director of the School of Middle Eastern and North African Studies

ABSTRACTS

Abstracts of keynote presentations *(ordered according to their sequence in the schedule)*

The Syntax of Polarity and Coordination in Arabic
Ahmad Alqassas, Georgetown University
aa1453@georgetown.edu

Coordinate complexes display a variety of empirical puzzles that challenge our standard assumptions about the nature of syntactic categories, operations and structural configurations available in syntax. The theoretical debates diverge on whether coordinate complexes have a special syntactic status necessitating the addition of syntactic categories, operations or configurations. Another debatable issue is whether the clause structure of coordinated DPs involves phrasal or clausal coordination at LF (the Logical Form). This talk focuses on empirical puzzles in Arabic involving the structure of coordinate complexes and their interaction with negation, phi-agreement, and ellipsis. A close look at the syntax of Arabic coordinate complexes reveals that, unlike their English counterparts, coordinators such as *wala* ‘nor’ are the lexicalization of a syntactic head that takes the internal conjunct as a complement and the external one as a specifier. Furthermore, the categorial status of the coordination phrase is determined by the syntactic category of the external conjunct and the coordination head is specified for the formal features of number and negation. We shall see that in negatively disjoined DPs [*laa* DP...*wala* DP], the coordinators are Negative Concord Items (NCIs) carrying uninterpretable negation features which are licensed by a negative operator under Agree. These coordinators are conspicuous analogues of the person-NCI *wala*-NP ‘no body’ in their ability to co-occur with negative adverbs in the CP layers without triggering a double negation reading. I will show that the same simple invisibles behind licensing NCIs are also the licensing conditions for these coordinate complexes. I then show that Closest Conjunct Agreement (CCA) in negatively disjoined DPs involves clausal coordination and ellipsis, while CCA elsewhere does not.

Do languages really ‘undress’? Arabic varieties between simplicity and complexity
Samira Farwaneh, University of Arizona
farwaneh@email.arizona.edu

The question of grammatical complexity in language and how relative complexity might be measured has moved to the forefront of linguistic investigation in the last decade. This is a stark departure from the central axiom of linguistics which has long been that languages are equally complex in their structure. Hence, there are no “primitive” or simplex languages. With respect to Arabic, the notion of complexity is inexorably linked with the diglossic use of Arabic. The now familiar diglossic situation in the Arab World is characterized by a hierarchical power relation between two distinct varieties of Arabic: Classical/Modern Standard Arabic (CA), perceived as superior and complex, and regional Arabic (RA) -comprising the various regional dialects- deemed non-standard, reduced, and more simplex (Ferguson 1959). Consequent to this ideological attitude, the intractable myth of CA superior complexity has prevented the emergence of systematic scientific approaches to measuring structural complexity objectively in Arabic. The aim of this contribution is to objectively investigate and measure the degree of complexity exhibited in the various dialects, in the hopes that discussion of complexity becomes linguistically inclusive and removed from ideological considerations.

Many definitions and measures have been proposed to assess the degree of linguistic complexity of a language compared to others. These include McWhorter’s (2001, 2007) criteria of overspecification and structural elaboration, Dahl (2004) Length of Description principle, Miestomo’s (2008) fewer distinction and One-Meaning– One-Form principles, among others. Most of these principles state that more overt markers and more transformation rules mediating between deep and surface structures

entail more complexity. This presentation challenges this assumption, by showing that the grammar of RA is in some aspects equally if not more complex than that of CA, concluding that the difference between CA and RA is not always one of subtraction but often of addition or elaboration.

I will examine the contribution of grammaticalization and degrammaticalization processes to the structural complexity of RA. Grammaticalization is a process whereby lexical items gradually turn into grammatical categories, such as auxiliaries, pronouns, or even inflectional affixes (Norde 2019). The presentation will demonstrate that, while Arabic varieties may have shed a number of overt category markers of CA, e.g., case, nominal complementizers, feminine plural, and dual, they have developed innovative grammatical distinctions from lexical categories which led to the expansion of the grammatical system of the dialects; e.g., verb to auxiliary or noun to modal. I will then discuss the effect of degrammaticalization, where a grammatical marker or function word, gains autonomy or substance at the phonological, morphological, semantic or syntactic level (Norde 2009). Degrammaticalization in Arabic can be exemplified by the evolution of prepositions into possessives, modals and existential, and particles into inflected auxiliaries.

Moreover, it will be shown that while CA might outrank RA in terms of the number of overt distinctions, RA is more complex than CA in terms of rule opacity, which derives one-to-many relationships between form and function. Reductive morphosyntactic and morphophonological processes in RA render rule environments increasingly opaque, contributing to grammatical complexity, as evident in negation and verb allomorphy. Thus, we witness a trade-off between CA's Opacity of form and RA's opacity of process.

Furthermore, some syntactic distinctions expressed analytically in CA are expressed synthetically in the dialects, rendering the RA verbal system more complex than its CA counterpart; e.g., CA's syntactic dative /negative constructions are synthesized morphologically in RA *ma-katab-t-ha-l-ha-š*.

Within the realm of phonology, Complexity is evident from the innovative syllable types and phonemicization of native and borrowed phones, as well as opacity of metrical structure.

Variation in Syrian Arabic
Rania Habib, Syracuse University
rhabib@syr.edu

Syrian Arabic is an umbrella term for many urban and rural varieties that are often in contact. In this paper, I analyze the findings of a number of phonological variables that are affected by the spread of urbanization to rural areas in Syria. Then, I compare these findings to those of certain discourse variables. The data are derived from the naturally occurring speech of 50 children and adolescents and 22 parents. In most of these variations, we observe age and gender differences, but these differences are not always unidirectional; they may differ from one variable to another and from one generation to another, which requires future exploration.

In the phonological variables (q), (o), (o:), (e), and (e:), we see competing linguistic changes. While there are changes in progress towards the urban variants, local identity plays a significant role in reversing these changes among males in rural areas. Boys and girls follow the same patterns of linguistic use of their fathers and mothers respectively, but do not faithfully acquire the frequencies present in their parents' speech, a strong indication that developmental and socio-psychological factors are playing a role in their linguistic behavior. Gender emerged as significant for both the older and younger generations regarding the variable (q) and for the younger generation regarding the vowel variables. Although gender emerged as statistically insignificant for the older speakers regarding the vowel variables, the

boys and girls show similar linguistic pattern to that of their fathers and mothers respectively, further indicating acquisition of linguistic pattern, but not frequencies.

In discourse variables, we see competing uses of the discourse markers *yaʕni* and *ʔinnu*: between males and females within the different age groups and between the older and younger generations. These competing uses reflect a reversed linguistic pattern between the children and their parents. This reversed pattern is intriguing and implies, in contrast to the phonological variables above, that these children are not acquiring their parents' linguistic pattern and are advancing the use of *ʔinnu*: over *yaʕni* that is preferred by the older generation. In other words, not only are these children not acquiring the frequencies of their parents faithfully, but they are also following a different pattern of linguistic use from their parents. Furthermore, gender emerged as statistically significant among parents, but not among children, indicating that the younger generation, both boys and girls, prefers a discourse marker that is disfavored by their parents. The much higher use of *ʔinnu*: by the younger generation compared to their parents indicates a change in progress towards *ʔinnu*:. This could be comparable to the increasing use of *like* by the younger generation in English, which is disliked by the older generation.

Further research is required on these and other variables to discover patterns of linguistic use in Arab society. It is my hope that this ongoing work will give a more comprehensive picture of what is going on linguistically in Arab areas that are highly influenced by contact and urbanization. It is also my hope that such work will encourage other researchers to conduct similar research in other parts of the Arab world, so we can understand the linguistic behavior of different parts of Syria and the Arab world and the trajectories of the different varieties that exit in the entire region.

**The Dynamic of the Sound System from Proto-Arabic to the Modern Dialects:
Between Retention and Innovation**

Mohamed Embarki, Université de Franche-Comté, Besançon
mohamed.embarki@univ-fcomte.fr

Arabic is the youngest language among the languages of the Semitic family. However, it is described by many Semitists as the most conservative language in the family. From a phonological point of view, it is the one that has preserved the most phonological features of Proto-Semitic. It has maintained a rich consonant system, composed of 28-29 phonemes, and a more reduced vowel system, composed of three phonemic vowels contrasting with duration and opposing three short vowels and three long vowels. This brief phonological overview should not hide successive phonetic developments and adaptations over two millennia.

The Arabic speaking area currently knows a duality between a standard variety and a dialectal variety, and this on the scale of each country, or even within each region of the same country. This duality is not new; we can find it at each stage of the Arabic history. If the specialists agree to note the current differences between on the one hand Standard Arabic and on the other hand the Dialectal Arabic, the origin of the separation is not unanimously accepted. Some have explained these differences within the framework of the PAC (pidginization and creolization) which occurred in military camps during the Arab conquests after the advent of Islam. Whereas, some research rather favor the existence of pre-Islamic Arabic varieties which have stood alongside the *fusha* and which nowadays continue alongside the standard variety.

The reconstruction work raises the question of a common ancestor of the Standard and the Dialectal Arabic varieties (Old Arabic or Proto-Arabic) vs. different ancestors from different branches of

Semitic. However, the indisputable phonological evidence remains limited and the phonetic proofs are impossible to find.

Meanwhile, studies in synchrony on the way in which phonological systems are daily implemented by Arabic speakers provide some interesting insights showing a skillful balance, which has always prevailed in the history of Arabic varieties, between a relative retention of inherited features and controlled areal innovation and diffusion.

Abstracts of paper presentations (ordered according to their sequence in the schedule)

Deriving null objects in Egyptian Arabic: Evidence for Argument Ellipsis

Usama Soltan, Middlebury College

usoltan@middlebury.edu

Deriving null objects in Egyptian Arabic: Evidence for Argument Ellipsis

Usama Soltan (Middlebury College, usoltan@middlebury.edu)

Introduction. Egyptian Arabic (EA) allows **null objects** (NOs) with strongly transitive verbs, but only when the antecedent of the NO is an indefinite nominal, as illustrated by the contrast between (1) and (2), where Δ marks a NO coindexed with an antecedent.

This paper provides empirical evidence that NOs in EA are not instances of a null *pro* or variables A'-bound by a topic operator. Similarly, an analysis of NOs as derived via Verb-Stranding VP-Ellipsis is also shown to be empirically problematic. Rather, I argue that NOs are derived via **Argument Ellipsis (AE)**, an operation that targets arguments for deletion at PF under certain grammatical conditions. The AE analysis is then shown to account for the distribution of NOs in the language as well as a wide range of further empirical facts.

Three potential analyses of NOs and challenges to each.

First analysis: NOs as null *pro*: [antecedent_clause ... **Obj_i**] [NO_clause ... **pro_i**]. This analysis wrongly predicts that NOs can take definite nominals as antecedents, contrary to the ungrammaticality of (2), where a pronominal clitic is required in object position. Furthermore, this analysis cannot readily explain why NOs, unlike pronouns, are associated with a **different-entity interpretation**: (1), for example, has an interpretation where two different books were found, as opposed to the object clitic in (2), which has a same-entity interpretation.

Second analysis: NOs as variables A'-bound by a topic (Huang 1984): [[Topic_i] ... [NO_i]].

This analysis is challenged by the contrast in (1-2), and the definiteness requirement on Arabic topics (3-4) (Aoun et al 2010).

Also, NOs in EA can occur inside islands, as in (5b), in the presence of a linguistic (not pragmatic) antecedent (5a), indicating that such NOs are cases of *surface*, not deep, *anaphora* (Hankamer and Sag 1976).

1. Mona laʔ-it **kitāb_i** wi-Huda kamān **laʔ-it** Δ_i
Mona found-SGF book and-Huda also found-SGF
'Mona found a book, and Huda did too.'
2. Mona laʔ-it **?il-kitāb_i** wi-Huda kamān **laʔ-it-*(u_i)**
Mona found-SGF the-book and-Huda also found-SGF-it
'Mona found the book, and Huda found it too.'

3. **?il-kitāb_i** Mona laʔ-it-**u_i** 4. ***kitāb_i** Mona laʔ-it-**u_i**
the-book Mona found-SGF-it book Mona found-SGF-it
'The book, Mona found it.' 'A book, Mona found it.'

- 5a. yizhar ?inn Mona laʔ-it **waziifa_i**
seem.SGM that Mona found-SGF job
'It seems that Mona found a job.'

- 5b. ?aywa wi-fii [island ?išāʔa ?inn Huda kamān laʔ-it Δ_i]
yes and-there rumor that Huda also found-SGF
'Yes, and there is a rumor that Huda did too.'

Third analysis: NOs via Verb-Stranding VP-Ellipsis (VSVPE): [ellipsis_clause [TP T+V_j+V_i [V_P t_j [VP-~~t_j~~-**Obj**]]]] (Huang 1991, McCloskey 1991, Goldberg 2005). Under this analysis, V raises out of VP, followed by VP-deletion.

A VSVPE analysis is challenged by the fact that EA does not have Aux-Stranding VP-ellipsis, as shown in (6).

Also, if VP-ellipsis were involved, we expect all VP-internal material to get deleted, contrary to (7), where an indirect object surfaces after a NO.

6. Ahmad kān bi-yilʕab koora wi-Ali kamān **(*kān)**
Ahmad was.3SGM ASP-play.3SGM soccer and-Ali also was.3SGM
'Ahmad was playing soccer, and Ali too.'
7. Mona ?idd-at **filūs_i** li-Ali wi Huda ?idd-at Δ_i **li-Hasan** kamān
Mona gave-SGF money-to-Ali and Huda gave-SGF to-Hasan too
'Did Mona give money to Ali?'

Proposed analysis: NOs are elided arguments: [antecedent_clause ... [VP **Obj_i**]] ... [ellipsis_clause ... [VP **Obj_i**]].

In this paper, I argue that NOs are derived via **Argument Ellipsis (AE)**, whereby an indefinite object is PF-deleted under identity with an antecedent. This has immediate empirical advantages. **First**, it explains the different-entity reading of NOs (1). **Second**, it explains why VP-internal material such as an indirect object can follow a NO (7). **Third**, it is not surprising that NOs can occur inside islands (5b). **Fourth**, the (in)definiteness-asymmetry for NOs can be accounted for in terms of variance in case-licensing mechanisms: structural case for definite nominals and inherent case for indefinite nominals (Saito 2007, Cheng 2013). This, in turn, has a **fifth** empirical consequence: We predict that PP and CP arguments, not needing structural case, can undergo AE, which is indeed the case, as shown by (8), which has a strict as well as a sloppy identity reading, a property often associated with ellipsis phenomena.

8. Eman baʕat-it **filūs** [PP li-ʔahla-hā]_i wi Huda baʕat-it **hadāyā** Δ_i

Eman sent-3SGF money to-family-her and Huda sent-3SGF gifts

'Eman_i sent money to her family, and Huda_j sent gifts [to her_{i/j} family].' (✓Strict; ✓Sloppy)

Conclusion. In sum, NOs in EA follow from an Argument Ellipsis operation that targets indefinite objects at PF. Such an analysis overcomes the challenges faced by alternative analyses and is shown to be empirically superior.

Selected references

- Aoun, Joseph, Benmamoun, Elabbas, and Choueiri, Lina. 2010. *The Syntax of Arabic*. Cambridge: Cambridge University Press.
- Cheng, Hsu-Te Johnny. 2013. *Argument Ellipsis, Classifier Phrases, and the DP Parameter*. PhD Dissertation, University of Connecticut, Storrs.
- Goldberg, Lotus. 2005. *Verb-Stranding VP Ellipsis: A Cross-linguistic Study*. PhD Dissertation, McGill University.
- Hankamer, Jorge, and Sag, Ivan. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7, 391–428.
- Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 531–574.
- Huang, C.-T. James. 1991. Remarks on the status of the null object. In Robert Freidin (ed.), *Principles and Parameters in Comparative Grammar*, 56–76. Cambridge, MA: MIT Press.
- McCloskey, James. 1991. Clause structure, ellipsis, and proper government in Irish. *Lingua* 85, 259–302.
- Saito, Mamoru. 2007. Notes on East Asian argument ellipsis. *Language Research* 43, 203–227.

The correlation between verb movement and tense in Arabic

Hamid Ouali, University of Wisconsin-Milwaukee

ouali@uwm.edu

It is well-established that Arabic is a verb-raising language (Ouhalla 1994). It has also been shown that the main perfective and the imperfective verbs behave differently in simple declarative clauses when it comes to their target of movement. Benmamoun (2000), Soltan (2007), and recently Ouali (2018), among others, argue that what sets past tense clauses and present tense clauses apart is V-to-T for past and lack of V-to-T for present. According to this literature then, there is a link between perfective verb, past tense, and V-to-T. The imperfective verb in present tense contexts moves only to Asp (sultan 2007). In this paper, I revisit these claims, argue that the link between V-to-T and past tense is not entirely accurate and does not always hold. I will discuss a set of stative verbs, change of state verbs, and achievement verbs, which when used in the perfective form, express states that hold true in the present tense. I will then propose that V-to-T is not linked to past. My analysis has wide implications on verb-movement and the syntax of tense and aspect in Arabic in general. The proposal is schematized below:

- (1) a. $T_{[\pm PAST]} - Asp_{Prog/Hab} - \text{imperfective } V$ (eventive verbs)
- b. $T_{[\pm PAST]} - Asp_{PFV} - \text{perfective } V$ (eventive verbs)
- c. $T_{[\pm PAST]} - \text{perfective } V$ (stative, change of state, and achievement verbs.)

By showing that verb morphology does not mark tense in Arabic, Benmamoun (2000), Aoun et al. (2010), and Ouali (2018), argue that tense correlates with V-movement. Past tense clauses involve V-to-T and present tense clauses do not; consequently, perfective verb forms are syntactically derived by V-to-T (Suffixed-V (2-4)), and imperfective verbs are derived by V-to-Asp (Prefixed-V (5-7)):

- | | | |
|-------------------------|-------------------------|-------------------------|
| (2) xaraj-at (SA) | (3) xrtj-at (MA) | (4) xaraj-it (JA) |
| left. PFV-3SF | left. PFV-3SF | left. PFV-3SF |
| ‘She left/went out.’ | ‘She left/went out.’ | ‘She smiled.’ |
| (5) ta-btasim (SA) | (6) ka t-btasm (MA) | (7) b ti-btisim (JA) |
| 3SF-smile.IPRF | ka 3SF-smile.IPRF | bi 3SF-smile. IPRF |
| ‘she is smiling/smiles’ | ‘she is smiling/smiles’ | ‘she is smiling/smiles’ |

Benmamoun (2000) argues that: “in the past tense the verb: (a) must merge with negation (b) prefers (for some speakers) to precede the subject in sentences and idioms, and (c) has person agreement prefix”. The correlation between perfective morphology and past tense does not always hold. Stative verbs such as *understand* as in (8-9), the change of state verbs such as *red* as in (10-11), and achievement verbs such as *recognize* as in the question-answer pairs in (12) in SA and MA respectively are in the perfective form and express states that hold true in the present as shown by the adverb *now*; in fact the imperfective form yields infelicitous readings in some cases (13).

- | | | | |
|----------------------------------|------------------------------------|-----------------------------------|------------------|
| (8) fahim-tu l-ʔaana (SA) | (9) fhəmt daba (MA) | (10) l-ʔaana, hmarr-at (SA) | (11) hmar-t (MA) |
| understand.PFV-1S now | understand.PFV-1S now | now redder.PFV-3SF | redder.PFV-3SF |
| ‘I understand now.’ | ‘I understand now.’ | ‘now, it is red/ it’s become red’ | |
| (12) Q: hal ʕarif-ta-ni l-ʔaana? | A: naʕam, ʕarif-tu-ka l-ʔaana (SA) | | |
| waf ʕraf.PFV-ti-ni daba? | ʔah, ʕraf.PFV-t-k daba (MA) | | |
| Q recognize-2SM-me now | yes, recognize-3S-you now | | |

- ‘do you recognize me now?’ ‘Yes, I recognize you now.’
 (13) # ka n-fhəm
 ka-lsm-understand.IPRF
 ‘# I understand’ (The only possible reading here is the habitual reading like: I usually understand)

These stative, change of state, and achievement verbs can also merge with negation (14), precede the subject (15), and have person agreement prefix in the present tense (14-15). Idiomatic and optative expressions, I argue, cannot be used as a diagnostic for past tense.

- (14) mazal ma-fhəm-t-ʃ aʃ ka-t-gul (15) ʔrəf-t had r-raʒəl
 still neg-understand-1SM-neg what ka-2SM-say know.PFV-1s this the-man
 ‘I still don’t understand what you’re saying’ ‘I know this man.’

In some Arabic dialects, the active participle of the verbs in (14) and (15) is used instead.

I conclude that the correlation then is not between the past tense and the perfective verb form, but rather it is between the perfective verb form and V-to-T, regardless of the tense value of T (present or past). As schematized in (1), in sentences with eventive verbs, Asp, which is the locus of grammatical aspect (Hallman 2016) is selected by T. The Progressive/Habitual Asp is selected by T that can bear either [+Past] or [-Past]. The PFV Asp can only be under the scope of T[+Past]. Sentences with stative verbs, change of state verbs, and achievement verbs do not involve Asp.

On the Realization of the Glottal Stop in Sudanese Arabic

Abdel-Khalig Ali, University of Toronto

abdelkhalig.ali@utorontoca

The literature on the morphophonology of Classical Arabic (CA) and, by extension, Modern Standard Arabic (MSA) describes intricate patterns of realization of the glottal stop that can generally be explained in terms of the phonotactic constraints in these dialects. A cursory look at contemporary spoken dialects of Arabic reveals equally interesting patterns. Yet, systematic studies of this phenomenon in the latter group of dialects are lacking. In this talk, I examine the realization patterns of the glottal stop attested in Urban Central Sudanese Arabic (UCSA). To this end, I examine derivations of forms whose roots consist of glottal stops in initial, medial, and final positions, and I identify the phonotactic constraints responsible for the realization patterns. Due to space limitation, the patterns illustrated here are of initial and final positions.

The verb patterns in (1) show that the root-initial glottal stop is realized in forms where it can be syllabified as an onset. However, when this segment occurs in contexts where it must be syllabified as a coda, it is deleted and the vowel preceding it is lengthened. This is illustrated by the verb patterns in (2).

- | | | | | | | |
|-----|-------|---|-------------|---|------------|-------------|
| (1) | Root | Pattern | Input | → | Output | Gloss |
| a. | ʔ-k-l | I C ₁ aC ₂ aC ₃ | /ʔakal/ | → | ʔa.kal | ‘he ate’ |
| b. | ʔ-k-d | II -C ₁ aC ₂ C ₂ iC ₃ | /na-ʔakkid/ | → | na.ʔak.kid | ‘we affirm’ |
-
- | | | | | | | | |
|-----|----|-------|---|--------------|---|-------------|----------------------|
| (2) | a. | ʔ-k-l | I -C ₁ aC ₂ uC ₃ | /na-ʔkul/ | → | naa.kul | ‘we eat’ |
| | b. | ʔ-z-n | X -staC ₁ C ₂ aC ₃ | /na-staʔzan/ | → | nas.taa.zan | ‘we seek permission’ |

This segment exhibits two additional peculiar patterns. First, in a limited number of roots the glottal stop is not realized even when it could be syllabified as an onset. This occurs in (3a&b), where the glottal stop is deleted and the preceding consonant geminates, but not in (3c).

- | | | | | | | |
|-----|-------|---|-----------|---|-------------|----------------|
| (3) | Root | Pattern | Input | → | Output | Gloss |
| a. | ʔ-k-l | VII nC ₁ aC ₂ aC ₃ | /nʔakal/ | → | ʔin.na.kal | ‘it was eaten’ |
| b. | ʔ-k-l | VI tC ₁ aaC ₂ aC ₃ | /tʔaakal/ | → | ʔit.taa.kal | ‘it eroded’ |
| c. | ʔ-m-r | VI tC ₁ aaC ₂ aC ₃ | /tʔaamar/ | → | ʔit.ʔaa.mar | ‘he conspired’ |

Second, this segment patterns like an epenthetic one with respect to phrasal syllabification. The glottal stop in the verb /ʔakal/ in (4a) is not realized at the phrase level. Instead, the preceding consonant is

syllabified as an onset across the word boundary. This is the same syllabification pattern of the definite article in the second word in (4b).

- (4) a. /al-walad/?akal/ → (?al.wa.la.d)(a.kal) b. /al-walad/al-kabiir/ →
 (?al.wa.la.d)(al.ka.biir)
 the-boy ate cf. *(?al.wa.lad.)(?a.kal) the boy the-big
 ‘The boy ate.’ ‘the big boy’

The realization pattern of the root-final glottal stop is peculiar in that this segment is categorically absent in the surface forms in UCSA. I illustrate this point with a partial paradigm of verbs derived from the root *b-d-ʔ* ‘having to do with beginning.’ The table in (5) shows the perfect and imperfect indicative forms of three verbs in CA/MSA and the corresponding forms in UCSA. The root-final glottal stop is absent in the latter even in contexts where it could be syllabified as an onset (5b Perfect).

(5)

Subject	Perfect		Imperfect Indicative		Gloss
	CA	UCSA	CA	UCSA	
a. 3MS	badaʔ-a	bada	ya-bdaʔ-u	yi-bda	‘he began/begins’
b. 3FS	badaʔ-at	bad-at	ta-bdaʔ-u	ta-bda	‘she began/begins’
c. 1P	badaʔ-na	badee-na	na-bdaʔ-u	na-bda	‘we began/begin’

The glottal stop is absent throughout the paradigm in all verb patterns and in the active and passive participle forms. Instead, the third consonant of the root is consistently realized as a palatal glide in the verbal noun forms (e.g., *bidaaya* ‘beginning’). Based on the latter pattern, I argue that the final glottal stop in these forms has been reanalyzed as a palatal glide *y* in UCSA. I also argue that the two peculiar patterns of the root-initial glottal stop in (3) & (4) are indicative of a similar reanalysis process.

Some Properties of the m- Definite Article in Upper Faifi Arabic
Abdullah Alfaifi (abdalfai@iu.edu) & Stuart Davis (davis@indiana.edu)
Indiana University

The definite article in Arabic dialects appears in different forms. In Modern Standard Arabic (MSA) and most Arabic dialects, the definite article is *l-* (or *al-*, *il-*, *lə-*), which has multiple assimilatory realizations which reflect a following coronal consonant, or alternatively, true coronal geminated realizations (Zaborski, 2006 for the assimilatory view; and Heselwood & Watson, 2015 for the alternative view). Some Arabic dialects exhibit other forms of the definite article that are often described as foreign to Arabic. These forms include *b-*, *n-*, and *m-* and appear in Arabic varieties in the southern part of the Arabian Peninsula (Rabin, 1951; Alqahtani, 2015; Behnstedt, 2016). This paper focuses mainly on the *m-* realization of the definite article in Upper Faifi Arabic (UFA), a dialect spoken in Faifa Mountains in the southwest of Saudi Arabia, and native to one of the co-authors. The *m-* definite article in UFA attaches to most nouns, including some proper nouns such as [yahya m-saalim] for ‘*Yahya Salem*’ (Alfaife, 2018) and adjectives [mbint imśabiira] ‘*the big girl*’. While Standard Arabic and most Arabic dialects have the definite article prefix /*l-*/, the definite article in UFA is almost exclusively prefixal /*m-*/ . As found in other dialects in the region, though, UFA does have the *l-* definite article but in limited usage. While UFA does possess both *l-* and *m-* as definite articles, there are certain semantic and syntactic conditions that determine the occurrence of either definite article. This paper examines a range of issues regarding the definite articles in UFA.

Phonologically, the *m-* definite article in UFA can sometimes contribute to forming word-initial biconsonantal clusters. For example, MSA [ʔaʃ-ʃams] ‘*the sun*’ and [ʔal-qa.mar] ‘*the moon*’ are respectively realized as [m-ʃams] and [m-ga.mar] in UFA (period indicates a syllable boundary). If the base word in UFA begins with a consonant cluster, the sequence [ʔi-] is inserted before the prefixal definite *m-* to avoid a word-initial triconsonantal cluster: MSA [ʔasʕ-sʕa.dir] ‘*the chest*’ and [ʔal-bi.laad] ‘*the land*’ are realized as [ʔim-sta.dir] and [ʔim-bi.laad] in UFA. The *m-* definite article can

additionally create initial [mm] geminates when prefixed to an m-initial word: UFA [m-miʃtʰ] for MSA [ʔal-miʃtʰ] ‘the comb’.

The use of the m- definite article (as opposed to the l- definite article) for most words in UFA makes it different from other varieties of southwestern Saudi Arabic (SSA) where there is optional variation on the same word between the use of the l- definite article and the m-definite article (see Alqahtani 2015 and Alahmari 2015 for two different accounts of this variation in two different (non-Faifi) varieties of SSA.) Nonetheless, for semantic reasons, the m-definite article is blocked in words with religious significance, the names of countries, days of the week, months, and names of non-local cities. For example, the MSA word [ʔal-ma.dii.na], referring to *Al-Madīnah Al-Munawwarah*, is realized similarly in UFA, but when the same word is used to mean ‘the city’, the UFA realization is [m-ma.dii.na]. Further, the days of the week take the l- definite article as in [yawm is-sabt] ‘Saturday’. While the names of the days appear in the *Idāfa* construction, the *Idāfa* normally takes the m- definite article as in [baab im-bajt] ‘door of the house’.

Further, as observed by Alfaifi & Behnstedt (2010: 60), the m- definite article is typically not used with a demonstrative in Faifi Arabic. While this observation is accurate for a pre-posed demonstrative as in (a), the m- definite article can be used with a post-posed definite article as in [m-marah ti] ‘this woman’, which has the same meaning as (a). Moreover, if the prefixal [m-] definite article were to be used with a demonstrative as in (b) then the expression would be understood as a complete sentence rather than as a noun phrase. This thus distinguishes between an inherent ambiguity that can be found in other dialects with respect to [til-marah] as either ‘this woman’ or ‘this is the woman’. This paper will examine these and other issues regarding the realization of the m-definite article in Upper Faifi Arabic.

- | | | | | | | | |
|----|------|-----|---------|---------------------|------|-----|--------|
| a. | [ti | l- | marah] | ‘this woman’ | *[ti | m- | marah] |
| | DEM | DEF | woman | | DEM | DEF | woman |
| b. | [ti | m- | marah] | ‘this is the woman’ | *[ti | l- | marah] |
| | DEM | DEF | woman | | DEM | DEF | woman |

References

- Alahmari, M. 2015. The development of the m- definite article in Arabic: A historical linguistic perspective. Unpublished Manuscript. Indiana University Bloomington.
- Alfaifi, A. 2016. *Aspects of the Morpho-syntax of the Faifa Dialect of Saudi Arabia: A Descriptive Study*. PhD thesis, Ulster University.
- Alfaifi, S. 2018. *A Grammar of Faifi*. Long Beach, CA: California State University, Long Beach MA thesis.
- Alfaifi, A. & Peter Behnstedt. 2010. First notes on the dialect of Ġabal Fayfā’ (Jazan province/Saudi Arabia). *Zeitschrift Für Arabische Linguistik* 52. 53–67.
- Alqahtani, K. 2015. *A Sociolinguistic Study of the Tihami Qahtani Dialect in Asir, Southern Arabia*. PhD thesis, University of Essex.
- Behnstedt, P. 2016. *Dialect Atlas of North Yemen and Adjacent Areas*. Leiden: Brill.
- Heselwood, B. & Janet Watson. 2015. The Arabic definite article: A synchronic and historical perspective. *Arabic and Semitic Linguistics Contextualized: A Festschrift for Jan Retsö*, 157–176.
- Rabin, C. 1951. *Ancient West Arabian*. London: Taylor's Foreign Press.
- Zaborski, A. 2006. Article, definite. In Kees Versteegh, Mushira Eid Alaa Elgibaly, Manfred Woidich, & Andrzej Zaborski (eds), *Encyclopedia of Arabic Language and Linguistics*, vol. 1, 187–188. Boston, MA: Brill.

Optionality in Avoidance of CCC Clusters in Qassimi Arabic

Ahmed Alnuqaydan, University of Utah

u1215249@utah.edu

Many Arabic varieties use epenthesis to break up triconsonantal clusters. Onset “CV” dialects such as Egyptian Arabic epenthesize a vowel after the second consonant (CCVC) while rime “VC” dialects

such as Iraqi Arabic break up the illicit cluster by inserting a vowel after the first consonant (CVCC) (Broselow, 1992; Selkirk, 1981; Kiparsky, 2003). On the other hand, Qassimi Arabic (QA) follows both patterns. As exemplified by the underlined vowels in (1), old QA speakers (age: above 65) show variation between these two patterns in avoidance of morpheme-internal -CCC- in trisyllabic verbs.

(1)

- /ʃbk/ [i.ʃəb.kuh] (CVCC) ~ [iʃ.bə.kuh] (CCVC) “plug it in”
- /ʃlh/ [ja.ʃəl.ħuh] (CVCC) ~ [jaʃ.lə.ħuh] (CCVC) “may god guide him”
- /dʰhl/ [ma.dʰəh.luh] (CVCC) ~ [madʰ.ħə.luh] (CCVC) “mix it up”

QA optionality, I argue, is best accounted for by a partial order grammar (Antilla, 1997) involving an interaction between stress and epenthesis. This analysis both produces the optionality in (1) and improves on other analyses by providing an independent justification for epenthesis.

QA has a default-to-opposite (DTO) stress system where stress falls on the rightmost heavy syllable while in the absence of heavy syllables, stress falls on the initial syllable (Ingham, 1982). CVVC, CVCC and CVV are heavy both finally and non-finally whereas CVC is only heavy non-finally. Bakovic’s (1998) analysis of DTO main stress, where Parse- σ is dominated by WSP and FTBIN (at the moraic level) and HDFT-R outranks HDFT-L, only predicts one surface form [ja.(ʃə́l).ħuh] but not the other possible form [(jaʃ).lə.ħuh] (2). Ranking WSP higher than Parse- σ ensures that heavy syllables are footed and stressed whereas Parse- σ is dominated by FTBIN to ensure that only bimoraic syllables are parsed into feet. HDFT-R and HDFT-L are gradient head-alignment constraints where each syllable separating the designated edge of the head foot from the designated edge of the word incurs a violation. Ranking HDFT-R above HDFT-L requires main stress to fall on the rightmost available foot which is defined by its interaction with other constraints. In (2), [ja.(ʃə́l).ħuh] wins because it satisfies HDFT-R better than the other possible output [(jaʃ).lə.ħuh]; this means, though, that an epenthetic vowel is stressed, something languages often avoid (Kager, 1999). Stressed epenthetic vowels violate Kager’s HEADDEP. The other variant [(jaʃ).lə.ħuh] avoids stressing an epenthetic vowel while still stressing the rightmost heavy syllable; the cost, though, is a new violation of HDFT-R because that rightmost heavy syllable is farther from the right edge.

Given that, the two optional forms can be accounted for by a partial order grammar where the ranking between the two competing constraints HDFT-L & HEADDEP is not specified as illustrated in (3) where each output is possible depending on how the partial ranking is resolved.

Unlike older speakers, younger speakers show no variation and produce only [jaʃ.lə.ħuh] (CCVC). This can be explained if the ranking HEADDEP >> HDFT-R is invariant in their grammar.

This analysis also accounts for other dialects with fixed epenthesis sites: Egyptian Arabic has the invariant ranking HEADDEP >> HDFT-R, and Iraqi Arabic has the opposite. It is superior to previous analyses (Farwanah, 1996; Ito, 1989; Rose, 2000) as it explains what motivates each epenthesis site rather than merely stipulating where epenthetic vowels go.

(2)

jə.ʃlhuh	WSP	FTBIN	Parse- σ	HDFT-R	HDFT-L
→ ja.(ʃə́l).ħuh			**	*	*
? (jaʃ).lə.ħuh			**	**!	
(ja).(ʃə́l).(ħuh)		*!*		*	*
(jaʃ).(lə).(ħuh)		*!*		**	
(ja).ʃə́l.(ħuh)	*!	**	*		
jaʃ.(lə).ħuh	*!		*		*
ja.(ʃə́l).ħuh		*!	*		*
(jaʃ).lə.ħuh		*!	*	*	

3-Stress and Epenthesis Interaction; Optionality/ Partial Order Grammar

jə.ʃlhuh	WSP	FTBIN	Parse- σ	HDFT- R	HEADDEP	HDFT-L
→ ja.(ʃəl).huh			**	*	*(!)	*
→ (jaʃ).lə.huh			**	**(!)		

References

- Anttila, A. (1997). Deriving variation from grammar. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 35-68.
- Bakovic, E. (1998). Unbounded stress and factorial typology. Ms. Rutgers University. [available on Rutgers Optimality Archive, ROA-244, <http://roa.rutgers.edu/files/244-0298/roa-244-bakovic-2.pdf>]
- Broselow, E. (1992). Parametric variation in Arabic dialect phonology. In E. Broselow, M. Eid, & J. McCarthy (Eds.), *Perspectives on Arabic linguistics* (Vol. IV, pp. 7-45). Amsterdam, The Netherlands: John Benjamins.
- Farwaneh, S. (1996). Directionality effects in Arabic dialect syllable structure. (Doctoral dissertation). University of Utah, Salt Lake City, UT.
- Ingham, B. (1982). *North east Arabian dialects*. Routledge.
- Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language & Linguistic Theory*, 7(2), 217-259.
- Kager, R. (1999). Surface opacity of metrical structure in Optimality Theory. B. Hermans and M. van Oostendorp, (eds). *The Derivational Residue in Phonology*. Amsterdam: John Benjamins, 207-245.
- Kiparsky, P. (2003). Syllables and moras in Arabic. In C. Fery & R. van de Vijver (Eds.), *The syllable in optimality theory* (pp. 147-182). Cambridge, United Kingdom: Cambridge University Press.
- Rose, S. (2000). Epenthesis positioning and syllable contact in chaha. *Phonology*, 17(3):397-425.
- Selkirk, E. (1981). Epenthesis and degenerate syllables in Cairene Arabic. In H. Borer & J. Aoun (Eds.), *Theoretical issues in the grammar of Semitic languages: MIT working papers in linguistics* (3, pp. 209-232). Cambridge, MA: Department of Linguistics and Philosophy, MIT.

ArabiLight: A Corpus-Based Model for Spoken Modern Standard Arabic

Khaled Elghamry
Ain Shams University
kelghamr@indiana.edu

Attia Youseif
Indiana University
ayouseif@iu.edu

One of the recommendations for the conversational agents being currently developed for Arabic is to use a language variant that sounds educated yet friendly and natural. Modern Standard Arabic (MSA) has been adopted so far as the initial candidate for such task. One of the issues here is how to handle word-final vowels in spoken MSA. There are two extremes in this respect: [a] to drop these vowels across the board, and [b] to pronounce these vowels regardless of the context. Building on previous studies (e.g., Hallberg 2016 and Bassiouney 2010), this paper argues that real-world corpora do not support any of these two scenarios. The contribution of this paper is two-fold. It presents a corpus-based analysis of the realization of word-final vowels. Using these results, it proposes and describes *ArabiLight*, a conversational-friendly model for spoken MSA. This model is based on what we call the “Sufficient Necessary Minimum”. The authors also discuss possible applications of this model in learning and teaching Arabic.

The corpus was collected from: [a] official news channels from all Arab countries, [b] channels covering pan-Arab issues such as Aljazeera, AlArabiya and Skynews, [c] and the Arabic-speaking versions of France 24, CNN, BBC, Euronews and CNBC. The total sample contains 224 minutes, divided equally between male and female news anchors.

The sample was manually transcribed by the authors. Each word was labelled for the presence or absence of a word-final vowel that mark the mood or case of such word, given its context. Long and short pauses were also marked. Each news piece was labelled with the anchor's gender and nationality.

The preliminary results show that the full use of case endings seems to be an idealized scenario that exists only in the Arabic grammar textbooks. The suggestion of dropping case across the board seems hard to be realized in real-world corpora. For example, our analysis shows that Iraqi anchors pronounce word-final vowels the least with %42, and the Tunisians the highest with %61. The effect of topic and gender on this ratio is currently under investigation.

The transcribed corpus was statistically analyzed to identify the contexts where word-final vowels were consistently pronounced by the anchors from every source in the corpus. These include, for example, “tanween nasb”, the vowel marking case or mood in words immediately followed by “alif wasl”, and the final vowels in perfective verbs. The authors argue that these contexts represent the necessary minimum of pronouncing these vowels that is sufficient for the realization of a spoken MSA that sounds educated yet friendly and natural. We call it *Arabilight*.

The paper also discusses the advantages of using this model in teaching and learning spoken MSA that sounds educated, yet friendly and natural. In addition, it could also help learners and teachers focus mainly on the necessary minimum of the grammar rules for word-final vowel realization, and hence reduce the learner's ‘fear’ of the grammar, and consequently increasing their motivation, and the teacher's time teaching an idealized textbook grammar model.

This model is part of an ongoing project concerned with the corpus-based construction of resources and models for Arabic that also includes *Mashhoor*, a learner's corpus-based familiarity dictionary for Arabic, and *Darajaat*, a graded dictionary for Arabic learning and teaching. We plan to improve our model using more data from other MSA sources.

References

- Bassiouny, R. (2010). *Arabic and the media: linguistic analyses and applications*. Brill.
Hallberg, A. (2016). *Case Endings in Spoken Standard Arabic*. Lund University, Faculties of Humanities and Theology.

An Emergent Grammar model for diglossic variations between Arabic and Faifi
Essa Alfaifi, University of Arizona & Jazan University
alfaifiessa@email.arizona.edu; alfaifiessa@gmail.com

The study of grammar in ‘diglossic speech communities’ in which two or more varieties of the same language are used in different contexts (Ferguson, 1959, 1991, etc.), raises important questions, such as: How do these varieties interact? Are there clear-cut boundaries between language varieties in these contexts? Can we posit that speakers, in certain circumstances, may possess more than one internal grammar of their language that is associated with different social contexts? If so, how do speakers develop the knowledge of this type of grammar? And, more importantly, how can we systematically account for the massive scale of variation that is part of any diglossic speech grammar?

Faifi, spoken in southwestern Saudi Arabia, represents an interesting case of language contact and diglossia in which Faifi and Arabic, closely-related linguistic systems, feed into a hybrid grammar that language speakers develop across time. Intense language contact and diglossia in Faifi have resulted in mixed language properties that can be divided into three categories with no clear-cut boundaries: (i) properties specific to Faifi, (ii) properties specific to Arabic and (iii) shared properties between the two. All three are illustrated in (1) with the distribution of the reverse-affricate [st̪] and the emphatic [sˤ]. Data, which consist of elicited forms from Faifi native speakers and a 10.000-word corpus of spontaneous Faifi speech, were gathered during a fieldwork in the community.

(1) *Distribution of [st̩] and [sʰ] in Upper Faiifi*

[st̩] Only		[sʰ] Only		Variation [st̩]/[sʰ]		
Faifi word	Gloss	Arabic Loan	Gloss	Faifi (Low)	Arabic (High)	Gloss
a. [st̩irfa]	‘rack’	f. [sʰu:ra]	‘picture’	k. [st̩ala]	[sʰala]	‘prayer’
b. [st̩awar-]	‘noise’	g. [sʰadaga]	‘charity’	l. [st̩aum-]	[sʰaum]	‘fasting’
c. [st̩afga]	‘hit’	h. [sʰa:ru:x]	‘missile’	m. [st̩abir-]	[sʰabir]	‘patience’
d. [ʃast̩al-]	‘muscle sore’	i. [fasʰl]	‘classroom’	n. [st̩afara]	[sʰafra]	‘yellow’
e. [ʃast̩i:m-]	‘skin’	j. [ʃasʰi:r]	‘juice’	o. [ʔast̩laʃ]	[ʔasʰlaʃ]	‘bold’

Evidence from the language shows that each of these three categories is governed by a set of learned regulatory principles (i.e. grammar) of its own. The study maintains that the grammar which governs the phonemic spilt patterns attested in Faiifi is *emergent*. This means that the hybrid grammar of Faiifi is derived from domain-general cognitive processes that are not specific to language including: memory, attention to frequencies and similarities, categorization, establishing connections between similar categories, making generalizations from the observed patters, etc.

The analyses are grounded on the core of a relatively new framework called *Emergent Grammar* (EG) which drifts away from using the concept of Universal Grammar and genetic endowment specific to language as central in driving research questions and the way of interpreting data (Archangeli, Mielke & Pulleyblank, 2012a, b; Archangeli & Pulleyblank, 2014, 2015, 2016, 2017, etc.). Instead, EG allows abstract symbolic grammatical structures to emerge from experience with the language, using the general cognitive capabilities of the language learner. This idea is not completely new to the field considering the long-standing emergentist approach to language (e.g., Hopper 1987, 1998; Bybee 1998, 1999; Lacerda, 2003; Ellis & Larsen-Freeman 2006). What is new and appealing about the EG framework is that it accounts for language through a set of bottom-up regulatory principles (i.e. emerging conditions and relations among morphemes) that filter out unwanted surface combinations and give an optimal output. The regulatory principles of EG are: Morph Sets, Morph Set Relations, Phonotactics, Morphotactics and Morphophonotactics among other mechanisms discussed in this study.

This study demonstrates how speakers build a symbolic system (grammar) in a bottom-up fashion from different speech contexts such as home, public spaces, school, formal and informal situations. When grammar is abstract, it governs word formation processes (i.e. what is allowed and disallowed regarding the linguistic output). The study also discusses why the hybrid grammar of Faiifi must be viewed as a constantly-updating system that is shaped by language experience. This view has implications for the study of language gradience and change, and how forms diffuse, compete and decay across time.

References

- Archangeli, D., Mielke, J. & Pulleyblank, D. (2012a). Greater than noise: Frequency effects in Bantu height harmony. In Botma Bert & Roland Noske (eds.), *Phonological Explorations. Empirical, Theoretical and Diachronic Issues* (pp. 191-222). Berlin/Boston: De Gruyter.
- Archangeli, D., Mielke, J. & Pulleyblank, D. (2012b). From sequence frequencies to conditions in bantu vowel harmony: Building a grammar from the ground up. *McGill Working Papers in Linguistics*, 22(1), 1-26.
- Archangeli, D. & Pulleyblank, D. (2014). Autosegments in Emergent Grammar. In Elaine Espindola & Abhishek Kumar Kashyap (Eds.), *Proceedings of the 7th International Free Linguistics Conference* (199–205), Hong Kong Polytechnic University. Free Linguistics Conference.
- Archangeli, D. & Pulleyblank, D. (2015). Phonology without universal grammar. *Frontiers in psychology*, 6, Article 1229.
- Archangeli, D. & Pulleyblank, D. (2016). Emergent morphology. In H. Harley, & D. Siddiqi (Eds.), *Morphological metatheory* (pp. 237-270). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Bybee, J. (1998). The emergent lexicon. In *Chicago Linguistic Society*, 34, 421–35.

- Bybee, J. (1999). Usage-based phonology. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, & K. Wheatley (Eds.), *Functionalism and Formalism in Linguistics* (pp.211–242). Amsterdam: John Benjamins.
- Ellis, N., and Larsen-Freeman, D. (2006). Language emergence: implications for applied linguistics—introduction to the special issue. *Applied linguistics*, 27(4): 558–89.
- Ferguson, C. (1959). Diglossia. *Word*, 15(2), 325–340.
- Ferguson, C. (1991). Diglossia revisited. *Southwest Journal of Linguistics*, 10(1), 214–234.
- Hopper, P. (1987). Emergent grammar. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 13, pp. 139–157).
- Hopper, P. (1998). Emergent grammar. In Michael Tomasello (Ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure* (155–175). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lacerda, F. (2003). Phonology: An emergent consequence of memory constraints and sensory input. *Reading and Writing*, 16(1–2), 41–59.

When the leak becomes a flood: Destabilizing diglossia in Tunisia

Karen McNeil, Georgetown University,

km1542@georgetown.edu

In 2003, Keith Walters lauded the foresight of Charles Ferguson’s seminal 1959 article “Diglossia,” which “offered not only an admittedly idealized characterization of diglossia in Arabic, but also pointed out how and why it might change” (77); Walters posited that this change was in fact taking place in Tunisia. However, the changing language situation in Tunisia (e.g. in Sayahi, 2019) has not received as much attention as that of Egypt and Morocco (Elinson, 2013; Kindt & Kebede, 2017; Miller, 2017) and studies on Tunisia are more likely to focus on the dynamics between Arabic and French (Daoud, 2011; Guellouz, 2016; Walters, 2011). Although it is likely still premature in 2020 to declare Tunisia “postdiglossic”—as Walters did in 2003—I argue that the incipient changes in the status of Tunisian diglossia he described have greatly progressed and, in fact, accelerated. This change is reflected by the encroachment of Tunisian Arabic into previously *fushā*-only domains and can be explained by: 1) the near-universal literacy rates of young Tunisians; 2) the increased opportunities for informal writing in Tunisian Arabic, especially online; and 3) the national pride and sense of Tunisian exceptionalism following the 2011 revolution. Tunisia thus provides a distinctive point of comparison to the rest of the Arab world, and these developments have significant implications for the study of both diglossia and the processes of language standardization.

One of Ferguson’s conditions for the development of a stable diglossia was widespread illiteracy (1959:338) and this was certainly the case in pre-modern Tunisia: only 9% of Tunisians born before 1914 were literate in Arabic. By 1994, however, youth literacy had reached 93.5% (Walters 2003:85–87), and today it is nearly universal at 97%. This is significantly higher than in Egypt (88%) or Morocco (82%) (UNICEF 2015). As Ferguson predicted more than half a century ago, widespread literacy destabilizes diglossia because, when ordinary people write to each other about ordinary topics, a genre of *informal writing* is created; in other words, opportunities to write in the “spoken” language arise. Although Walters noted that “occasions for writing the dialect ... are expanding” (2003:101), these occasions were still limited in 2003. Because of the internet, in 2020 this is no longer the case. The informality of online writing and text messages cues the use of colloquial Arabic rather than the formal variety (Abu Elijah, 2014) and the availability of such genres in Tunisia has greatly expanded: almost half of the population has internet access, and there are more mobile phones than there are people in Tunisia (UNICEF 2017). This provides daily opportunities for Tunisians to write in their colloquial variety and for informal writing in general. In addition to short-form genres like text messages and Twitter in Tunisian Arabic, we also find Tunisian blogs with long-form essays and amateur fiction.

All of this digital writing has served to erode the cultural taboos against reading and writing in the colloquial and has led to the increasing acceptability of *printed* works in Tunisian Arabic. Unlike

Egypt, Tunisia does not have a long tradition of producing literature in colloquial Arabic: the first Tunisian Arabic novel was not published until 2013 (Ben Brik) and prior to that even the use of colloquial Tunisian in dialogues in novels was rare. But Ben Brik's novel appears to have opened the floodgates: on a 2019 research trip to Tunisia I collected *eight* Tunisian Arabic novels, dealing with subjects as varied as women's rights (Fazaâ, 2017), religious extremism (Ezzine, 2017), and LGBT rights (Charfeddine, 2019). I also obtained two recent translations, including of classical philosophy, into Tunisian (Boulselmi, 2019; Mistoura, 2017). Judging by the current editions offered, the Tunisian novels have been remarkably successful, selling thousands more copies than the average Arabic-language novel in Tunisia.

It is notable that all of these works were published *after* the revolution of 2011 and subsequent democratic transition. Although a latecomer in the expansion of colloquial into fiction (compared to, especially, Egypt), Tunisia has—uniquely in the Arab world—seen an expansion of colloquial into government. Most notably, the new Tunisian constitution was issued in three languages: Arabic, French, and Tunisian Arabic (Mejri 2017). Tunisian Arabic is also used in the annual budget and discussions in the National Assembly, in service of government transparency and accessibility (Sayahi, 2014, 2019). Of the three conditions mentioned by Ferguson that would lead to diglossia being viewed as a “problem” and thus to the eventual breakdown of the system, one was “desire for a full-fledged standard ‘national’ language as an attribute of autonomy or of sovereignty” (1959:338). It could be argued that the revolution—and its resulting sense of national pride and Tunisian exceptionalism—has begun to fulfil this condition. So while it may still be premature to describe Tunisia as “postdiglossic,” there is no denying that the leaks Ferguson acknowledged between the idealized High and Low varieties in 1959 have now become a flood: Tunisia is now a prominent exemplar of the breakdown of diglossia in the Arab world.

Detecting and Measuring Hate Speech in Arabic Social Media
Khaled Elghamry, Ain Shams University, Egypt
kelghamr@indiana.edu

There is an increasing interest in detecting and monitoring hate speech on social media platforms (e.g. Facebook and Twitter), yet research on the subject so far is very limited for Arabic social media. This paper describes [1] bootstrapping a hate speech lexicon for Arabic, [2] a novel method for calculating a “seriousness” score for hate terms that predicts the likelihood of this virtual hate speech leading to violent acts in reality, and [3] using this lexicon in a template-based fashion to identify potential targets of hate on the Arabic-speaking Twitter.

Assuming that hate speech is abusive speech targeting specific group characteristics, such as ethnic origin, religion, gender, or sexual orientation (Warner & Hirschberg 2012), we constructed a hate lexicon for Arabic as follows. A bootstrapping process was initiated using a seed of 25 frequent hate speech terms in the Arabic discourse, such as terms representing the target as “low/سافل”, “dirty/قذر”, “infidel/كافر”, or as subhuman through using dehumanizing animal metaphors of targets, for example, as “pigs/خنزير” or “dogs/كلاب” (Bahaa-eddin, 2012). The seed was expanded by adding the (near) synonyms of these terms, when available. These terms were used as search terms in Twitter from January/1/2011 to January/1/2016. The terms coordinating in the search results with at least 3 terms in the seed were extracted and manually filtered to identify potential hate terms. This resulted in a total of 527 hate terms.

Search results were processed further to identify potential hate patterns. A word sequence is a potential pattern if it coordinates with or immediately follows at least 3 unique terms in the core 527-term lexicon. Expressions satisfying this condition were identified then grouped, resulting in the following 3 patterns. The “Son_of” pattern: the word “ابن/son” or “بنت/daughter”, followed a noun expressing “lowness” and “sub-humanness” (e.g. حمار/donkey, سفاح/adultery and شوارع/streets). The Cursing pattern: cursing expressions (e.g. لعنة الله على/May Allah curse and شئت، ابعد، احرق، دمر، شرد، اللهم اهلك/Oh Allah: destroy, disperse, burn, wipe out, eliminate, eradicate). The “Deserve” pattern: an

expression with the general meaning of “deserve” (e.g. *فيه حلال/يستحق*) (followed by a noun (e.g. *المشنقة*/gallows), verbal noun (e.g. *القتل*/killing) or an infinitive-phrase in the passive form (e.g. *أن يقتل*/to be killed) that expresses a violent act.

It was observed that the nouns, clauses or phrases that follow the “deserve” expressions seem to correlate positively with the seriousness of the hate term. Accordingly, this pattern was used to calculate a seriousness score for each term. This score can be interpreted as an indicator of the potential violent acts that the hate speech could instigate in reality. To calculate this score, the words following the deserve word/phrase were compiled then manually grouped into classes, where phrases or nouns that refer to similar “punishments/verdicts” were grouped into the same “verdict class”. For example, the nouns *القتل*/killing, *الإعدام*/execution, *الذبح*/slaughtering and *التصليب*/crucifixion were grouped under DEATH. This resulted in a total of 15 classes such as DEATH, HUMILIATION and HARASSMENT. We calculated a score for each term in our lexicon based on its frequency of co-occurrence with these classes. This resulted for example in the following scores 0.81, 0.625, 0.58 and 0.36 for the terms *كافر*/infidel, *خائن*/traitor, *خنزير*/pig and *حقير*/base, respectively, with a score of 1 indicating maximum seriousness.

Identifying hate targets here expands on the template-based method in (Silva et. al 2016), which identifies hate in social media in English using the expression “I hate” and its synonyms. The assumption is that words or phrases immediately following these expressions are very likely targets of hate speech. In addition to “I hate”, we also used the terms in our lexicon and the 3 patterns of hate to identify potential targets. We were able to identify a total of 293 likely targets of hate speech in our corpus. The top 20 targets included groups such as: *الشيعة*/Shia, *الإخوان*/Ikhwan *المسيحيون*/Christians, *اليهود*/Jews, *الملحدون*/Atheists and *المثليون*/homosexuals.

References

- Bahaa-eddin, M. (2012). *HateSpeak in contemporary Arabic discourse*. Cambridge Scholars Publishing.
- Silva, L., Mondal, M., Correa, D., Benevenuto, F., & Weber, I. (2016, March). *Analyzing the targets of hate in online social media*. In 10 Int'l AAAI Conference on Web and Social Media.
- Warner, W., & Hirschberg, J. (2012, June). *Detecting hate speech on the world wide web*. In Proceedings of the workshop on language in social media (pp.19-26). Association for Computational Linguistics.

Trajectories of change: The formation of lexical raising predicates and subj- to-subj raising constructions in vernacular Arabic

Maris Camilleri (mcamil@essex.ac.uk) Louisa Sadler (louisa@essex.ac.uk)

University of Essex

We here start off from the premise that across the different Arabic vernaculars, raising structures, particularly subj-to-subj raising constructions that consist of lexical raising predicates (rather than raising predicates with functional bearings, such as phrasal verbs, for instance) are available, and that the presence of such constructions is distinct from the wider distribution of (straightforward) copy raising structures, which very often, but not necessarily always so, involve the presence of the complementisers *kaPinn*/*kinna*/*killi*, depending on the dialect. At focus here will be lexical raising predicates and the step-by-step progressions that yield to the formation of the raising constructions they participate in.

While the verbs *dher* and *bēn* both meaning ‘seem, appear’ do function as SSR predicates, at least in different Western dialects (Camilleri and Sadler, 2018; 2019), other lexical predicates have grammaticalised into ‘seem, appear, as though’ predicates as a result of changes in the semantics. The imperfective forms of *ʔāl* lit. ‘say’, as discussed specifically for Egyptian Woidich (2007), but which are quite widespread eastwards, and in the Gulf, are perhaps the best known examples. In principle, while in accordance with typological observations made in Heine and Kuteva (2002) and Kuteva et al. (2019), where from the semantics of ‘say’ we end up with ‘like, as though’ functions, the semantic development of *ʔāl*, to the best of our knowledge has not yet developed raising type properties, even if the new semantics it takes on is shared with lexical predicates associated with the realm of ‘seem, appear, be as though’, etc. Yet perhaps the most notable change in semantics (and argument-structure) from an intransitive verb into a raising verb is that associated with *Talaʔ* lit. ‘go (up)’, discussed to our knowledge for the first time in its use as a raising verb, in Camilleri and Sadler (2019). What’s special about this verb is that the semantic change associated with the original lexical meaning ‘go (up)’ into a predicate meaning ‘seem, appear, turn out to be’ is unprecedented, and has not been shown to be present elsewhere, crosslinguistically. What we are after here are such similar sorts of lexical-to-lexical changes, given that the end result, we argue, is that of a lexical raising predicate, but where the changes observed and involved go much beyond mere changes in the semantics and associated argument-structure differences.

We here hypothesise two distinct diachronic paths that have led to the grammaticalisation of raising predicates in vernacular Arabic, and hence, the formation of the subj-to-subj raising constructions in which they appear.

1. A categorial change from P or N, to V, as in (1a)-(1c), resulting in the formation of a lexical pseudo-verb with raising properties;
2. A change involving more structural complexity: a P (typically meaning ‘like; as’ (Taine-Cheikh, 2004)) + Complementiser, which have fused morphophonologically, primarily as a result of procliticisation of the P onto the C, and have been (categorially) reanalysed as a V that then takes its own complement. This, we argue, however, to be only one possible direction of this source construction. The other alternative grammaticalisation is the formation of a C, which, we argue is what underlies the formation of the complementisers *kaPinn*, *killi*, and *bèalli* across the different dialects, which prototypically are present in copy raising constructions.

- (1) a. *zei-na mberrd-în*
 like-1pl.gen cold-pl
 We seem cold. Djidjelli Algerian: Marcais (1954, p. 524)
- b. *šor-ha habl-a*
 towards-3sgf.gen crazy-sgf
 She seems like she's crazy. Benghazi Libyan: Saad (2019)
- c. *šikil-ak ma haDDar-it-š malih*
 form-2sgm.gen neg prepare.pfv-2sgm-neg well
 You seem to not have prepared well (for the exam). Jordanian: Jarrah and Alshamari (2017, p. 32)

We will be expanding our hypothesis by providing a reconstruction of how these changes could have progressed, highlighting behaviours associated with, and attributed to a prominent internal possessor, in the case of *šekl*, for instance, which is a grammaticalisation trigger already observed in other constructions across the vernaculars involving the grammaticalisation of predicative prepositions used in possessive constructions (Camilleri, 2019), as well as changes from adjunction to embedding structures, which we argue is a path well-throdden in the grammaticalisation of a number of syntactic structures as identified elsewhere for other constructions across the vernaculars in Camilleri and Sadler (2018; accepted).

References

- Camilleri (2019). The development of the universal perfect in Arabic. *Language*, 95(4), 683-710.
- Camilleri and Sadler (2018). Schematising (Morpho)Syntactic Change in LFG: Insights from grammaticalisation in Arabic. In M. Butt and T. H. King (eds). *Proceedings of the LFG18* (pp. 129-149). Stanford: CSLI.
- Camilleri and Sadler (2019). A Non-Canonical Diachronic Formation of Raising Predicates. In M. Butt, T. H. King, and I. Toivonen (eds). *Proceedings of the LFG19* (pp. 90-110). Stanford: CSLI.
- Camilleri and Sadler (accepted). The grammaticalisation of an auxiliary and a copula: The Arabic 'sit' participle. *Journal of Historical Syntax*.
- Heine, B. and T. Kuteva (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Jarrah, M. and M. R. Alshamari (2017). The syntax of the evidential particle *šikil* in Jordanian Arabic. *Italian Journal of Linguistics* 29(2), 29-56.
- Kuteva, T., B. Heine, B. Hong, H. Long, S. Rhee, and H. Narrog (2019). *World lexicon of grammaticalization*. Cambridge University Press.
- Marçais, P. (1954). *Textes arabes de Djidjelli*. Paris: Presses universitaires de France.
- Saad, A. (2019, June). L'évolution sémantique de *šor* (vers, en direction de) dans le dialecte de Benghazi (Libye). Poster presented at NACAL47, Paris.
- Taine-Cheikh, C. (2004). De la grammaticalisation de 'comme' (comparatif) en arabe. In R. d. J. M. Haak and K. Versteegh (Eds.), *Approaches to Arabic Dialects*, pp. 309-328. Leiden: Brill.
- Woidich, M. (2007). Everything you always wanted to know about 'äl, yi'ül 'to say' in Egyptian Arabic. In *Approaches to Arabic Linguistics*, pp. 675-700. Leiden: Brill.

Is Morphological Case a Feature of Individual Nominal Elements?

Evidence from Standard Arabic

Basem Al-Raba'a, KIMEP University

balrabaa@kimep.kz

According to Chomsky (1995, 2000, 2001), Case is a byproduct of complete phi-feature (number, person, gender) agreement between functional heads (probes) and noun phrases (goals); for example, the agreement (Agree) relation between a verb, particularly little *v*, and its direct object results in the realization of accusative Case on the object, from which it is transmitted via agreement to other elements like a modifying adjective within the determiner phrase (DP), if any (see also Fassi Fehri 1999, Carstens 2000). Another view entertains the idea that the Case feature spreads to other elements of the DP by way of a feature-sharing mechanism (Pollard and Sag 1994, LeTourneau 1995, Danon 2011). The transmission of Case, whether through agreement or feature sharing, is described as a two-part model as it involves two successive steps: first, a head assigns Case to a noun phrase (NP), and second, that Case is distributed to the rest of nominal elements inside the DP (Brattico 2010). While this approach can straightforwardly account for cases in which nouns and adjectives bear the same morphological case, as in **Error! Reference source not found.**, it fails to capture Case marking within DPs whose nominal elements receive different morphological cases, as in **Error! Reference source not found.**–**Error! Reference source not found.**

- (1) Badr-un ʔakala ʔaʕaam-an laḏiiḏ-an.
Badr-NOM ate food-ACC delicious-ACC
'Badr ate delicious food.'

- (2) marar-tu bi-Zajd-in wa-ʕamr-an.
passed-1P.SG.NOM by-Zayd-GEN and-Amr-ACC
'I passed by Zayd and Amr.'

(Ibn Jinni n.d., cited in Wagih 2012: 138; my translation)

- (3) marar-tu bi-Zajd-in aḏ-ḏariif-a.
passed-1P.SG.NOM by-Zayd-GEN the-funny-ACC
'I passed by the funny Zayd.'

(Wagih 2012: 138; my translation)

Note that *Zajd-in* is assigned genitive in **Error! Reference source not found.**–**Error! Reference source not found.**, whereas the second conjunct *ʕamr-an* in **Error! Reference source not found.** and the adjective *aḏ-ḏariif-a* in **Error! Reference source not found.** are assigned accusative. To uniformly capture the distribution of morphological case in Standard Arabic, whether the nominal elements of the DP carry matching or mismatching overt cases, as respectively shown in **Error! Reference source not found.**–**Error! Reference source not found.**, I argue, by applying Kayne's (2002) approach to Case assignment, that a certain probe establishes an Agree relation with each nominal element within the DP separately, resulting in the valuation of Case on each lexical item independently of the other (see also Brattico 2008, 2010, 2011). This indicates, as the data in **Error! Reference source not found.**–**Error! Reference source not found.** show, that Case is a property of individual nominal elements rather than DPs, which is in conformity with Kayne's model. However, such data show that the accusative Case in **Error! Reference source not found.**–**Error! Reference source not found.** is assigned by a non-local c-commanding probe (little *v*) rather than by the local preposition (P), which runs counter to the locality restriction on Case assignment in Kayne's approach.

Drawing on Arabic data (including the above examples as well as other instances in Arabic where we find morphological case mismatches such as negative noun phrases, complex adjective phrases, and quantified construct-states), this study shows that Case assignment concerns only lexical items (e.g., noun, adjective, quantifier) rather than DPs, and that being phi-complete is not a necessary condition for the valuation of Case (Soltan 2007, Aldholmi, Ouali and Trinh 2019), contra Chomsky (2000, 2001). Finally, I argue that the locality constraint on Case assignment (Kayne 2002, Brattico 2010)

cannot capture all the Arabic facts and therefore should be reformulated with respect to Arabic in such a way that the locality and non-locality of Case assignment should represent unmarked and marked structures, respectively.

Variation in verbal negation in Jordanian Arabic
Ekab Al-Shawashreh
Yarmouk University
shawashreh@yu.edu.jo
Marwan Jarrah
University of Jordan
m.jarrah@ju.edu.jo

This study investigates the variable expression of verbal sentences negation in Jordanian Arabic (JA). Verbal sentences in JA can be negated through three different ways (with no effect on meaning). The pre-verbal negation particle *ma:* (not) can be used alone (see (1)); it could also be reinforced by the negation suffix *-f* which is encliticized onto the verb (see (2)); or the suffixal negation clitic, *-f*, can be used alone (see (3)).

(1) ?ana	ma:	baħibb	?il-hali:b	(the preverbal negation variant)
I	NEG	like	the-milk	
'I don't like milk.'				
(2) ?ana	ma:	baħib-bif	?il-hali:b	(the discontinuous negation variant)
I	NEG	like-NEG	the-milk	
(3) ?ana	baħib-bif	?il-hali:b		(the post-verbal negation variant)
I	like-NEG	the-milk		

While negation variation has been extensively investigated in many languages (e.g., Childs 2017), it can be noticed that much previous research on negation in Arabic has mainly focused on syntactic, morphological and semantic aspects of negation (e.g., Al-Salem 2012, Benmamoun et al. 2013). Systematic analyses of morpho-syntactic variation in negation based on spontaneous speech data are given sparse attention in contemporary studies of syntactic variation in Arabic. Such paucity of quantitative studies on the variable negation of verbal sentences provides the primary motivation for the present study (see Al-qassas 2012 for a recent work that investigates some aspects of this variation in JA. Note that Al-qassas 2012 based his work on a convenient sample and limited his analysis to distributional analysis which is widely deemed insufficient to capture the variation in precise terms).

Following the tenets of variationist sociolinguistics (Labov 1972), we investigate verbal sentences negation in a corpus of spontaneous JA recorded in the Irbid metropolitan area in the fall of 2018. The corpus (over 50 hours of audio-recordings obtained from 50 speakers stratified by age, sex, level of educational attainment, and locality) explores the distribution of the different variants of verbal sentences negation in JA and decides how social factors affect the choice of verbal negation variants. The number of tokens retained and coded for analysis is 1069. The distributional results show that pre-verbal negation is the most frequent variant (55.2%), followed by the discontinuous negation variant (31.8%), which in turn is followed by the post-verbal negation variant (13%). Young speakers slightly favor post-verbal and discontinuous negation more than old speakers who favor pre-verbal negation more than their young counterparts. Females favor pre-verbal negation more than males who favor discontinuous and post-verbal negation. For region, urban speakers favor pre-verbal negation more than rural speakers who, in turn, favor discontinuous and post-verbal negation more than their urban counterparts.

The results further reveal that educated speakers favor pre-verbal negation more than uneducated speakers who, in turn, favor discontinuous and post-verbal negation. Cross-tabulations of age and the other social factors provide some remarkable insights. First, a cross-tabulation of age and gender shows that young females, old females and old males favor pre-verbal negation while young males favor discontinuous negation. Similarly, a cross-tabulation of age and region shows that young urban speakers favor pre-verbal negation while young rural speakers favor discontinuous negation. A third cross-tabulation of age and level of education reveals that educated speakers (young and old) favor

pre-verbal negation while uneducated speakers (young and old) favor discontinuous negation. This result can be interpreted in light of two viewpoints. The first view relates to the effect of Standard Arabic, which only permits pre-verbal negation. Educated speakers favor pre-verbal negation as it is syntactically similar to that in Standard Arabic. Educated speakers might feel badly when they use post-verbal or discontinuous negation as they are not used in Standard Arabic (the language of educated speakers). Although using the non-standard negation particle -f is not stigmatized in JA, it can be taken as a sign of colloquial speech (see Coveney 2002, along these lines for similar pattern found in French). The other interpretation of this result can be discussed within Al-Wer's (2002) viewpoint where education is viewed as a factor that increases the chance of expanding the speaker's social network and mobility. Variable rule analyses of the social factors (age, sex, education and locality) show that verbal negation variation in JA is socially constrained. This result is in line with the ones reported in previous studies for other languages (Meisner, 2010). Variable rule analyses of the social factors reveal that level of education, region and sex are statistically significant in constraining variant choice. Based on the range values of the social factor groups, level of education has the strongest effect in determining variant choice.

The constraints hierarchy within this factor group shows that speakers with post-secondary education favor pre-verbal negation (.67) while those with intermediate education and no education disfavor pre-verbal negation (.44) and (.29) respectively. The factor group with the second strongest effects on variant choice is region (locality), with urban speakers favoring pre-verbal negation (.67) while rural speakers disfavoring it (.34). The results also show that speaker's sex has the third strongest effects on variant choice. Females favor pre-verbal negation (.62) whereas males disfavor it (.39). Age is found to be statistically non-significant in conditioning variant choice. The results constitute a strong piece of evidence to the effectiveness of accountable analyses of spontaneous speech in uncovering key patterns of morpho-syntactic variation in one of contemporary spoken varieties of Arabic.

Generational changes in VOT in Qatari Arabic

Vladimir Kulikov (vkulikov@qu.edu.qa), **Najla Hamad M H Al-Hajri**
(na1509967@student.qu.edu.qa), and **Buthaina Mohammad A J AL-kuwari**
(ba1507760@student.qu.edu.qa)
Qatar University

Previous research on Voice Onset Time (VOT) in the Arabic dialect of Qatar (Kulikov, 2019) revealed that VOT in voiceless stops ($M = 55$ ms) is longer in Qatari Arabic than VOT in other Arabic dialects of the Gulf (e.g., AlDahri, 2013, for Saudi Arabic). Longer VOT in stops in this dialect suggests that voiceless stops in Qatari Arabic, unlike in other dialects, are aspirated (Kulikov, 2019), which indicates a linguistic innovation.

Some findings (Kulikov, 2016) suggest that variation in VOT can be, in part, explained by social factors. Urban female Qatari speakers were shown to produce longer VOT than female speakers of the Bedouin dialect. Therefore, VOT lengthening may be an indicator of an on-going sound change. This scenario is predicted by Labov's (2003) Cascade model and Trudgill's (1974) Gravity model of diffusion, who argue that a sound change is more likely to originate in urban areas. In addition, previous studies (Eckert, 1989; Grégoire, 2006; Labov, 1990, 2010; Milroy & Milroy, 1993) showed that innovative linguistic features are more frequently used by women.

We argue that VOT lengthening in Qatari Arabic is a new linguistic variable that distinguishes the two major communities in the Qatari society, Hadar (or urban dwellers) and Bedouin (or rural dwellers), in addition to traditional variables, e.g., /j-dʒ/ or /k-tʃ/ (Al-Amadidhi, 1985). Longer aspiration in voiceless stops is a prestige variant that marks the urban (Hadar) variety. We predict that the sound change is led by young Hadar speakers. Speakers of the rural (Bedouin) variety use a more traditional variant with shorter aspiration. Young Bedouin speakers, however, are departing from the norms of the rural community and are developing longer aspiration in voiceless stops. This trend should be more prominent in female speakers.

These predictions were tested in an experimental production study. The acoustical data came from 40 female speakers of Qatari dialect, who were divided into two groups according to their tribal affiliation (Hadar or Bedouin) and two age groups (younger than 30 years and older than 30 years). The participants pronounced words with initial /t/ ($n = 12$) and /k/ ($n = 12$) in stressed syllables before long vowels /i:, a:, u:/ in their native dialect. Distractors ($n = 15$) were used to mask the interest in the voiceless stop category. VOT intervals were manually marked in PRAAT, and the total of 960 tokens were submitted to an acoustic analysis.

The results showed that variation in VOT can be explained by both tribal affiliation and age of the speakers (Fig. 1). Both main effect of Tribe ($F(1,36) = 5.17, p < .05$) and main effect of Age ($F(1,36) = 11.03, p < .01$) were significant. VOT in voiceless stops in Hadar speakers was 8 ms longer than in Bedouin speakers. Younger speakers had longer aspiration than older speakers ($MD = 12$ ms). A Vowel*Tribe interaction ($F(2,72) = 5.03, p < .05$) revealed that the difference in VOT between tribes was maintained before /a/ and /u/ ($p < .01$), but was neutralized before /i/ ($p = .388$).

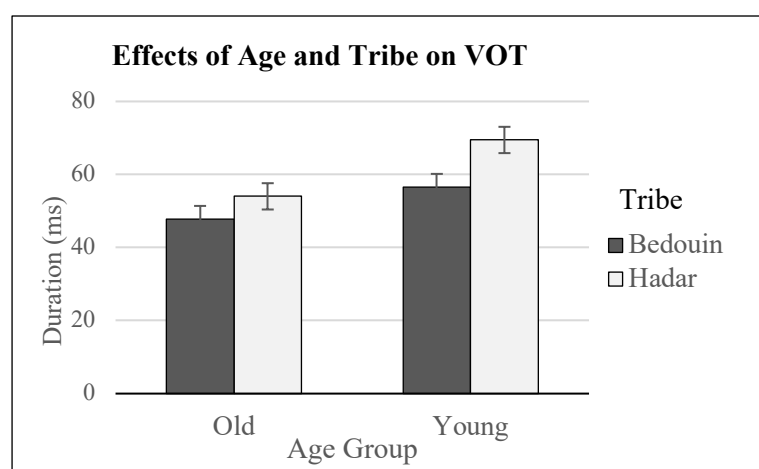


Figure 1. Differences in VOT between young and old speakers in the Hadar and Bedouin communities.

The results are consistent with previous studies (Kulikov, 2016) that reported a 10 ms difference in VOT between Hadar and Bedouin speakers. The findings support our predictions and suggest (1) that the longer aspiration marks the speech of young urban speakers, and (2) that longer VOT is likely to be a prestige variant as young Bedouin speakers are also acquiring longer aspiration in some vocalic contexts.

Linguistic Globalization and New Urbanism: Mapping a Single Indexical Field of [tʃ] in French, Cairene Arabic, and Casablanqan Arabic
Sarah Ruth Schwartz, University of Texas at Austin
srschwartz@utexas.edu

This paper studies palatalization of /t/ as a sociolinguistic variable in Cairene Arabic, Casablanqan Arabic, and Standard French, combining both prior sociolinguistic research on the indexicality of [tʃ] with new empirical data. It maps the n and $n+1$ order indexes of this feature as observed in each of these three languages into a single indexical field that provides insight into how indexicality travels across borders, and how globalization is creating a new urbanism in non-Western cities. It answers the following questions: How can we construct and interpret a single indexical field for [tʃ] across French, Casablanqan Arabic, and Cairene Arabic? What are the implications of [tʃ] being simultaneously indexed as both rural and urban, and both masculine and feminine? Mapping the indexical field of [tʃ] between languages and across borders will contribute both to our understanding of the linguistic effects of globalism and transcultural flows as well as to the development of third wave sociolinguistic methodology.

Data for the paper comes from two main sources: a metastudy of prior sociolinguistic work into the indexicality of [tʃ] in Cairene Arabic and French, and the author's ongoing variationist study of /t/ in Casablancon Arabic. Haeri (1997) first analyzed sociolinguistic variation in the realization of /t/ in Cairene Arabic. She found that variation is dependent upon gender and class, with upper middle-class women producing light palatalization [tʃ], and middle and lower-middle class women producing full palatalization [tʃ], with [tʃ] becoming stigmatized as "low-class". Jamin (2004) reported on the indexicality of [tʃ] in France, showing that the variable's indexical field has shifted in the wake of recent North African immigration. Whereas the palatalization of /t/ preceding a high front vowel in French predates the arrival of North African immigrants, the fact that these immigrants natively produced [tʃ] in their language shifted the perception of the variable; it was reanalyzed as North African style, and when they settled in the low-income areas it was reanalyzed to index "ghetto" talk, roughness, and low socioeconomic class. Little has been written on /t/ in Casablancon Arabic, other than the author's ongoing variationist study of Moroccan rappers and slam poets. Data was taken from 42 speakers, aged 23-41, in Moroccan Arabic, who were observed in two styles: verbal art performances and semi-structured interviews. The data show that Casablancon [tʃ] is a Hip-Hop performance feature, with n+1 indexes of toughness and masculinity. However, in the broader Moroccan culture the feature is considered "rural," entering the dialect due to internal migration; "rural" features are indexed as rough, masculine, and lower class, whereas traditionally "urban" features are indexed as soft, feminine, and upper class throughout the Arabic-speaking world (Hachimi 2007). Thus, rapid urbanization in Casablanca and the spread of global Hip-Hop culture has resulted in the construction of a new Moroccan urbanism.

As such, we see overlap in the indexicality of [tʃ] between these three linguistic codes – even across linguistic boundaries, this single feature has a coherent indexical field. This paper maps the unified indexical field of [tʃ], and explores two significant implications of this overlap. First, as Arab cities have grown in the past 50 years (Miller 2007) as a result of the globalization of industry, a new form of urbanism has developed that resembles the working-class urbanism attested in Western cities (Trudgill 1972). This implicates the effect of globalization on linguistic style. Second, a feature's index can change across borders as a result of international migration. Finally, this project also expands the indexical field model of third-wave sociolinguistics because it shows indexical connections of a single feature across different languages, whereas prior applications have focused solely on registers of a language (Eckert 2008), or regional dialects of a language (Walker et. al. 2014).

References

- Eckert, Penelope. 2008. "Variation and the indexical field." *Journal of Sociolinguistics* 12(4), pp. 453-476.
- Haeri, Niloofar. 1997. *The Sociolinguistic Market of Cairo: Gender, Class, and Education*. London: Kegan Paul International.
- Hachimi, Atiqa. 2007. "Becoming Casablancon: Fessis in Casablanca as a Case Study." In *Arabic in the City: Issues in Dialect Contact and Language Variation*, edited by Catherine Miller, Enam al-Wer, Dominique Caubet, and Janet C.E. Watson. London: Routledge, pp. 97-122.
- Jamin, Mikaël. 2004. "'Beurs' and Accent Des Cités: A Case Study of Linguistic Diffusion in La Courneuve." *Contemporary French and Francophone Studies* 8 (2): 169–76.
- Miller, Catherine. 2007. "Arabic Urban Vernaculars: Development and Change." In *Arabic in the City: Issues in Dialect Contact and Language Variation*, edited by Catherine Miller, Enam Al-Wer, Dominique Caubet, and Janet C.E. Watson, 1–31. New York: Routledge.

Trudgill, Peter. 1972. “Sex, Covert Prestige, and Linguistic Change in the Urban British English of Norwich.” *Language In Society* 1(2): pp. 179-195.

Walker, Abby, et. al. 2014. “Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change* 24: pp. 169-189.

Conditions on Definiteness in Arabic Superlatives
Peter Hallman, Austrian Research Institute for Artificial Intelligence
peter.hallman@ofai.at

This paper brings data from contemporary Arabic to bear on two related issues in the relation between superlatives and definiteness, namely 1) whether definiteness in superlatives is semantically contentful, and 2) whether adverbial superlatives are fundamentally nominal. Arabic superlatives alternate between a morphologically definite format and an indefinite format, yet this alternation is not associated with any differences in interpretation, militating against analyses that strictly connect morphological definiteness to semantic definiteness, at least for Arabic. Adverbial superlatives, though, are never morphologically definite, even though they are explicitly nominal. The occurrence of morphological definiteness in Arabic superlatives is contingent on both DP-internal order and the use of the DP as an argument or an adverb. I claim that morphological definiteness is morphosyntactically conditioned but not semantically contentful in superlatives. The paper treats Syrian Arabic, which is typical of the Arabic dialects in the relevant respects.

Szabolcsi (1986) and Heim (1999) point out that a superlative construction like (1) has two readings, one where we compare mountains in terms of height (the ‘absolute’ reading) and another where we compare Sarah to other mountain climbers in terms of the how high a mountain they climbed (the ‘relative’ reading).

- (1) Sarah climbed the highest mountain.

Szabolcsi and Heim observe that definiteness is incompatible with the composition they attribute to the relative reading, and claim that *the highest mountain* is semantically indefinite on that reading, but definite on the absolute reading. Other analyses seek to maintain a consistent relationship between form and meaning by attributing a meaning to *the* which is either different in the two interpretations (Sharvit and Stateva 2002, Krasikova 2012) or weak enough to be compatible with both, given an appropriate analysis of the two readings (Coppock and Beaver 2014, 2015).

In Arabic, adjectives typically follow the noun they modify. Definiteness is marked by the definite article *l-*, which occurs on the head noun and is copied onto all adjectival modifiers of the noun (2a). Further, a relative clause modifying a definite noun begins obligatorily with the relativizing particle *lli*. Indefinites are unmarked (2b), and relative clauses modifying an indefinite only optionally begin with *lli*.

- (2) a. *l-zabal l-ʕa:li *(lli) axad-na ʃu:ra minn-u.*
the-mountain the-high *(that) took-1PL picture of-it
‘The high mountain that we took a picture of’
b. *zabal ʕa:li (lli) axad-na ʃu:ra minn-u.*
mountain high (that) took-1PL picture of-it
‘A high mountain that we took a picture of’

Superlative adjectives are formed morphologically by inserting the base adjective into the ‘relative’ prosodic template $aC_1C_2aC_3$, where C_1 - C_3 represent the root consonants of the adjectival base, deriving, for example *abʕad* ‘furthest’ from *baʕi:d* ‘far’, *ashal* ‘easiest’ from *sahl* ‘easy’, *aʕla* (/aʕlay/) ‘highest’ from *ʕa:li* (/ʕa:liy/) ‘high’, etc. They may occur in the definite format in (3) with both an absolute and relative reading, though the relative clause blocks the relative reading (Sharvit and Stateva 2002).

- (3) sa:ra ʔaʕʕ-it ʕala l-zabal l-aʕʕa *(lli) axad-na ʕu:ra minn-u.
 Sarah climbed-3FS on the-mountain the-highest *(that) took-1PL picture of-it
 ‘Sarah climbed the highest mountain we took a picture of.’

But the superlative may also be indefinite (4), though here the adjective precedes the noun. In this case, the definite article may not appear and the relative pronoun *lli* is optional, as is typical for relative clauses with indefinite heads. This construction, too, has both an absolute and relative reading, though again the relative clause blocks the relative reading.

- (4) sa:ra ʔaʕʕ-it ʕala aʕʕa zabal (lli) axad-na ʕu:ra minn-u.
 Sarah climbed-3FS on highest mountain (that) took-1PL picture of-it
 ‘Sarah climbed the highest mountain that we took a picture of.’

The construction in (4) is morphologically indefinite both by virtue of the absence of the definite article and by virtue of the optionality of *lli* in the relative clause. Yet, there is no difference in use between the two formats in (3) and (4). It is easy to imagine that if absolute superlatives are definite and relative superlatives indefinite in English, as Szabolcsi and Heim claim, that a language with both definite and indefinite superlatives would apply them along these lines, using the definite superlative to express absolute readings and the indefinite superlative to express relative readings. But the morphological definiteness of superlative noun phrases correlates only with aspects of the internal structure of the superlative DP here, not with any aspect of its interpretation. And if definiteness is contentful, as Coppock and Beaver and others claim, then any vacuous alternation with indefiniteness is unexpected.

What, then, governs the alternation in (3)-(4), and what does this alternation tell us about the definiteness of superlatives in other languages? I approach this question by investigating adverbial superlatives in Arabic. The adverbial superlative phrase *aktar wa:ʕid(e)*, literally ‘most one-(F)’ in (5) is overtly morphologically nominal, as Matushansky (2008) claims adverbial superlatives are in general. It is modelled after the indefinite argument superlative in (4), but displays the dummy noun *wa:ʕid(e)* ‘one-(FEM)’, which agrees with the subject of comparison and therefore disambiguates two readings of the English translation to (5).

- (5) sa:ra bi-t-ʕibb muḥammad ʕalla:ḥ aktar waḥd-e / aktar wa:ʕid.
 Sarah IND-3FS-love Muhammad Sallah most one-FEM / most one_{MASC}
 ‘Sarah loves (the football star) Muhammad Sallah the most.’
 with *waḥde* (FEM): more than anyone else loves him
 with *wa:ʕid* (MASC): more than she loves anyone else

Unlike the argument superlatives in (3) and (4), the adverbial superlative can never be morphologically definite.

- (6) *sa:ra bi-t-ʕibb muḥammad ʕalla:ḥ l-wa:ʕid-(e) l-aktar.
 Sarah IND-3FS-love Muhammad Sallah the-one-(FEM) the-most

I construct an analysis of adverbial superlatives in which *wa:ʕid(e)* ‘one-(FEM)’ is restricted by a covert relative clause whose content is retrieved anaphorically from the degree relation that the superlative puts the subject of comparison in. With the feminine adverbial *aktar waḥde* ‘most one-FEM’ in (5) that relation is [$\lambda d \lambda x . x$ loves Muhammad Sallah *d*-much], so that (5) can be (very clumsily) paraphrased as *Sarah is the most one who loves Muhammad Sallah*.

The fact that the adverbial superlative cannot be morphologically definite falls under Longobardi’s (1994) generalization that morphological definiteness in DPs is licensed by argumenthood. I pursue an analysis along these lines, that attributes only a morphosyntactic function to the definite article in superlatives, but no semantic content. The definite article fills the determiner slot in DP by default

when DP stands in the Agree relation to a licensing head (or is ‘lexically governed’ in Longobardi’s terms). The indefinite construction in (4) is a ‘construct state’ construction in which the preposed adjective occurs in [spec,DP] and fulfills the requirement that DP have phonological substance in the appropriate environment, obviating the determiner itself (based on Ritter 1988). The absence of a construct state construction in English requires that D itself be filled. The overt indefinite article is a poor choice to fulfill this syntactic condition on D because it brings with it the implicature (which I claim is part of its lexical meaning) that the discourse referent it introduces is discourse-new. This conflicts with the presupposition of *-est* that the subject of comparison is included in a contrast set that is presupposed, and therefore not strictly discourse-new; the computation of an NP like *highest mountain* precedes the computation of a larger indefinite phrase *#a highest mountain*, so the introduction of the contrast set precedes the introduction of the new discourse referent associated with *a*, violating the novelty condition of the indefinite article. This implicature can be defeated under certain conditions, as Herdan and Sharvit (2006) point out. In summary, I claim that the definite article is not semantically contentful in superlative constructions in at least Arabic, but that its occurrence is conditioned by morphosyntactic factors, and I endeavour to present a morphosyntactic analysis of those factors and spell out their significance for other languages.

References:

- Coppock, Elizabeth and David Beaver (2014) A superlative argument for a minimal theory of definiteness. In Snider, Todd, Sarah D’Antonio and Mia Weigand (eds.) *Proceedings of SALT XIV*, 177-196. LSA, Washington, DC.
- Coppock, Elizabeth and David Beaver (2015) Definiteness and determinacy. *Linguistics and Philosophy* 38:377-435.
- Heim, Irene (1999) Notes on superlatives. Ms., U. of Texas at Austin.
- Herdan, S and Yael Sharvit (2006) Definite and nondefinite superlatives and NPI licensing. *Syntax* 9:1-31.
- Krasikova, Sveta (2012) Definiteness in superlatives. In Aloni, Maria et al. (eds.) *Logic, language and meaning*, 411-420. Springer, Dordrecht.
- Longobardi, Giuseppe (1994) Reference and proper names. *Linguistic Inquiry* 25:609-666.
- Matushansky, Ora (2008) On the attributive nature of superlatives. *Syntax* 21:26-90.
- Ritter, Elisabeth (1988) A head-movement approach to construct state noun phrases. *Linguistics* 26:909-929.
- Sharvit, Yael and Penka Stateva (2002) Superlative expressions, context, and focus. *Linguistics and Philosophy* 25:453-504.
- Szabolcsi, Anna (1986) Comparative superlatives. In Fukui, Naoko (ed.) *MIT Working Papers in Linguistics*, vol. 8, 245-265.

“Ps” in the varied perfect expressing constructions across Arabic

Maris Camilleri, University of Essex

mcamil@essex.ac.uk

The perfect is primarily understood as a specific construction whose core structure involves an auxiliary along with a participial form (Portner, 2011), which in Germanic/Romance associates with distinct semantic interpretations depending on the presence/absence of adjuncts (e.g. Dowty (1979)). In alternative readings of what the perfect is, it is treated as an aspectual category that bears relevance to tense, and is either associated with a four-fold categorisation in e.g. Comrie (1976); Brinton (1988), or with the broader semantic distinction between existential/experiential vs. universal/continuous (McCawley (1971); McCoard (1978)). In the case of Germanic/Romance, it is the same core structure, grammaticalised out of the possessive construction (Heine and Kuteva (2006); Drinka (2017)), that associates with, and expresses, the different readings. This is not true of what one finds in Arabic or in Celtic languages, for instances, with which Arabic draws a lot of parallels, as will be additionally shown.

Through a comparative array of Arabic dialectal data, I will demonstrate that while a semantic split between the Universal vs. the Existential perfect reading suffices, the syntax does not merely match up in an isomorphic manner. Rather, structural differences which go beyond the mere adjunction of adverbs render finer-grained readings, and that they are additionally indicative of distinct grammaticalisation trajectories. While the classic split that has been discussed so far in the literature on the vernaculars, summarised and brought together in Hallman (2016) and Camilleri (2019), makes reference to active participles and perfective forms as expressing an Existential reading, as opposed to designate constructions expressing a Universal reading, I here demonstrate that there is much more to be said about the landscape of perfect expressing construction in Arabic. There appear to be (at least) three specific constructions (beyond other structures involving mere adverbs in relation to perfective, or active participle forms) with varied distributions across the different dialects that express distinct Perfect readings. These are unified by virtue of the obligatory presence of prepositions (Ps), at least, formally; hence “Ps”, since it will be argued that these are functionally/categorially not such.

1. Universal perfect: “Ps” la/il ‘to, for’ (1), ʔand ‘at’
2. Existential perfect: “Ps” la/il ‘to, for’, bi ‘with, in, by’, maʔ ‘with’ (2)
3. Recent perfect/Short perfect (Johnson, 1981): “P” baʔd ‘after’, k/kī(f) ‘like’ (3)

(1) muna il-a xams iyyām bi-l-xabis

Muna to-3sgf.gen five day.pl in-def-prison

Muna has been in prison for five days.

(Damascene) Syrian: (Hallman, 2016, p. 77)

(2) aní máʔ-i ma-fhúm

inna as-sultan ...

I with-1sg.gen pass.ptcp-understand.sgm comp def-king

I have understood that the king ...

Andalusian: Corriente (2008, p. 375)

(3) kī usal

like arrive.pfv.3sgm

He’s just arrived.

(Tlemcen) Algerian: Marçais (1902, p. 192)

It will be argued that while both the Universal and Existential perfect constructions have grammaticalised from the same type of source structure, the Hot news/Short/Recent perfect reading has developed out of a separate construction. The split finds further support in parallels from Celtic, which, while the different languages within the family grammaticalise the same three distinct structures from different sources, intriguingly, the Ps ‘to’ and ‘at’, and the P ‘after’ are present in the expression of the Existential, and the Short perfect readings, respectively.

References

- Brinton, L. J. (1988). *The development of English aspectual systems: Aspectualizers and post-verbal particles*. Cambridge: Cambridge University Press.
- Camilleri (2019). The development of the universal perfect in Arabic. *Language*, 95(4), 683-710.
- Comrie, B. (1976). *Aspect*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Corriente, F. (2008). Árabe andalusí. In F. Corriente and Á. Vicente (Eds.), *Manual de Dialectología neoárabe*, pp. 353–378. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo.
- Dowty, D. R. (1979). *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*, Volume 7 of *Synthese Language Library*. Dordrecht: D. Reidel.
- Drinka, B. (2017). *Language Contact in Europe: The periphrastic perfect through history*. Cambridge: Cambridge University Press.
- Feghali, M. T. (1928). *Syntaxe des parlers arabes actuels du Liban*. Paris: Paul Geuthner.
- Hallman, P. (2016). The universal perfect in Syrian Arabic. *Brill's Annual of Afroasiatic Languages and Linguistics* 8 (1), 76–100.
- Haspelmath, M. (1998). How young is Standard Average European? *Language Sciences* 20(3), 271–287.
- Heine, B. and T. Kuteva (2006). *The Changing Languages of Europe*. Oxford: Oxford University Press.
- Heine, B. and M. Nomachi (2010). Is Europe a linguistic area? In M. Nomachi (Ed.), *Grammaticalization in Slavic Languages: From Areal and Typological Perspectives*, pp. 1–26. Hokkaido: Hokkaido University.
- Johnson, M. (1981). A unified temporal theory of Tense and Aspect. In P. Tedeschi and A. Zaenan (Eds.), *Tense and Aspect*, pp. 145–175. New York: Academic Press.
- Marçais, W. (1902). *Dialecte arabe parlé à Tlemcen*. Paris: Ernest Leroux.
- McCawley, J. D. (1971). Tense and time reference in English. In C. Filmore and D. L. Langendoen (Eds.), *Studies in Linguistic Semantics*, pp. 97–113. New York: Holt Rinehart.
- McCoard, R. W. (1978). *The English perfect: Tense-choice and pragmatic inferences*. Amsterdam/New York: North Holland Publishing.
- Portner, P. (2011). Perfect and progressive. In K. von Stechow, C. Maienborn, and P. Portner (Eds.), *Semantics: An International Handbook of Meaning*, pp. 1217–1261. Berlin: Walter de Gruyter.

Agree blocks movement: Evidence from Arabic

Marwan Jarrah, The University of Jordan

m.jarrah@ju.edu.jo

Based on the findings of Miyagawa (2010), Alshamari (2017) and Jarrah (2019), we argue that Agree as viewed in Chomsky (2001) may bleed movement in Arabic grammar. According to Jarrah (2019), if the probe assigns Morphological Case to its goal, the probe may not bear a ϕ -affix that expresses the ϕ -content of the goal. This indicates that Morphological case on the goal (assigned by the probe) blocks overt ϕ -agreement (of the goal) on the probe. Jarrah makes use of this analysis to account for the observation that the complementizer *ʔinn* in Arabic varieties (Modern Standard Arabic (MSA), Jordanian Arabic and Lebanese Arabic) does not bear a ϕ -affix that expresses the ϕ -content of the goal when it assigns its goal Morphological case (consider (1a), an example from MSA).

(1) ʔinna(*hu)	ʔal-walad-a	qaraʔa	ʔad-dars-a
Comp	DEF-boy-ACC	read.Past. 3SG.M	DEF-lesson-ACC
‘The boy read the lesson.’			

When the goal cannot receive a morphological case because the goal is a pro (as in (2a), another example of MSA) or because the language does not obtain Morphological case (including Arabic dialects) (consider (2b), an example from Jordanian Arabic), *ʔinn* should bear a ϕ -affix that expresses the ϕ -content of the goal.

(2) a. *ʔinna*(hu)* *qaraʔa* *ʔal-walad-u* *ʔad-dars-a*
 Comp read.Past.3SG.M DEF-boy-NOM DEF-lesson-
 ACC

‘The boy read the lesson.’

b. *ʔif-fab* *ʔiʕtaraf* *ʔinn-hum* *ʔixwat-uh* *zʰarab-u* *ʔil-walad*
 DEF-young man confess.Past.3SG.M Comp-3PL.M brothers-his hit.Past-3PL.M DEF-boy
 ‘The young man confessed that his brothers hit the boy.’

According to Jarrah, *ʔinn* agrees with a pro that is situated in Spec,TP in (2a). Given that the pro does not have a phonological form and hence cannot bear Morphological case, *ʔinn* bears a ϕ -affix (-*hu*) that shows the ϕ -content of the pro. This affix is taken as a morphological outcome of the Agree relation between *ʔinn* and the pro. In (1a), *ʔinn* agrees with the preverbal subject (which is assigned Accusative Case by *ʔinn*). Here, *ʔinn* is not allowed to bear a ϕ -affix displaying the ϕ -content of its goal as the Agree relation between them is RECORDED through the morphological Case. The complementary distribution between Morphological case and ϕ -agreement is ruled by the postulated interface principle *Agree Chain Record* which demands that Agree relations must have a PF record (i.e. an overt Case marking on the goal or, if not, a ϕ -affix on the probe). In an independent work, AlShamari (2017) shows that there are two types of topic particles (lexical words that mark the element serving as a topic in the sentence) in North Hail Arabic: agreeing particles and non-agreeing particles. Agreeing particles can bear a ϕ -affix that expresses the ϕ -content of the element that functions as the topic of the sentence as in (3a,b) below, whereas non-agreeing particles do not bear such a ϕ -affix; see (3c).

(3) a. *ʔedi-hin* *l-banaat* *ʕaf-an* *as-sayarah* *bi-a-sa:ħah*
 PRT-3P.F DEF-girls see.PST-PL.F Def-car in-Def-yard
 ‘The girls saw the car in the yard.’
 b. *al-hazi:mah* *ʕad* *ʔal-laʕibi:n* *lazim* *jidʒtahdu:n* *l-taweet-ah*
 the-defeat PRT Def-players must work harder to-compensate-it
 ‘As for the defeat, the players must work harder to compensate for it.’

The agreeing particle *ʔedi* agrees with the subject in (3a). According to Alshamari (2017), the subject in (3a) is a topic which is marked by *ʔedi* that functions as a topicalizer heading the Topic Phrase (Rizzi 1997). On the other hand, in (3b) the topicalized element, i.e. *alhazi:mah*, should move to the Spec position of the Topic Phrase headed by the non-agreeing particle, *ʕad*. When the topicalizer is an agreeing particle, the topic remains in situ. We interpret AlShamari’s data as follows. When the probe can agree with its goal and bear a ϕ -affix that displays the ϕ -content of its goal, the goal should NOT move to the Specifier position of the XP headed by the probe. This implies that an overt ϕ -agreement blocks movement; otherwise it would become mysterious why topics should not move to the Spec position of Topic Phrase (at least at PF). On the other hand, when an overt ϕ -agreement is not an option because the topicalizer does not bear uninterpretable ϕ -features, the goal should move to Specifier position of the Topic Phrase.

Combining the results of Jarrah (2019) and Al-Shamari (2017), we reach the following situation. Morphological Case blocks an overt ϕ -agreement, which in turn blocks overt movement (Morphological Case > overt ϕ -agreement > overt movement). This analysis, if true, implies that overt movement is just a mechanism to record the Agree relations between the probe and the goal. Miyagawa (2010) argues for a similar finding regarding this function of Move. However, Miyagawa does not link Move to morphological Case or ϕ -Agree in that Move may not occur at PF when the Agree relations between the probe and the goal are secured through an overt ϕ -Affix on the probe, a result that Al-Shamari (2017) implicitly argues for. However, Al-Shamari (2017) does not examine the instances where the topicalizer may assign Morphological Case to its goal. That might be

attributed to the assumption that topicalizers do not assign Case, and North Hail Arabic (of whose data Ashamari (2017) draws on) does not obtain Morphological case. Jarrah (2019) examines this relation between Case and ϕ -Agree, given that MSA is a morphological Case language, hence the interaction between Case and overt ϕ -agreement becomes possible.

References

AlShamari (2017). Topic particles in the North Hail Dialect of Najdi Arabic. Ph.D thesis. Newcastle University. UK. Chomsky, N. 2001. 'Derivation by phase'. In M. Kenstowicz (ed) *Ken Hale. A Life in Language*. Cambridge: MIT Press. 1-52. Jarrah, M. (2019). 'Record your Agree: A case study of Arabic complementizer *ʔinn*'. *Journal of Linguistics*. Miyagawa, S. 2010. Why agree? Why move? Unifying Agreement-Based and Discourse- Configurational Languages. Cambridge: MIT Press. Rizzi, L. 1997. 'The fine structure of the left periphery'. In L. Haegeman (ed). *Elements of Grammar*. Dordrecht: Kluwer Academic Publishers. 281-337.

Arabic and English plural formation in typically developing heritage speakers of Arabic: A longitudinal study

Reem Khamis-Dakwar (khamis-dakwar@adelphi.edu), Iman Salam
(Imansalam@mail.adelphi.edu), Amanda Nagler (Amandanagler@mail.adelphi.edu)
Adelphi University
Karen Froud (kf2119@columbia.edu)
Columbia University

Arabic is the second most represented language at homes across the U.S. spoken by 2.6% (129,400) of all K-12 English learner students in the U.S. Many second generation Arab Americans are referred to as heritage language learners (HLLs). HLLs are usually exposed to Arabic as their native language in early childhood, however their exposure to and use of Arabic lessens or ceases later (Albirini, Benmamoun, & Saadah, 2011). The study of the effects of the functional shift in language use in HLL children is an emerging new field (Benmamoun, Montrul, & Polinsky, 2016; Montrul, 2012) and the effects are still not fully understood.

Studies on language acquisition in HLLs commonly report the presence of incomplete language acquisition or attrition due to reduced exposure to L1, mostly affecting grammatical proficiency. It is hypothesized that with reduction of L1 exposure and use, two outcomes might result: HLLs may either reorganize their linguistic system to look like that of monolingual English speakers, or they may maintain their L1 settings.

There are few studies on language development in Arabic heritage speakers in general and in Arabic heritage speaking children in particular. One study by Khamis-Dakwar, Ahmar & Froud (2019) examined production of Arabic dual, regular and irregular plurals in Arabic and English in 21 typically developing Arabic HLL children aged 5; 04 to 9; 11 using an elicited production task. The findings revealed better performance on English plural elicitation tasks compared to Arabic. In the Arabic elicitation tasks, typically developing HLL children showed superior performance on production of dual and feminine plural compared to masculine and broken plural forms. These findings were interpreted as revealing the systematic interrelationships between L1 loss and L2 learning in the course of language acquisition in heritage language learner contexts, and perhaps L1 maintenance of early-acquired features since loss of the dual and feminine plural forms was not exhibited. Moreover, this study shows that examination of the dual plural could have utility as a marker for typical L1 acquisition in clinical speech and language evaluations of heritage speakers, who are mandated to be examined in both their L1 and L2. However, such theoretical and applied applications were constrained by the methodological cross-sectional design of the study, making it difficult to determine the course of maintenance through the different stages of language exposure in HLLs.

In this presentation, we describe a two-year longitudinal examination of Arabic plural and dual

productions and English plural productions in two female siblings, starting one month after they immigrated to the U.S. The children were aged 3;01 and 5;06 at the start of the study, with 2;9 and 5;1 years respectively of Arabic exposure in an Arabic majority speech community. The two children joined an English-only schooling system one month post their immigration to the U.S. The elicitation task includes 120 items including singular, dual, feminine and masculine regular plurals, and irregular plurals in real words and pseudowords. The English items include 60 items targeting production of English irregular and regular plurals, half real and half pseudowords in each category.

Results to date reveal effects of age of immigration, child's age, levels of language exposure and use, and languages and structures assessed. As expected, a systematic interrelationship between L1 attrition and L2 learning was exhibited. Moreover, the Arabic plural formation in these two HLLs correlated with the developmental trajectory observed in Arabic monolingual acquisition. That is, masculine and broken plural (known to be acquired later than dual and feminine plural) were not acquired for the younger child and were lost for the older one. The feminine plural was maintained for the older child, but not the younger one. Dual formation was maintained for real words but not generalized to pseudowords. Moreover, homeland visits correlated with intensive exposure to Arabic and changes to Arabic performance, supporting a process-oriented approach rather than an emphasis on end results in the study of linguistic competence in HLLs (e.g. Putnam & Sanchez, 2013).

References

- Albirini, A., Benmamoun, E., & Saadah, E. (2011). Grammatical Features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition*, 33(2), 273–303.
- Albirini, A., Benmamoun, E., & Chakrani, B. (2013). Gender and number agreement in the oral production of Arabic Heritage speakers. *Bilingualism*, 16, 1, 1-18.
- Benmamoun, E., Montrul, S., & Polinsky, M. (2010). *White paper: Prolegomena to heritage linguistics*. Retrieved from <http://nhlrc.ucla.edu/pdf/HL-whitepaper.pdf>
- Khamis-Dakwar, R., Ahmar, M., & Froud, K. (2019). Maintenance of Arabic dual formation in typically developing heritage speakers of Arabic. In Haddad, Y., & Khalfaoui, A. (Eds.), *Perspectives on Arabic Linguistics XXXI. Papers from the annual symposium on Arabic Linguistics Norman, Oklahoma*. (pp. 207-233). Amsterdam: John Benjamins.
- Montrul, S. (2012). Is the heritage language like a second language? *EUROSLA Yearbook*, 12, 1-29.
- National Center for Education Statistics (2019). *Digest of education statistics 2018. Table 204.27. English language learner (ELL) students enrolled in public elementary and secondary schools, by home language, grade, and selected student characteristics: Selected years, 2008-2009 through fall 2016*. Retrieved from https://nces.ed.gov/programs/digest/d18/tables/dt18_204.27.asp
- Polinsky, M. (2016). Bilingual children and adult heritage speakers: The range of comparison. *International Journal of Bilingualism*, doi:10.1177/1367006916656048.
- Putnam, M.T., & Sanchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478-508.

Are Arabic Listeners 'Stress Deaf' to their Own L2 Pronunciation?

Iman Albadar, University of Delaware
iman@udel.edu

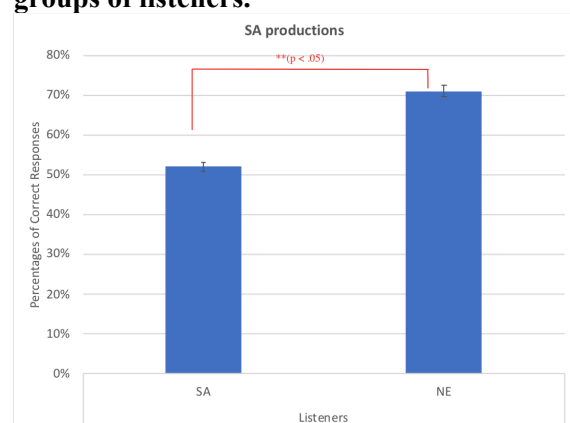
Second Language (L2) learners' ability to perceive L2 word stress is proposed to be determined by whether they have encoded stress in their L1 phonological representations [3]. Since stress in Arabic is predictable (i.e. not encoded in their phonological representation), Arabic listeners have been found to be "stress deaf", in that they are insensitive to L2 stress properties compared to speakers whose L1s are with unpredictable stress like English [4]. However, differences between L1 and L2 prosodic properties can also influence L2 learners' perception of stress (e.g., [5]). Arabic is quantity-sensitive

in which stress is usually assigned to a heavy syllable. The acoustic cues for Arabic stress include mainly F0 and less importantly duration, where duration is also used for phonemic vowel and consonant distinction [6]. This raises the question of whether and to what extent, the prosodic properties of learners' native language would affect the perceptibility of stress and stress deafness? The present study aims at answering this question by 1) investigating the performances of Arabic learners of English and English native speakers in perceiving English stress produced by Arabic L2 speakers (between-group), and 2) by comparing Arabic learners' perception of English and Arabic L2 speakers' productions (within-group).

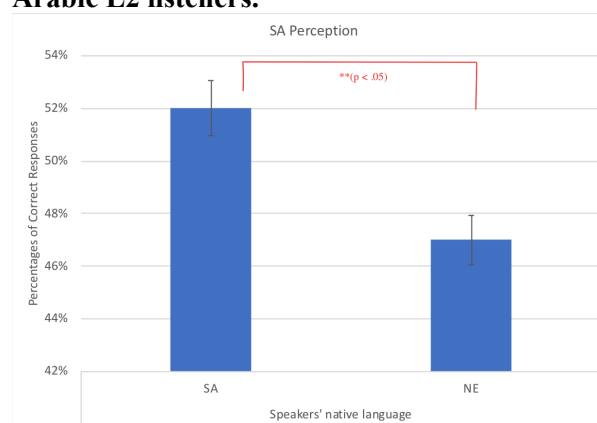
Ten advanced Saudi Arabic (SA) learners of English and ten native English speakers (NE) (avg. age: 25.4 years) performed an identification task, in which they listened to 180 English nonce three-syllable words and had to identify the stressed syllable. These words were produced by other NE (90 words) and SA (90 words) learners of English who knew unambiguously where stress was placed (stressed syllables were marked, e.g. zoubawsa (1st syllable should be stressed)). All the tested words contained only open syllables with one consonant in onset position. The syllables were classified as either light (weak W) or heavy (heavy H) depending on the kind of vowel (i.e. a light syllable contains lax vowel, and a heavy syllable contains diphthong or tense vowel). The position of stress was systematically varied in the different types of words in which one heavy syllable was stressed in each word, and if there is another heavy syllable in the same word, it will be a secondary stress. The test responses were collected from each participant and accuracy (correct responses) was determined according to whether the identified stress is the intended stress as pronounced by the speaker.

The results revealed that SA listeners sustained stress deafness even when stress was produced from speakers with the same native language (i.e. SA). As shown in [1], SA had significantly more difficulties in identifying the stressed syllable than NE, confirming previous results [4]. However, the effect of prosodic properties was found when comparing SA's responses to word stress produced by SA with those produced by NE. The accuracy rates of SA's responses showed a better performance in perceiving SA's L2 production of stress than in perceiving NE's productions [2]. These results suggested that although listeners' experience with prosodic cues does enhance sensitivity to stress perception, it is not robust enough to surmount stress deafness.

[1] Percentages and standard error of correct responses to L2 Arabic productions by the two groups of listeners.



[2] Percentages and standard error of correct responses to L1 and L2 Arabic productions by Arabic L2 listeners.



References

Peperkamp, S. & Dupoux, E. (2002). A typological study of stress 'deafness'. In C. Gussenhoven & N. W arner (eds.): *Laboratory Phonology 7*: 203-240. Berlin : Mouton de Gruyter.

- Altmann, H. (2006). *The Perception and Production of Second Language Stress: A Cross-linguistic Experimental Study*. Ph.D. Dissertation, University of Delaware.
- Tremblay, A. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics*, 29, 553-584.
- Vogel, Irene & Athanasopoulou, Angeliki & Pincus, Nadya. (2016). Acoustic Properties of Prominence and Foot Structure in Jordanian Arabic.

The acquisition of word order in early Moroccan Arabic
Anna Gavarró and Iman El Hadeff Hammad
Universitat Autònoma de Barcelona
anna.gavarró@uab.cat; iman.elhadeff@e-campus.uab.cat

In this paper we investigate the acquisition of word order in young children exposed to Moroccan Arabic, by means of an act-out task. The word orders tested are exemplified in (1) and include SVO (1a), OSVcl (1b) and VclS (1c).

- (1) a. L-begra dŋ-at d-džāža.
D cow push past.f D hen
‘The cow pushed the hen.’
b. L-begra, d-džāža bās-t-ha.
D cow D hen kiss past.f f.sg.cl
‘The cow, the hen kissed her.’
c. Tbŋ-at-ha d-džāža.
chase past.f f.sg.cl D hen
‘The hen chased her.’

The aim of the experiment reported is to test the hypothesis by Friedmann et al. (2009) according to which children follow a stricter version of Relativised Minimality (RM) in sentence comprehension, to the effect that dependencies in which lexical objects move across lexical subjects are misinterpreted. The claim of these authors is that child RM can account for the asymmetry attested in the acquisition of many languages (including Palestinian Arabic) between subject and object relatives – with object relatives being less well understood than subject relatives. While the results for relative clause comprehension seem robust across languages, the results from studies on the acquisition of topicalisation are less consistent (see, for example, Hu et al. 2018, based on results from Chinese). The prediction for Moroccan Arabic if child RM holds would be that (1a) would be better understood than (1b), since the object moves across the subject in (1b), an instance of object topicalization (Ennaji et al. 2004); for (1c), since the object is pronominal and not lexical, one might expect the same performance as for (1a).

The experiment was administered to 40 children (10 2-year-olds, 10 3-year-olds, 10 4-year-olds and 10 5-year-olds, age range: 2;02,06–5;11,11) and 10 adults as control group. The experimental items that the children had to act-out were 15 reversible sentences corresponding to the word orders above, either gender-matched or gender-unmatched subject and object. The sentences were presented in discourse contexts that rendered them felicitous.

The adult control group performed at ceiling. The results for the children appear in the table and show that children are at ceiling at age 5, but even before that they perform at very high levels. Although no statistical analysis has been performed yet, one can observe an age effect, with 2-year-olds performing below 3-year-olds, and 3-year-olds performing below 4-year-olds. The percentage of non-correct answers corresponded mainly to theta role reversal and, in the youngest group, to non-answers.

	SVO	OSVcl	VclS
2-year-olds	71.6%	70%	72%
3-year-olds	86.6%	85%	86%
4-year-olds	95%	90%	96%
5-year-olds	100%	100%	100%

Percentage correct performance, children

As for the theoretical implications of the study, for no age group did we find a contrast between sentence types: SVO was not better understood than object topicalization, and postverbal subjects preceded by a V+object clitic were understood at the same rates. No difference emerged either as a consequence of the matched vs. mismatched gender of subject and object in (1b), while some work shows that gender mismatch boosts comprehension of object relative clauses in Hebrew (see Belletti et al. 2012). These are, to our knowledge, the first results on the acquisition of word order in Moroccan Arabic; they indicate that young children’s comprehension is much more adult-like than predicted by child RM.

References

- Belletti, A., Friedmann, N., Brunato, D. & Rizzi, L. (2012) Does gender make a difference? Comparing the effect of gender on children’s comprehension of relative clauses in Hebrew and Italian. *Lingua*, 122: 1053–1069.
- Ennaji, M., Makhoukh, A., Es-Saiydy, H., Moubtassime, M. & Slaoui, S. (2004) *A Grammar of Moroccan Arabic*, Pars Lettres 25, Fes.
- Friedmann, N., Belletti, A. & Rizzi, L. (2009) Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119: 67-88.
- Hu, S., Guasti, M. & Gavarró, A. (2018) Chinese children’s knowledge of topicalization: Experimental evidence from a comprehension study. *Journal of Psycholinguistic Research*, 47(6): 1279-1300.

Multiple Origins, Convergence, and the Earliest Maghribi Varieties: The Case of Šibilant-Merging Adam Benkato, UC Berkeley abenkato@berkeley.edu

Descriptive work on Maghribi Arabic varieties, particularly varieties spoken by Jewish communities, has noted a particular phenomenon, usually depicted as the assimilation of the sibilants *š* and *ẓ* to *s* and *z*, or vice-versa. Pointing out that this phenomenon is not a synchronic process but a diachronic one, Heath refers to it as “sibilant-merging”: the complete phonemic merger of *s* with *š* and *z* with *ẓ*. In such varieties, there is only one voiced and one voiceless (non-emphatic) sibilant, occupying a place of articulation ranging from alveolar [s,z] to pre-palatal [ɕ,z] to palato-alveolar [ʃ/ʒ] depending on the variety.

Sibilant-merging is not, however a speech characteristic particular to Jewish Maghribi dialects, but is also reported for the dialects of the Egyptian oases of Farafra and Baḥariyya, as well as certain varieties of the southern Hijaz in the Arabian Peninsula. All reported cases of sibilant-merging are depicted on the adjacent map.

In this paper, first give a detailed overview of sibilant-merging in all dialects for which it is reported. I then argue that it may useful to view the phenomenon from a broader, historical linguistic perspective. I suggest that its presence in “archaic” dialects of northern Africa (those usually grouped as “pre-Hilali” and attributed to the first layer of Arabic dialects there) *and* in dialects of the Arabian Peninsula means that it may be an **innovation** shared by these dialects rather than unrelated individual developments. This innovation may have taken place in varieties of the Arabian Peninsula prior to their dispersal into northern Africa.

I then compare this feature with several other features in Maghribi dialects held to be archaic retentions or early shared innovations, especially retention of a voiceless reflex *t̥* for usual *d* and *ð* and near-total loss of *h.

Feature	Sibilant-Merging	*š / *ṭ > <i>t̥</i> (not <i>d</i> / <i>ð</i>)	*h > <i>ø</i> (or <i>h̃</i>)
Dialects	Rabat, Meknes, Fes ^M ; Tafilalt, Sefrou ^J (Morocco); Algiers ^J (Algeria); Tunis, Gabes, Djerba ^J (Tunisia); Farafra, Baḥariyya ^M (Egypt); south Hijaz (Arabia)	Tangier, Tetuan, Branes, Mtiou ^M ; Tafilalt, Sefrou ^J	Algiers ^J ; Tripoli ^J (Libya); Maltese

		(Morocco); Jijel ^M (Algeria)	
--	--	--	--

Looking at how these features overlap, I argue for reconsidering the canonical narrative of the two-wave Arabization process of northern Africa, consisting of two largely homogeneous dialect waves. I suggest that the initial scenario is more likely to have consisted of **multiple original dialects**, and their post-dispersion interactions in northern Africa may have resulted in the **convergence** of some features. Using sibilant-merging as a case study, combined with the few other studies of shared phonological rarities among first-layer dialects in northern Africa, I propose that a more careful assessment of their features, without attempting to first relate them to the canonical narrative and while integrating work on polygenesis and other language origin models, can better help us to understand the history of Arabic in the region.

Subgrouping Arabic Dialects via Historical Developments in Closed-Class Morphemes
Alexander Magidow, University of Rhode Island
amagidow@uri.edu

The classification of Arabic dialects has traditionally been undertaken via a qualitative approach focusing on a small number of isoglosses, especially the realization of certain phonemic variables such as the reflexes of /*q/, /*ʒ/, etc. (see e.g. Taine-Cheikh 1998; Palva 2006). In addition to focusing strongly on phonological isoglosses, many of the distinguishing features used to classify dialects are actually genetic retentions, not innovations. While phonemic variables are extremely important sociolinguistically, they are also subject to rapid change in the course of a small number of generations. Morphological elements on the other hand tend to be better preserved over time since they are relatively deeply integrated into a dialect, are less likely to be borrowed, while also changing more slowly over time (Hetzron 1976). This study uses closed-class morphemes – pronouns, demonstratives and interrogatives – to provide historically motivated subgroupings of Arabic dialects, and to analyze the importance of these subgroupings for our understanding of the history of Arabic.

This study uses a database of dialect material from 82 dialects. Through historical analysis of the development of the pronouns, demonstratives and interrogatives, 55 historical isoglosses were identified (for example, the lengthening of the first syllable in the pronoun *a:na* ‘I’, the realization of the 3ms suffix pronouns with vowels /u/, /a/, /i/, etc.). In this presentation, I will briefly discuss a small number of the potentially more controversial isoglosses, but the focus will be on the identification of dialect subgroupings. One contribution of this study is that it provides a new set of isoglosses for dialect classification, whether synchronic or diachronic.

Dialects were grouped using two methods – first they were grouped manually, using a cross-tabulation of percentages of dialects with a given feature. This method identified three major dialect groups: a North African dialect group, a Peninsular Bedouin group, and Sedentary Levantine group. Following this, the matrix of innovations was used as input to a Bayesian phylogenetic algorithm, which provided support further support for the North African group, and identified two further groups – a Southern Arabian group which includes Andalusian and Classical Arabic, and a Syrio-Iraqi group, of which the Sedentary Levantine group is a subgroup. This Syrio-Iraqi group also includes Central Asian dialects, reflecting existing intuitions in the literature (Zimmermann 2006; Ingham 2006). See the Figure 1 for a map of these groups.

These groupings call into question some important beliefs in the field of Arabic dialectology. For example, the North African group includes dialects on both sides of the supposed Pre-Hilalian/Hilalian dialect division and indeed this is the single most cohesive dialect group, suggesting that our narrative for the dialect history of North African should be reconsidered. Similarly, the most common innovations found in the study suggest that the Southern Arabian group is not the most

proximate ancestor of most modern dialect, and instead is a very early branch away from the main ancestors of most Arabic dialects today.

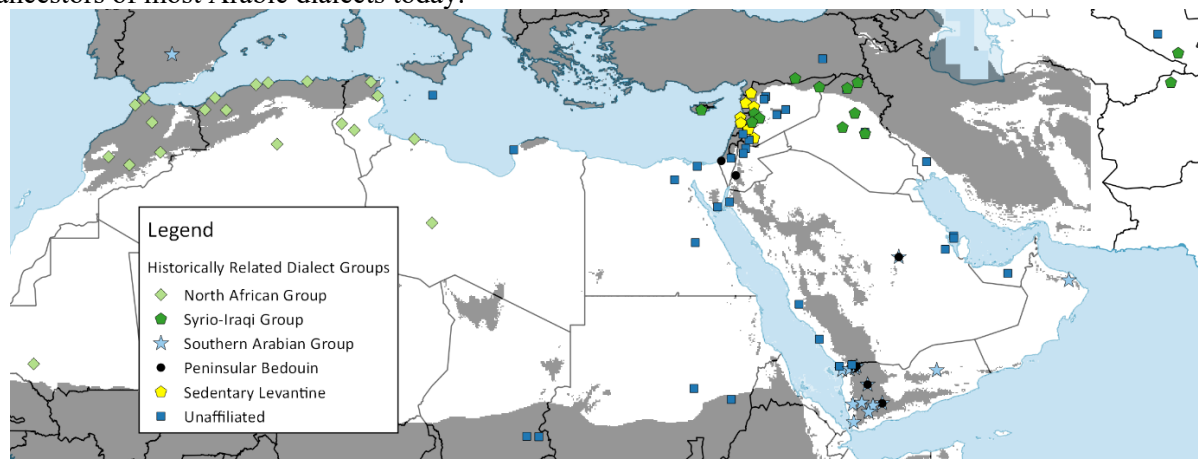


Figure 1: Map of major dialect grouping determined by manual comparison and phylogenetic methods. Note that multiple symbols on the same dialect represent membership in multiple groups.

References

- Hetzron, Robert. 1976. Two Principles of Genetic Reconstruction. *Lingua* 38. 89–108.
- Ingham, Bruce. 2006. Afghanistan Arabic. In Kees Versteegh (ed.), *Encyclopedia of Arabic Language and Linguistics*, vol. 1, 28–35. Online edition. Leiden: Brill.
- Palva, Heikki. 2006. Dialects: Classification. In Kees Versteegh (ed.), *Encyclopedia of Arabic Language and Linguistics*, vol. 1, 604–611. Leiden: Brill.
- Taine-Cheikh, Catherine. 1998. Deux macro-discriminants de la dialectologie arabe (la réalisation du qâf et des interdentes). *atériaux arabes et sudarabiques (GELLAS)* 9. 11–51.
- Zimmermann, Gerit. 2006. Uzbekistan Arabic. In Kees Versteegh (ed.), *Encyclopedia of Arabic Language and Linguistics*, vol. 4, 612–623. Online edition. Leiden: Brill.

Models of Root Co-Occurrence Restrictions and Their Implications for Arabic Phonology

Laura R. Faircloth, *The University of Texas at Austin*

lrfaircloth@utexas.edu

BACKGROUND Arabic has a system of roots and patterns, where each root generally consists of three consonants, $C_1C_2C_3$ (Holes, 2004; Ryding, 2005; Watson, 2007). There are limits on adjacent consonants at the root level, which have been analyzed several different ways. Greenberg & Al-Ani (1950) focused on the restrictions against homorganic clusters of laryngeal, pharyngeal, post-velar, velar, liquids, sibilant, inter-dental, dental, and labial consonants. McCarthy's (1988) feature geometric analysis relied on primary and secondary place features but also incorporated the relationship between primary place and sonorance. Alternatively, the co-occurrence restrictions can be analyzed as a diachronic process of dissimilation. The Proto-Semitic analysis relies on similarities between reconstructed place and manner (Huehnergard, 2019). However, the proposed analyses of co-occurrence restrictions do not accurately account for all of the observed patterns (McCarthy, 1988; Pierrehumbert, 1992). This study compared three models of the co-occurrence restrictions, indicating the strengths and weakness of each and providing directions for future analyses of the consonantal patterns. If the co-occurrence restrictions are more clearly understood, the phonological analyses of Arabic based on these claims will be improved.

ANALYSIS 3770 tri-consonantal roots from Lane's (1876) dictionary were divided into C_1C_2 and C_2C_3 pairs and the observed frequency of each pair tallied. The expected frequencies for each pair were based on the total occurrences of each sound in the pair (Greenberg & Al-Ani, 1950). Differences between these frequencies would indicate root co-occurrence restrictions. Three models were constructed to predict the observed patterns of co-occurrence based on the expected frequency and the

hypothesized restrictions. **Model 1:** The Greenberg & Al-Ani (1950) analysis compared two models: one with only primary place and one with place, voicing, manner, and secondary place. **Model 2:** The McCarthy (1988) model had primary and secondary place features as well as an interaction between primary place and sonorance of consonant pairs. **Model 3:** The Proto-Semitic Model used the features of reconstructed sounds from Huehnergard (2019).

RESULTS There was a significant difference between the observed and expected frequencies of the consonant pairs ($p < 0.0001$), suggesting that there are co-occurrence restrictions on adjacent root consonants. **Model 1:** The model with multiple features was significantly better ($p < 0.0001$), suggesting that primary place alone does not account for the co-occurrence restrictions. This analysis correctly prohibited labial /b f m/ co-occurrence but allowed adjacent gutturals such as /k x q ʔ h ʕ ʕ/, though these consonants rarely co-occur in roots. **Model 2:** Interactions between factors, such as place and continuant, significantly improved the model ($p < 0.0001$), supporting the inclusion of other factors. However, it did not capture asymmetries such as the occurrence of /ʃ/ with alveolar /t tʰ d/ but not with /dʕ/. **Model 3:** The effect of primary place was significant ($p = 0.024$) and improved the analysis of co-occurrence of fricatives and affricates. Arabic /ʃ/ and /dʕ/ rarely co-occur, likely because they had the same reconstructed place and manner as lateral fricatives, /*ɬ/ and /*ɮ/ respectively. The feature geometric model ($p < 0.0001$) had a significantly better fit (demonstrated by a lower AIC) to the data (AIC = 4191.7) than the Greenberg (AIC = 4564.7) or Proto-Semitic (AIC = 4432.3) models.

DISCUSSION Root consonants with fewer common features were the most likely to co-occur, demonstrated by Model 2. The co-occurrence restrictions must include more than a prohibition against adjacent homorganic consonants. An accurate analysis will improve the related models of the underlying phonological structure of Arabic, such as which consonants pattern as gutturals (McCarthy, 1988). The gutturals /k x q ʔ h ʕ ʕ/ rarely co-occur, suggesting that they form a natural class to the exclusion of the emphatic consonants /tʰ sʰ ʔ dʕ/ (Herzallah, 1990; Zawaydeh, 2003), possibly to avoid sequences of similar sounds that are difficult to perceive (Frisch, Pierrehumbert & Broe, 2004). The co-occurrence restrictions are a part of the grammar of modern speakers (Frisch & Zawaydeh, 2001), so these co-occurrence restrictions may be related to the articulatory, acoustic, and perceptual characteristics of different consonants (Wright, 2004). The comprehensive analysis of the co-occurrence restrictions will incorporate perceptual factors to further clarify natural classes, improving the understanding of the phonological system of Arabic.

References

- Frisch, Stefan A., Janet B. Pierrehumbert & Michael B. Broe. 2004. Similarity Avoidance and the OCP. *Natural Language & Linguistic Theory* 22(1). 179–228.
doi:10.1023/B:NALA.0000005557.78535.3c.
<https://doi.org/10.1023/B:NALA.0000005557.78535.3c>.
- Frisch, Stefan A & Bushra Adnan Zawaydeh. 2001. The Psychological Reality of OCP-Place in Arabic. *Language* 77(1). 91–106.
- Greenberg, John & Salman H. Al-Ani. 1950. The patterning of root morphemes in Semitic. In *Readings in Arabic Linguistics*, Indiana University.
- Herzallah, Rukayyah S. 1990. *Aspects of Palestinian Arabic Phonology: a non-linear approach*. Ithaca, NY: Cornell University Dissertation.
- Holes, Clive. 2004. *Modern Arabic Structures, Functions, and Varieties* Georgetown Classics in Arabic Language and Linguistics. Washington DC: Georgetown University Press.
- Huehnergard, J. 2019. Proto-Semitic. In J. Huehnergard & N. Pat-El (eds.), *The Semitic languages*, 51–80. London: Cambridge University Press. <https://doi.org/10.4324/9780429025563>.
- Lane, Edward William. 1876. *An Arabic-English Lexicon: Derived from the Best and the Most Copious Eastern Sources*. <https://lanelexicon.com/updates/>.
- McCarthy, John J. 1988. Feature Geometry and Dependency: A Review. *Phonetica* 43(45). 84–108.
- Pierrehumbert, Janet Breckenridge. 1992. Dissimilarity in the Arabic Verbal Roots. *Proceedings of NELS*

- Ryding, Karin C. 2005. *A Reference Grammar of Modern Standard Arabic*. New York, New York: Cambridge University Press.
- Watson, Janet. 2007. *The Phonology and Morphology of Arabic* Phonology of the World's Languages. Oxford, United Kingdom: Oxford University Press.
- Wright, Richard. 2004. A review of perceptual cues and cue robustness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically-based Phonology*, 34–57. Cambridge, UK: Cambridge University Press.
- Zawaydeh, Bushra Adnan. 2003. The interaction of the phonetics and phonology of gutturals. In J Local, R. Ogden & R Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI*, 278–292. Cambridge: Cambridge University Press.

Morphosyntactic derivation of the Form VIII verb: stuck in the middle again

Samantha Wray, New York University, Abu Dhabi

samantha.wray@nyu.edu

Of the Arabic measures—morphosyntactic paradigms—none has flummoxed those who seek to explain their phonological shape like Form VIII, hence McCarthy’s appeal to an “Eighth Binyan Flop” (1979, 1981) to account for the emergence of the infixal /t/. Within Distributed Morphology, various accounts for this (such as Tucker 2011, Zukoff 2017) are consistent with an explanation that predicts the pattern and CV skeleton as an emergent property at *Spell-out*, the phase at which syntactic structures are joined with phonological features (Chomsky 1993 inter alia).

Zukoff (2017) proposes that the infixal /t/ of Form VIII is the same /t/ as the prefix in Form V taCVCVC and the prefix in Form X staCVCVC. Its emergence as either prefix or infix is determined by the morpheme-ordering *Mirror Alignment Principle (MAP)* which appeals to relationships of c-command in the word’s internal structure. When we move beyond Zukoff’s analysis of Modern Standard Arabic, this is borne out by data from Cairene (a) and Moroccan (b), shown in (1):

- (1) Form VII with ‘t’ a. tkasar ‘was broken’ (Watson 2002)
 b. tharaq al-xubz ‘the bread was burnt’

These Form VII “t” verbs have been suggested as a likely historical merger between Forms VII and VIII (Watson 2002). Under Zukoff’s proposal, Form VII is a middle verb with [+mid] in an identical hierarchical position to [+refl]: it is in asymmetrical c-command with root. Therefore, following MAP, a middle-voice /t/ could indeed emerge as a prefix, not an infix.

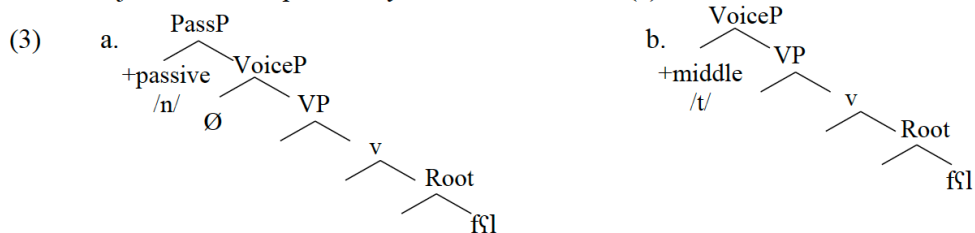
However, this does not explain the emergence infixal /t/ in Yemeni (a) or Gulf Arabic dialects, including Emirati (b), shown in (2). These verbs do not have a reflexive reading, which Zukoff’s MAP would necessitate in order to have the /t/ emerge as an infix. Rather, they also are in middle voice:

- (2) Form VIII a. htarag as-siyarra ‘the car was burnt’
 b. ftayal at-telfizyawn ‘the TV worked’ (Al Kaabi 2015)

I solve this in the following ways:

- Form VII verbs (and the Form VII “t” variants in (2)) are not middle voice; they are passive. Form VII’s designation as passive has been suggested before (c.f. Watson 2002)
- Following Bruening’s (2013) analysis for English, borrowed by Kastner (2016) for Hebrew: I posit that the passive is not a variant of Voice; rather there is a separate Pass phrase above Voice
- Swapping of adjacent components, true-middle /t/ and the root, in *Spell-out*, using Local Dislocation (Embick & Noyer 2001, henceforth LD), again utilized by Kastner (2016) for Hebrew

This proposal results in confirmation of the structure proposed for Form VIII (b) by Al Kaabi (2015), with a major revision as previously noted for Form VII (a):



As seen in (X), as [+passive] is in its own phrase in Form VII, it is not adjacent to the root (/n/-Voice-root), therefore they are not able to be swapped via LD, leading to the emergence of the Form VII /n/ (or /t/) as a prefix in *Spell-out*. Contrast these with the structure in (3b), which shows adjacency (/t/-root), permitting a swap of these elements via LD. Along with phonological constraints determining a left-aligned root (c.f. McCarthy & Prince 1993), this predicts the emergence of an infixal /t/ in Form VIII.

References

- Al Kaabi, Meera. 2015. *The Anatomy of Arabic Words: The Role of the Root in Representations and Processing*. Doctoral Dissertation, New York University.
- Bruening, Benjamin. 2013. By phrases in passives and nominals. *Syntax*, 16.1:1-41.

- Chomsky, N. 1993. *A Minimalist Program for Linguistic Theory*, MIT occasional papers in linguistics:1-67.
- Embick, D., & Noyer, R. 2001. Movement operations after syntax. *Linguistic inquiry*, 32(4):555-595.
- Kastner, Itamar. 2016. *Form and meaning in the Hebrew verb*. Doctoral Dissertation, New York University.
- McCarthy, John J. 1979. *Formal Problems in Semitic Phonology and Morphology*. Doctoral dissertation, MIT.
- McCarthy, John J. 1981. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry*, 12:373–418.
- McCarthy, J. J., & Prince, A. 1993. Generalized alignment. In *Yearbook of morphology 1993*, 79-153. Springer, Dordrecht.
- Tucker, Matthew A. 2011. The Morphosyntax of the Arabic Verb: Toward a Unified Syntax-Prosody. In *Morphology at Santa Cruz: Papers in Honor of Jorge Hankamer*. University of California, Santa Cruz: Linguistics Research Center.
- Watson, J. C. 2002. *The phonology and morphology of Arabic*. Oxford University Press on Demand.
- Zukoff, S. 2017. Arabic Nonconcatenative Morphology and the Syntax-Phonology Interface. In *NELS*(47).

The possible root of consonantal root effects in Arabic: Evidence from neurophysiological data

Ali Idrissi^a aidrissi@qu.edu.qa

Eiman Mustafawi^a eimanmust@qu.edu.qa

Tariq Khwaileh^a tariq.khwaileh@qu.edu.qa

John Drury^b john.drury@stonybrook.edu

^a *Qatar University*

^b *Jiangsu Normal University*

Transposed letter priming (jugde-JUDGE) and other orthographic/form priming found in Indo-European languages (Forster et al. 1987; Ferrand & Grainger 1992; Perea & Rosa 2000; Brysbaert 2001) have not been reliably found for Hebrew and Arabic (Frost et al. 2005; Valen & Frost 2009, 2011; Perea et al. 2010). Coupling this with demonstrations of Semitic root priming (Frost et al. 2000; Boudelaa & Marslen-Wilson 2005) has led to the suggestion that lexical memory for Hebrew/Arabic may be qualitatively different in organization. However, non-word primes derived from real word targets by single root letter replacement have been shown to yield priming in Arabic (Perea et al. 2014), dovetailing with literature questioning the status of the consonantal root as a morphological unit of lexical organization in Semitic (Idrissi 2018).

We used an ERP masked priming paradigm in which we manipulated prime duration (40/120 ms) between-participants and examined six conditions: (i) identity; (ii) root/meaning identity ([Saliib - maSluub] “cross – crucified”); (iii) root without meaning identity ([Saliib - Salaaba] “cross - hardness”), (iv) transposed real root orthographic overlap ([Saliib - baSal] “cross - onions”); (v) semantic relatedness ([Saliib – qasaawisa] “cross - pastors”); (vi) unrelated prime-target pairs. Condition-(iv) involved either local/adjacent root letter transpositions (e.g., tri-consonantal roots with 123 order preceded by real-word primes with identical consonants in 213 or 132 orders), or non-adjacent transpositions (e.g., 123 targets with 321, 312, or 231 primes). Real root primes like these have been previously found not to yield priming in Arabic (Perea et al. 2010; note other studies have used non-word primes).

Trials consisted of #-marks (#####; 500 ms), prime (40 or 120 ms), target (300 ms; using a larger font than the primes), and response/blink prompts. The task was go/no-go semantic categorization: half of the pairs (180) used names of animals/objects. Participants responded only if they saw animal name. Critical trials (30 items in each of (i)-(vi) = 180 critical pairs) did not require behavioral response. EEG was continuously recorded from 24 scalp electrodes (250 Hz EEG; 25 Ag/AgCl electrodes; Ground: AFZ; left-mastoid reference, re-referenced to linked mastoids offline. EEG pre-processed with a 0.1–30 Hz BPF). ERPs were examined within a 1000 ms epoch time-locked to prime

onset (-200 to 0 ms baseline). Preliminary results reported here are based on 29 Qatari Arabic native-speaker adult participants (N=15 for short-prime (40 ms); N=14 for long-prime (120 ms)).

Our results show that when prime duration was short (40 ms) priming reduced ERP amplitudes for all conditions except semantics (v). Importantly, this includes our root letter transpositions/(iv), contra expectations based on most previous literature. With long prime exposure (120 ms) all conditions (including semantics/(v)) showed priming but with differences effect size. Identity showed the largest effects, followed by root priming (ii/iii), with semantic primes and root-transpositions (iv/v) showing (equivalently) the smallest effects. We argue (1) that there may not, in fact, be such qualitative crosslinguistic differences with respect to orthographic priming, and (2) previously observed Arabic consonantal root effects in masked priming may be only due to orthographic rather than morphological overlap.

Audiovisual integration in spoken Arabic: The role of lexicality and syllabic affiliation

Michael Andrew Grosvald(mgrosvald@qu.edu.qa)

Noora Al-Ansari(na1601018@student.qu.edu.qa)

Ali Idrissi(aidrissi@qu.edu.qa)

Qatar University

Speech perception involves both auditory and visual information, as exemplified by the so-called McGurk illusion (McGurk & MacDonald, 1976) arising from a conflict between the auditory and visual signals, in which the listener perceives a fusion of the two streams, implying that information from both has been integrated. When observing cross-spliced videos showing a person saying /ga/ with an overdubbed /ba/, many people perceive a ‘fused’ /da/. The alveolar /d/ percept appears to be a compromise between /b/, produced at the front of the articulatory space and /g/, produced near the rear. This effect constitutes powerful evidence that multi-modal processing of speech occurs at a relatively early stage of perception.

Research on this phenomenon indicates that lexicality plays a role in audiovisual integration. Specifically, more fusion responses were observed when the fused percept coincided with a word than with a non-word (Brancazio, 2004). An effect of syllabic affiliation on audiovisual integration has also been shown. In a study with English stimuli, Ali (2003) reported greater fusion rates for coda than onset consonants. However, in a study comparing Arabic and English, Ali et al. (2005) found little difference in fusion rates amongst onsets and codas in Arabic, suggesting that the nature of Arabic syllabic representations might be different. The present research using Qatari Arabic is one of extremely few studies of the McGurk effect in any variety of Arabic (but see Alsalmi, 2016; Scott & Idrissi, 2018), in which we further explore the effects of lexicality and syllable affiliation on the McGurk illusion.

Our study was conducted with 22 native or early users of Qatari Arabic, which has all of /b/, /d/, and /g/ in its phoneme inventory. The 24 target stimuli of form C1VC2VC3 varied along the dimensions of (1) lexicality and (2) consonant position in the word (initial, medial and final). Factor (1) involved four possible combinations of audio/video and fused percept lexicality. In the first of these, ‘RRF’, real-word video with a ‘g’ sound was overdubbed with real-word audio with a ‘b’ sound, for which fusion would result in a pseudo-word (‘fake’ word) with a ‘d’ sound (i.e. ‘Real’ video input + ‘Real’ audio input leads to ‘Fake’ pseudo-word percept, hence ‘RRF’). The other three types were FFR (‘fake’ pseudo-word video and audio leading to a real-word percept), FFF, and RRR. As for (2), consonant position, our 24 target stimuli were divided into three groups of eight stimuli each, with the b/d/g sound either in C1, C2 or C3 position. To disguise the nature of the target stimuli, 48 filler stimuli were also included (evenly divided into real and fake Qatari words). Subjects performed two tasks after viewing each stimulus: lexical decision (“Was this item a ‘real’ or ‘fake’ Qatari word?”), followed by a 3-way forced-choice sound discrimination task. The sequence of 72 trials was separated into 3 randomized blocks of 24 trials each (8 targets and 16 fillers), one for each of C1, C2 and C3 sound discrimination.

Findings from our data include the following. Overall, auditory capture (AC, hearing a ‘b’) happened 17% of the time, visual capture (VC, hearing ‘g’) happened 44% of the time, and McGurk fusion (F, hearing ‘d’) happened 39% of the time. With respect to consonant position within the stimulus: the number of majority-F subjects (i.e. subjects who perceived McGurk ‘fusion’ on a majority of trials) was 17 for C1, down to 10 for C2, and just one for C3. The corresponding numbers for VC were 3, 8 and 19, and for AC, those numbers were 2, 4, and 2. In other words, fused percepts were less and less common at later consonant positions, and visual capture percepts more and more so. Auditory capture was overall uncommon. A significant lexicality effect was also found, in that fusion was much more likely to happen if the result was “real word from fake” (i.e. “FFR”), than vice versa (“RRF”) ($p < 0.001$).

This study reports original data from a relatively little-studied dialect of Arabic, adding to the crosslinguistic evidence for the McGurk effect. It also explores, in a novel way, the influences of lexicality and consonant position and syllabic affiliation on fusion rates. These results do not seem to coincide with those reported by Ali et al. (2005) and suggest, instead, that distance from the word onset rather than syllabic affiliation is at play in determining the degree of audiovisual integration and the nature of the percept. We discuss these results in light of the literature on the McGurk phenomenon and speech perception, in general.

Production and Perception of Consonant Clusters in Nonwords by Iraqi and Najdi Speakers
Saja Albuarabi (saja@uwm.edu) & Hanyong Park (park27@uwm.edu)
University of Wisconsin – Milwaukee

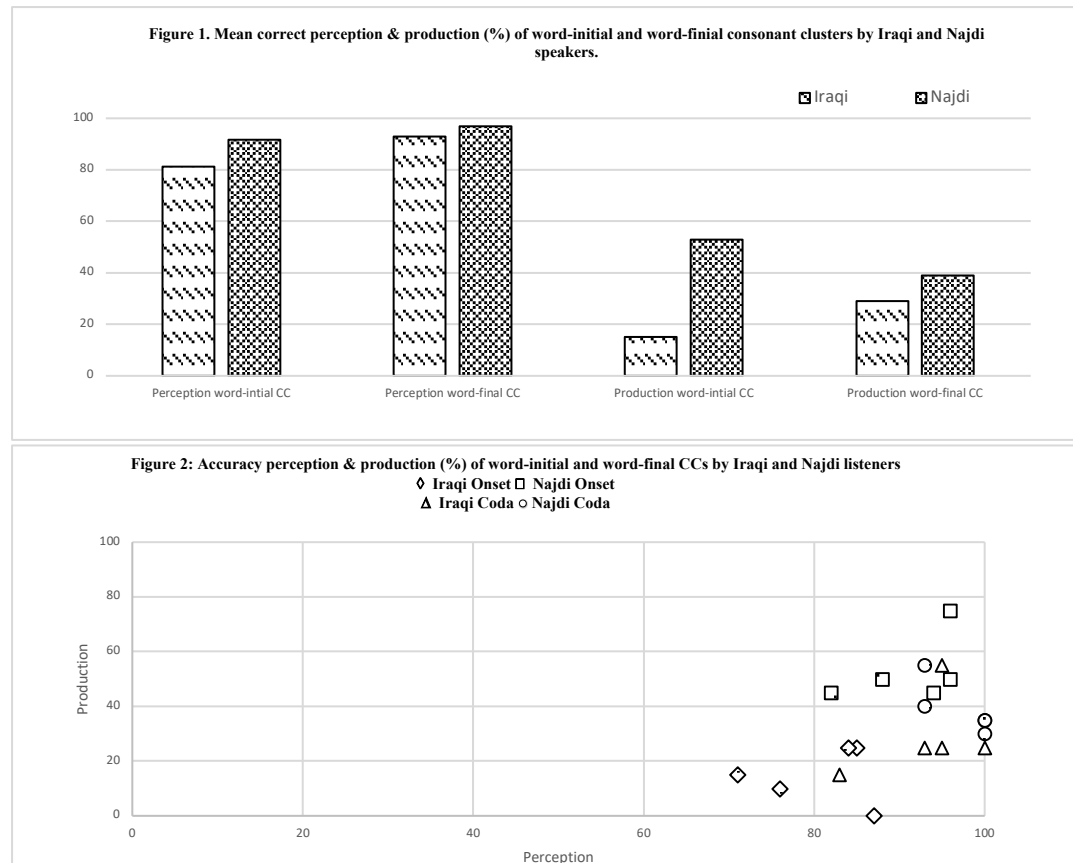
This study has a two-fold purpose: comparing two dialects of Arabic, Iraqi and Najdi, with nonwords and examining the entire phonological system (perception and production) of these two dialects. With these purposes in mind, we investigated the production and perception of word-initial and word-final consonant clusters (CCs) by Iraqi and Najdi speakers whose phonotactic constraints differ regarding these. It has been argued that Iraqi speakers have difficulty in producing CCs in word-initial and word-final positions (Blanc, 1964; Erwin, 2004; Al-Bazi, 2006), whereas, Najdi speakers do not have such difficulty (Alghmaiz, 2013; Alkhonini, 2014; Salem, 2014). However, these results are based on different methods (e.g., Iraqi studies without any empirical data vs. Najdi studies based on Najdi speakers’ production of Arabic words) and it is rare to consider the perception of CCs. Therefore, we investigated the production and perception of CCs of both Iraqi and Najdi speakers, using the same method (i.e., the same set of stimuli and elicitation procedure) for a better comparison of these two dialects. Importantly, we used nonwords, hoping to reveal hidden phonological constraints which may not be obvious by examining only lexical items in the language.

Native speakers Iraqi ($N = 5$) and Najdi ($N = 5$) dialects participated in both perception and production tasks. In the AX discrimination task (perception), after listening to a list of nonword pairs ($N = 152$), the participants decided whether the two paired items were the same or different. The pairs consisted of CC[a]CV-C_vC[a]CV, CC[a]CV-ʔ_vCC[a]CV for word-initial sequences and CVCC-CVC_vC for word-final ones. In the reading task (production), we had the participants read a list of nonsense words in the form of CC[a]CV for word-initial sequences and CVCC for word-final ones. The list in Arabic script included 10 CC types that have been argued to be allowed in Najdi Arabic. Each type had two stimuli, resulting in a list of 20 word-initial and 20 word-final CCs. We examined the accuracy rates for their perception and production. The accurate production rates were based on spectrogram examination.

Figure 1 summarizes the production and perception results. Production results replicated previous findings for the Iraqi dialect: the Iraqi speakers had difficulty in producing word-initial CCs. However, the Iraqi speakers pronounced some word-final CCs without much difficulty. Although the Najdi speakers were better than the Iraqi speakers, they also had difficulty in producing CCs, contrary to the previous findings. Perception results were more surprising in that both Najdi and Iraqi speakers were good at discriminating CCs from non-CCs (e.g., C_vC, ʔ_vCC, etc.) in both word-initial and word-

final positions. We also observed a relationship between perception and production; the one with a higher perception score tends to show a higher production score and vice versa (see Figure 2).

Our study with nonwords suggests that the absence of certain CC sequences in the lexicon does not necessarily indicate phonological constraints on those sequences. This may explain the discrepancy between our results with nonwords and previous results with real words. Also, our study clearly demonstrates that we should examine both perception and production to understand the entire phonological system; the perception system is not identical to the production system.



Reference

- Al-Bazi, M. P. (2006). Iraqi dialect versus standard Arabic. *United States: Medius Corporation*.
- Alghmaiz, B. A. (2013). *Word-initial consonant cluster patterns in the Arabic Najdi Dialect*. Southern Illinois University Carbondale.
- Alkhonini, O. A. (2014). *Coda consonant cluster patterns in the Arabic Najdi dialect*. Southern Illinois University at Carbondale.
- Blanc, H. (1964). *Communal dialects in Baghdad*. Cambridge, MA: Harvard University Press.
- Erwin, W. M. (2004). *A short reference grammar of Iraqi Arabic*. Georgetown University Press.
- Salem, S. T. (2014). *The modification of English/s/+ consonant onset clusters by Levant Arabic speakers*. Lethbridge, Alta.: University of Lethbridge.

Decomposing Arabic Complementizers

Saleem Abdelhady, smabdelhady@mun.ca

Phil Branigan, branigan@mun.ca

Hussein Al-Bataineh, hhalbataineh@mun.ca

Memorial University of Newfoundland, Canada

This study investigates the internal structure of Arabic declarative complementizers (Cs). Most studies have looked at the role of Arabic Cs in forming complex constructions and how case-assignment takes

place (Ahmed, 2015; Jarrah, 2019b; Ouhalla, 1994; Shlonsky, 1997; Soltan, 2006). However, comparing the C systems of Standard Arabic (SA) with Jordanian Arabic (JA) and other varieties of Arabic shows that the systems are more complex than previously stated. The complexity can be realized clearly with the use of clause-initial Cs.

- (1) *ʔinna al-banāt-i waṣal-na* (SA) (2) **ʔanna al-banāt-i waṣal-na* (SA)
 COMP the-girls-ACC arrived-3F.PL
 ‘Indeed, the girls arrived.’ COMP the-girls-ACC arrived-3F.PL
 ‘Indeed, the girls arrived.’
- (3) **ʔinn-u ʔil-walad ʔiḏa* (JA) (4) **ʔanno ʔil-walad ʔiḏa* (LA)
 COMP-3M.SG the-boy.ACC arrived COMP the-boy.ACC arrived
 ‘Intended: the boy arrived.’ ‘Intended: the boy arrived.’ (Jarrah, 2019b)

The examples above show that, in SA, two forms of Cs are used in the system *ʔinna* and *ʔanna*. In JA, nevertheless, only one form is used *ʔinn*. A notable difference between the two systems is that, in SA, it is acceptable to have the C in clause-initial position, yet the C must be *ʔinna*. The use of *ʔanna* leads to ungrammatical construction (2). However, the ungrammaticality of using *ʔanna* intersects with the ungrammaticality of using Cs in the same position in JA; the use of the C *ʔinn* is unacceptable, which makes it behave like the C *ʔanna* in SA. Ross (1970) attributes the behavior to the existence of a hidden performative verb of saying, whereas Fassi Feheri (2012) claims that the distinction is based on embedding contexts, overlooking the fact that both Cs can be embedded. Looking at data from JA questions the validity of previous conclusions (6).

- (5) (*ʔaqulu*) *ʔinna al-banāt-i waṣal-na* (SA) (6) *bagūl ʔinnu ʔil-bānt ʔiḏ-ū* (JA)
 (say.I) COMP the-girls-ACC arrived-3F.PL say.I COMP the-girls.ACC arrived—
 3F.PL
 ‘(I say) that the girls arrived.’ (Ross, 1970) ‘I say that the boy arrived.’ (Jarrah, 2019b)

The use of the C *ʔinn* along with an explicit verb of saying is acceptable in JA. However, assuming that the verb of saying is hidden does not lead to grammatical structure in this variety (compare (3) with (6)). Comparing the two systems highlights the following pressing research questions: (a) Why does the hidden performative verb hypothesis only apply to SA but not JA Cs? And (b) How is the SA C system different from that of Arabic varieties? If not, how do they fit the universal pattern? The study argues that Cs in Arabic are complex morphemes that are composed of nano-syntactic features that affect their size and function as triggers of presupposition (Baunaz & Lander, 2018; Baunaz, 2018; Starke, 2009). Syncretic Cs cannot be used without explicit verbs since their use results in uncontrolled presuppositions – as the case of JA. SA, on the other hand, licenses *ʔinna* but not *ʔanna* because *ʔinna*, unlike *ʔanna*, is not syncretic. The analysis tests constructions across factive and non-factive verbs. To arrive at their functional sequence (7), we compare Arabic Cs across different languages (Baunaz, 2018). The following table illustrates and orders Cs in away to avoid ABA patterns (Caha, 2009).

	Emotive F1	Factive	Semi-factive F2	Nonfactive F3	Desiderative F4
Modern Greek	pu		pu/oti	oti	na
SA	ʔanna		ʔanna	ʔinna	ʔan
Sason Arabic	le		le	le	ta
JA	ʔinn		ʔinn	ʔinn	ʔinn
Lebanese Arabic	ʔanno		ʔanno	ʔanno	ʔanno
French	que		que	que	que
Serbo-Croatian	da		da	da	da

The distribution is argued to pour out from the Cs internal semantic composition (Baunaz, 2018). Avoiding *ABA patterns (Caha, 2009) makes the following ordering an optimal option for the functional sequence of Arabic Cs.

(7)	SA	JA	LA	SnA
a. [F1P F1]	=> /ʔan/	/ʔin/	/ʔənnə/	/ta/
b. [F2P F2 [F1P F1]]	=> /ʔinna/	/ʔin/	/ʔənnə/	/le/
c. [F3P F3 [F2P F2 [F1P F1]]]	=> /ʔanna/	/ʔin/	/ʔənnə/	/le/
d. [F4P F4 [F3P F3 [F2P F2 [F1P F1]]]]	=> /ʔanna/	/ʔin/	/ʔənnə/	/le/

As a trigger of different degrees of presupposition (veridical > nonveridical), factivity affects Arabic Cs. Cs can be used in clause-initial positions, only if their use is not syncretic and enables the recoverability of presupposed information without intersection. In addition to supporting Baunaz' (2018) conclusion, the study shows how a cognizant point of view can affect veridicality and the selection of Cs cross-linguistically.

References

- Ahmed, H. (2015). *Verbal Complementizers in Arabic*. The University of Utah, United States.
- Baunaz, L., Lander, E. (2018). Deconstructing categories syncretic with the nominal complementizer. *Glossa: Journal of General Linguistics*, 3(1).
- Caha, P. (2009). *The nanosyntax of case* (Ph.D.). University of Tromsø.
- Jarrah, M. (2019a). Factivity and subject extraction in Jordanian Arabic. *Lingua*, 219, 106–126.
- Jarrah, M. (2019b). Record your Agree: A case study of the Arabic complementizer *ʔinn*. *Journal of Linguistics*, 55(1), 83–122.
- Lena, B. (2018). Decomposing Complementizers: The functional sequence of French, Modern Greek, Serbo-Croatian, and Bulgarian Complementizers. In L. Baunz, K. De Clercq, L. Haegman, & E. Lander (Eds.), *Exploring Nanosyntax*. Oxford University Press.
- Ouhalla, J. (1994). Verb movement and word order in Arabic. In D. Lightfoot & N. Hornstein (Eds.), *Verb Movement*.
- Ross, J. (1970). On Declarative Sentences. In R. Jacobs & P. Rosenbaum, *Readings in English Transformational Grammar* (pp. 22–227). Ginn and Company.
- Shlonsky, U. (1997). *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. Oxford University Press.
- Soltan, U. (2006). Standard Arabic subject-verb agreement asymmetry revisited in an Agree-based minimalist syntax. In C. Boeckx (Ed.), *Linguistik Aktuell/Linguistics Today* (Vol. 92, pp. 239–265).
- Starke, M. (2009). Nanosyntax A short primer to a new approach to language. *Nordlyd* (6) 1-10.

The syntax of confirmationals in Omani Arabic (OA)

Mohamed Jlassi, Sohar University

MJlassi@su.edu.om

Speakers of OA massively tag their utterances with the Discourse Marker (DM) *la:h*, which is in no way similar to the negative *la:* (no). In each of the utterances, *la:h* conveys the speaker's (S) calling on the hearer (H) for a confirmation of some type. In (1), S requests H to confirm S's prior knowledge and belief of H's visit to her friend.

- (1) ʔinti zurt-i ʂadi:q-tek el-ʔusbu:ʕ el-ma:ði:, **la:h** ?
 you visit-prf.2sf friend-cli (her) the-week the-last-gen, DM
 “Did you visit your friend, didn't you?”

In (2), S, knowing H, uses *la:h* to check out on some guess about H's possible action.

- (2) ʔawel marra t- uʔubx-i ha:ði e-ʔʔabxa, **la:h**
 first time ipfv-cook-2sf this the- food-acc, DM
 “It is the first time you cook this food, right?”

In (3), S is asserting a fact about a situation, expecting a confirmation from H.

- (3) la:, xala:s la: ʔa:t-i, naħn qum- na: b-el-wa:jeb w ʔakØer, **la:h**

no, over not worry-2sm/f we do-prf.1pm/f with-the-duty and more, DM
 “It’s over! No worries, we did the necessary; what do you think?”

In (1-3), S sets up through *la:^h* a process of negotiation to trigger H’s confirmation of the assertion conveyed by the utterance. In this respect, *la:^h* in OA is a DM that mediates between the speech participants, their beliefs about themselves, about the world and the content of the utterance. The DM is then at the edge between the world of discourse and the world of utterance. Such a process is known as grounding (Bavelas et al. 2012). Structurally, the DM is a right-peripheral constituent that does not seem to scope on the proposition or affect its truth values. It is not either sensitive to sentence typing given that (1), for example, remains interrogative in the presence of *la:^h*. From that position, *la:^h* grounds the root clause in the discourse situation where the interlocutors negotiate the assertions in each of (1-3). The relationship between the discourse effect of the DM and the utterance is best mediated by syntax. In the literature, a rising interest in studying the syntactic encoding of speech acts and more generally discourse has lately characterized the syntax-discourse interface literature. Noticeably, this has been counted in Wiltschko et al. (2014:1) as a ‘neo-performative view’ resurrecting Ross’s (1970) Performative Hypothesis. The tenet of this research tendency is that speech acts as expressed by DMs can be syntacticized into higher domain predicative structures that map discourse into syntax (Speas and Tenny 2003, Munaro and Poletto 2009, Hill 2013, Haegeman and Hill 2013). Relevant to the present talk, two opposite views can be distinguished. Speas & Tenny’s (2003) view of the topmost structure as a Speaker-oriented two-position layer; Wiltschko et al. (2014)’s and Hill (2006) view of the two-position layer to be Addressee-oriented. Contra to both, the talk claims that the higher structure is a three-position layer. A first projection bears no S or H orientation. It is about the inception of the idea of discourse in the mind of the speaker including the speaker’s thoughts of the topic of discourse including the speaker. It is the Ground (Grnd) projection that sets up the initial link between the world of discourse as schematized in the speaker’s mind and the utterance. A second projection is S-oriented and encodes the Speaker’s predictions of H’s possible reactions to the utterance. It is the Common Ground (CGrnd) projection where S devises the strategy of the utterance in terms of calling on H, considering the context, setting the tone and preparing responses. The third projection is Common Ground Management (CGrndM) where S manages the Common Ground to spell-out the utterance triggering H’s confirmation. This third position is H-oriented. Consider (4).

(4) [Grnd.....[CGrnd.....[CGrndM..... [Force.....[Fin....]]]]]

(4) will account for the distinct types of confirmational in OA in view of their variety in (1-3), reflecting their syntactic complexity. Basically, confirmation is one, but it is the object of confirmation that varies. In terms of derivation, to derive the right-periphery DM *la:^h* the talk argues that ForceP moves to the specifier position of Grnd according to the Clause Fronting Hypothesis (Munaro & Poletto 2003; Haegeman, 2014; and Wiltschko, 2014). The central contention will be that syntacticizing confirmational in OA is crosslinguistically valid evidence for syntacticizing DMs and discourse in general. Positing a higher fine-grained functional layer through the syntax of confirmational in OA is also evidence for the argument in generative Grammar that ‘the functional hierarchy derives in part from non-linguistic cognition’ (Wiltschko, 2014:2), and that syntacticizing discourse helps the understanding of ‘the ‘structure and functioning of the cognitive systems at the interface with the syntactic module’ (Cinque & Rizzi 2010:63).

Notes on Najdi Arabic (NA) Scope Taking

Mohammed Abuhaib, Imam Muhammad ibn Saud Islamic University & University of Wisconsin-Milwaukee
mabuhaib@uwm.edu

This work investigates some issues for scope taking, and distributivity of universal and existential quantifiers in NA. In this dialect, sentences with multiple quantifiers are limited to the surface scope interpretations only, while inverse readings are essential for complex DPs (construct state, CS). Regarding Arabic, Fehri (2018) proposes that it is similar to English in that it resorts to LF movements (Quantifier Raising (QR) May, 1977) which contribute ambiguous scope readings (surface & inverse). In contrast, my claim is that NA scope taking resembles German

(Bobaljik & Wurmbrand, 2012:381) where the scope rigidity is witnessed only in the clause level, but inverse linking is permitted at LF for DPs. As indicated, English is a QR language and LF scope ambiguity can be seen in the following¹:

- (1) **A boy climbed every tree.** ($\exists > \forall$; $\forall > \exists$)

As shown, the sentence above is ambiguous due to the shown possibility of scope taking and the indefinite co-variation. On the other hand, the scope rigid languages do not show this ambiguity since their quantifiers order is fixed. Specifically, the scope is determined by the narrow syntax due to the flexibility of their word order (scrambling) as in Hungarian, Japanese and German (Bobaljik & Wurmbrand, 2012; Szabolcsi, 2010; Beghelli & Stowell, 1997). Consider the following example from Japanese:

- (2) Dareka-ga subete-no hon-o yonda. ($\exists > \forall$)(? $\forall > \exists$)
 someone-NOM all-GEN book-ACC read
 “Someone read all the books.”

(Bobaljik & Wurmbrand, 2012: 374)

I. NA Clause Level Scope

The following points support the suggested claim about the scope rigidity at sentence level:

Inverse Scope: For NA, the below sentence cannot be construed as its English counterpart in (1) since the order of the quantifiers is fixed:

- (3) tasal:q **walad** (#muxtalif) **kul:** **fadžarah** ($\exists > \forall$) ($? \forall > \exists$)
 climbed **boy** (#different) **every** *indf.* **tree**
 “one boy climbed every tree.”

The only way for deriving the inverse scope readings is by a leftward syntactic movement:

- (4) **kul:** **fadžaratin** tasal:aq-ha **walad** ($? \exists > \forall$) ($\forall > \exists$) **Clitic Left Dislocation**
every **tree** climbed-it *indf.* **boy**
 “every single tree was climbed by one boy (boys vary with trees)”

Negation Scope: several sentences can be derived syntactically for representing (\forall, \exists , negation $^-$)scope interactions with one predicate. In the following example, the subject *kul:-u ta:lib-in* “every student” and the object *muṣal:im* “a teacher” are both affected by the leftward movement to the locations where they are interpreted:

- (5) (**fi:h**) **muṣal:im**_(g) **kul:** **ta:libin**_(i) ma: yuhib-ah_(g) *t*_(i) *t*_(g) ($\exists > \forall > \neg$)
 (there) **a teacher** **every student** not like-him
 “there is a specific teacher that every student does not like”

Replacing any trace with its co-indexed moved item changes the scope and the interpretations to ($\exists > \neg > \forall$), ($\forall > \neg > \exists$) or ($\neg > \forall > \exists$) without affecting the grammaticality. However, for object reconstruction, delete the clitic pronoun.

If-clause: when an indefinite is a subject of a distributive predicate like “die” in *if-clause*, it has to appear before *if* with a cliticized pronoun on the verb to obtain the specific reading of the indefinite and distributivity as:

- (6) **Ṯala:Ṯt rdza:lin min aqa:rb-i** lau: ma:t-u, ?a-badfaṣ mi?at dola:r lil-huku:mah
 ($\exists_3 > \text{if}$)
three men of relative-mine if die-**they**, will-pay-I hundred dollar to-government
 “if three specific men of my relatives die, I will pay one hundred dollars to the government”

The existential subject “three men...etc” has to appear before *if-clause* to show the specificity. if it appears after *if*, the indefinite will pick any three men relatives.

II. DP Level Scope

Wide Scope: On the other hand, the below CS causes the sentence ambiguity based on the subject nominals scope:

- (7) **sadi:q-at kul: walad** thib sai:arah ($\exists > \forall$) ($\forall > \exists$)
friend_{fem} **every boy** love *indf.* **car**
 “one single female friend for every boy loves a car” (only one female friend and one car)
 “for every boy, there is a female friend that loves a car.” (more than one female friend and one or more cars)

Binding: the embedded quantified possessor takes scope at sentence level and binds -ah “his” at LF. Still, the number of the female friends in the below sentence is affected by the scope location interactions, as in (7):

- (8) **sadi:q-at kul: walad**_(i) thib **sai:arat-ah**_(i) ($\exists > \forall$) ($\forall > \exists$)
friend_{fem} **every boy** love car-his

“one female friend of every boy loves their cars” “for every boy, there is a female friend who loves his car.”

¹ Symbol for scope representation (High>Narrow)

References:

- Beghelli, F., & Stowell, T. (1997). Distributivity and negation: The syntax of each and every. In *Ways of scope taking* (pp. 71-107). Springer, Dordrecht.
- Bobaljik, J. D., & Wurmbrand, S. (2012). Word order and scope: Transparent interfaces and the $\frac{3}{4}$ signature. *Linguistic Inquiry*, 43(3), 37-421.
- Fehri, F. (2018). *Constructing feminine to Mean: gender, number, numeral, and Quantifier Extensions in Arabic*. Rowman & Littlefield.
- May, R. (1977). *The Grammar of Quantification*, (Doctoral dissertation), MIT.
- Szabolcsi, A. (2010). *Quantification*. Cambridge University Press.

oalshehri@email.arizona.edu

Bruening, B. (2013). By-Phrases in Passives and Nominals. *Syntax*, 16(1):1-41.

Halle, M. and Marantz, A. (1993). Distributed Morphology and the Pieces of inflection. In *The View from building 20*, pages 111-176. The MIT Press.

- Halle, M. and Marantz, A. (1994). Some Key Features of Distributed Morphology. *MIT Working-Papers in Linguistics*, 21(275):88.
- Heim, I. (1982). The Semantics of Definite and Indefinite Noun Phrases. *Ph.D. Dissertation*. UMass.
- Heim, I. and Kratzer, A. (1998). Semantics in Generative Grammar. Blackwell.
- Kratzer, A. (1996). Severing the External Argument from Its Verb. In Rooryck, J., Zaring, L., et al., editors, *Phrase Structure and the Lexicon*, pages 109{137. Springer
- Marantz, A. (1997). No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your own Lexicon. *University of Pennsylvania working papers in linguistics*, 4(2):14.
- Wood, J. (2015). *Icelandic Morphosyntax and Argument Structure*, volume 90. Springer.

Conditional Clauses in Modern Standard Arabic and Jisr az-Zarqā Arabic:

A Semantic Comparison

Letizia Cerqueglini, Tel Aviv University

cerqueglini@tauex.tau.ac.il

Conditional clauses (CCs) specify hypotheses regarding something that could have taken place or will take place and consist of a protasis (condition) and an apodosis (governor). According to Comrie (1986), the oppositions among “real,” “possible”, and “unreal” conditional types do not account for the several choices observed cross-linguistically. “Hypotheticality” seems rather a continuum of degrees, from lower to higher probability, whose choice depends on “subjective evaluation” and sociocultural conventions. In general, CC types are classified formally on the basis of conditional markers and verbal patterns (Danks 2011). Different Arabic varieties have different CC systems (Al-Hilal 2017; Biṭunē 2015; Fischer 2002; Grigore 2005; 2008; Henkin 2000; Ingham 1991; Jalonen 2017; Peled 1987; Roth 2002; Sartori 2008-2009, 2010; Vanhove 2002; Wright 1986). I compare here CC types in Modern Standard Arabic (MSA) and in the Bedouin Arabic variety spoken in the coastal village of Jisr-az-Zarqā (JaZ, Central Israel). I refer to the traditional variety (TJaZ) spoken by elders over age 70. This study is based on a 260,000-word corpus of TJaZ oral narrative and daily interactions.

MSA CCs are generally sketched based on formal features, as in Fischer (2002) and Ryding (2005). Fischer differentiates between “real” and “unreal” CCs; perfect and jussive forms in the protasis indicate the reality of the utterance, without temporal implications. *Kāna* before the entire protasis marks the utterance as past. The conditional marker *ʔin* generally correlates with reality and *law* with possibility and irreality, while *ʔida* is associated with temporal conditional meanings. The apodosis may be marked by *fa* and *la*, with *la* correlating with irreality. Ryding divides possible condition (*ʔida*+ past tense; *ʔin*+ perfect or jussive) from impossible condition: *law*+ past tense (on ‘tense,’ see Cuvalay 1994).

Formally, the TJaZ distribution of verbal patterns and conditional markers is similar to that of Damascene Arabic (Jalonen 2017), with both *ʔida* and *ʔin* used for the real types. Yet TJaZ distinguishes four conditional types, two real (“more possible,” “less possible”) and two irreal (hypothetical, counterfactual), which are semantically closer to Baghdadi Arabic (Grigore 2005) and Classical Arabic (Versteegh 1991) CCs than to Damascene Arabic CCs. TJaZ widely exhibits temporal conditional CCs, which can be formally classified within the domain of real condition, as in Mardini (Grigore 2008), Sirti (Biṭunē 2015), Bedouin Peninsular varieties (Ingham 1991), and Classical Arabic (Thackston 1994). As in Baghdadi (Grigore 2005), Sirti (Biṭunē 2015) and Classical Arabic (Larcher 2000), asyndetic CCs mainly correlate with real conditional types. TJaZ protasis with *law*+ perfect indicates counterfactual CCs (Larcher 2003). Hypothetical types are marked by protasis

with *ʔida*+ perfect. *ʔin* marks only habitual real CCs, *ʔida* other real types. In TJaZ irreal types, the apodosis can be marked by: 1. inflected forms of *kān*+ the regent verb, both in the perfect; 2. *la*, for counterfactual apodosis; 3. *ʔin* in both protasis and apodosis. TJaZ CCs show strong semantic continuity with *qaltu* dialects and Classical and Quranic Arabic *facies* (Kinberg 1977; Larcher 2000, 2003; Lewin, 1970; Peled 1987, 1992; Sartori 2010, Thackston 1994; Versteegh 1991), attesting to an archaic, well preserved, conditional system, distant from the semantic models of MSA and sedentary Levantine Arabic.

References

- Al-Hilal, M. 2017. "Paratactic Conditionals in Syrian Arabic: A Study Based on Deir Ezzor Dialect", *Al-ʕArabiyya* 50. 81–103.
- Biṭunē, G. 2015. "The Conditional Clause in the Spoken Arabic of Siirt", *Folia Orientalia* 12. 63–74.
- Comrie, B. 1986. "Conditionals: A Typology". Traugott, E., Meulen, A., Snitzer Reilly, J. (eds.) *On Conditionals*. Cambridge: CUP. 79–99.
- Cuvalay, M. 1994. "On the role of "tense" in conditional sentences". Caubet, D., Vanhove, M. (eds.) *Actes des premières journées internationales de dialectologie arabe de Paris (27–30 janvier 1993)*.
- Danks, W. 2011. *The Arabic Verb: Form and Meaning in the Vowel-lengthening Patterns*. Amsterdam/Philadelphia: Benjamins.
- Fischer, W. 2002. *A Grammar of Classical Arabic*. Yale.
- Grigore, G. 2005. "Expressing conditionality in the Spoken Arabic of Baghdad". *Folia Orientalia* 47. 39–48.
- Grigore, G. 2008. "Conditional Structures in Mardini Arabic", *ZAL* 49. 63–77.
- Henkin, R. 2000. "Pragmatics of Conditionality in Negev Bedouin". Mifsud, M. (ed.) *Malta: Proceedings of the Third International Conference of AIDA*. 9–14.
- Ingham, B. 1991. "Subordinate Clauses of Time and Condition in Bedouin Dialects", *BSOAS* 54/1. 42–62.
- Jalonen, J. 2017. *Conditional Constructions in Damascus Arabic. Form and Meaning*. MA Th. University of Uppsala.
- Kinberg, N. 1977. *Hypothetical Conditionals in Arabic: A Study of law Clauses*. Unpubl. PhD. University of Michigan.
- Larcher, P. 2000. "Subordination vs. Coordination "sémantiques". L'exemple de l'arabe classique", *Annales Islamologiques* 34. 193–207.
- Larcher, P. 2003. "Les systèmes hypothétiques en "law" de l'arabe classique", *BéO* 55. 265–285.
- Lawler, J. 1975. "Elliptical Conditionals and/or Hyperbolic Imperatives", Grossman, R. et al. (eds.) *11th Regional Meeting of the Chicago Linguistic Society. University of Chicago Press*. 371–382.
- Lewin, B. 1970. "Non-Conditional if-Clauses in Arabic", *ZDMG* 120. 264.
- Peled, Y. 1987. "Conditional Sentences without a Conditional Particle in Classical Arabic Prose", *ZAL* 16. 31–43.
- Peled, Y. 1992. *Conditional Structures in Classical Arabic*. Wiesbaden: Harrassowitz.
- Roth, A. 2002. "Le conditionnel dans le parler arabe de Kormakiti (Chypre)". *AIDA 5th Proceedings*. Universidad de Cádiz.
- Ryding, K. 2005. *A Reference Grammar of Standard Arabic*. Cambridge: CUP.
- Sartori, M. 2010. "Pour une approche relationnelle de la conditionnelle en arabe littéraire moderne". *Arabica* 57. 68–98.
- Sartori, M. 2008-2009. "L'Évolution des Conditionnelles en Arabe Égyptien Contemporain", *BéO* 58. 233–257.
- Rosenhouse, J. 1984. *The Bedouin Arabic Dialects*. Wiesbaden: Harrassowitz.
- Thackston, W. 1994. *An Introduction to Koranic and Classical Arabic*. Bethesda: Ibex.
- Vanhove, M. 2002. "Conditionnelles et concessives en Arabe de Yafiṣ (Yemen)." *Sprich doch (...). Festschrift für Otto Jastrow zum 60. Geburtstag*. Wiesbaden: Harrassowitz.

- Versteegh, K. 1991. "Two Conceptions of Irreality in Arabic Grammar. Ibn Hisām and Ibn al-Hāgib on the Particule law." Larcher, P. (ed.) *De la grammaire de l'arabe aux grammaires des arabes*. BÉO 43. 77.
- Wright, W. 1951. *A Grammar of the Arabic Language*. Cambridge: CUP.

***Summarization of Moroccan Arabic News Texts Using Probabilistic Topic Modeling
for Arabic L2 Micro-learning Tasks***

Elsayed Issa
University of Arizona
elsayedissa@email.arizona.edu

Aitor Arronte Alvarez
University of Hawaii at Manoa
arronte@hawaii.edu

Background: Text summarization is the process of creating a concise and coherent summary of a longer text while preserving the meaning and the important information in the text (Allahyari et al., 2017). There are several reasons for the importance of automatic summarization of documents. Automated summaries reduce reading time, make the selection process easier, improve the effectiveness of indexing, result in less bias comparatively to human summarizers, and help in question-answering systems (Torres-Moreno, 2014). Automatic text summarization has been adopted in several studies (e.g., Douzidia & Lapalme, 2004; Froud et al., 2013; Ba-Alwi, 2015; Azmi & Al-Thanyyan, 2012; Haboush et al., 2012; Azmi & Al-Thanyyan, 2009; Belkebir & Guessoum, 2015 among others) that have used different algorithms (e.g., Rhetorical Structure Theory (RST), clustering techniques, adaboost and Latent Semantic Allocation (LSA) among others) for various purposes (e.g. Machine Translation, Information Retrieval among others). To our knowledge, none of the current studies have used Latent Dirichlet Allocation (LDA) algorithm.

Contribution and design: This paper contributes to this line of research on automated text summarization for L2 learning. It describes a text summarization method based on a different algorithm for a different purpose. It uses a Latent Dirichlet Allocation (LDA) algorithm that has not been used by any other study for L2 acquisition purpose. In other words, this paper contributes to L2 learning where summaries serve as small learning pieces that L2 learners read instead of larger documents. It is based on Probabilistic Topic Modeling and its Latent Dirichlet Allocation (LDA) algorithm as well as a sentence extraction approach to Arabic L2 micro-learning tasks. Micro-learning refers to a learning content that divides a larger topic into various small pieces of content that students can learn from. The extractive method selects and extracts the more relevant pieces or sentences than others in a longer text (Das & Martins, 2007). Using the two python packages of Scrapy and GenSim, this study was conducted in two main phases and it was tested on a Moroccan news corpus that was designed specifically for the purpose of this study.

Implementation: In the first phase, a Moroccan news corpus was designed by scraping more than 20 Moroccan newspapers. The scraping spiders and crawlers were implemented using python's Scrapy module. The corpus is a general-purpose 2 gigabytes of txt/json files, and it can be used in any other study. The corpus and Scrapy spiders are available on Github for download. Besides, a full documentation is provided on how to use Scrapy spiders to increase the size of the corpus. In the second phase, two things were done: First, the probabilistic topic model was built based on a Latent Dirichlet Allocation (LDA) algorithm, and second, the summarization algorithm based on a sentence extraction approach was implemented. It is used to discover the underlying topics in a text document or several documents. The basic assumption behind it is that a document can be represented by a set of latent topics, multinomial distributions over words, and assume that each document can be described as a mixture of these topics (Chang et al., 2009). Each document has then a set of topics and probability distributions associated with them. At the same time, each topic has a set of words and their probabilities of occurrence given that document and topic, i.e., topic models build bags for topics to extract information. In GenSim, topic modeling can be done using Latent Semantic Allocation (LSA) or Latent Dirichlet Allocation (LDA). This present system adopts LDA, which is a generative probabilistic model of a corpus. The idea is that "documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words" (Blei et al., 2003). Based

on the output of LDA, and the output of the desired summarization algorithm, the system adopts an extractive approach.

Evaluation: Taking into account that these summaries are extracted for L2 micro-learning tasks, the results of the system were evaluated using a text quality evaluation method by a human evaluator. This method provides four criteria to analyze several linguistic aspects of the text: grammaticality, non-redundancy, reference, and coherence. Grammaticality means that the text should not include any non-textual items such as punctuation errors or incorrect words, while non-redundancy indicates that redundant data should not be included. Reference clarity shows that nouns and pronouns should be clearly referred to in summary, and coherence refers to the coherent structure of the sentences (Steinberger & Jezek, 2012).

Results: A total of 100 summaries were evaluated using a 4-point scale ranging from 0 to 4 (0 = very poor, 1 = poor, 2 = acceptable, 3 = good, 4 = very good). The results of the evaluation method indicated that grammaticality and coherence scored 100% (4 points). As for non-redundancy, there are 88 sentences that scored 100% (4 points), while 11 sentences scored 3 points and one sentence scored 2 points due to some repetitions in the summarized data. A total of 63 sentences scored 4 points, 12 sentences scored 3 points, and 25 sentences scored 2 points. Since the extractive method extracts the most relevant sentence, nouns, pronouns and anaphors represent a problem. The extracted sentences included anaphors that refer to other antecedents in preceding sentence(s).

Limitations: The system has several imperfections that can be addressed in future research. First, more evaluation measures and human evaluators will be involved. Second, more improvements to the preprocessing phase of the raw data will be implemented to improve the performance of the topic model. Third, more improvements will be introduced to the summarization algorithm to overcome the errors caused by nouns, pronouns and anaphors in the summarized sentences. Finally, an abstractive method can be adopted to provide machine-generated summaries instead of extracting sentences. However, this method involves more complex technical implementations.

Code is available on GitHub via the following links:

The Moroccan news corpus and the Scrapy spiders: <https://github.com/Elsayedissa/The-Moroccan-News-Corpus>

The LDA and the summarization algorithms: <https://github.com/Elsayedissa/Arabic-News-Summaries>

References

- Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). Text summarization techniques: a brief survey. arXiv preprint arXiv:1707.02268.
- Azmi, A., & Al-Thanyyan, S. (2009, September). Ikhtasir—A user selected compression ratio Arabic text summarization system. In 2009 International Conference on Natural Language Processing and Knowledge Engineering (pp. 1-7). IEEE.
- Azmi, A. M., & Al-Thanyyan, S. (2012). A text summarizer for Arabic. *Computer Speech & Language*, 26(4), 260-273.
- Ba-Alwi, F. M., Gaphari, G. H., & Al-Duqaimi, F. N. (2015). Arabic text summarization using latent semantic analysis. *Current Journal of Applied Science and Technology*, 1-14.
- Belkebir, R., & Guessoum, A. (2015). A supervised approach to arabic text summarization using adaboost. In *New Contributions in Information Systems and Technologies* (pp. 227-236). Springer, Cham.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems* (pp. 288-296).
- Das, D., & Martins, A. (2007). A survey on automatic text summarization. literature survey for language and statistics. II Course at CMU.

- Douzidia, F. S., & Lapalme, G. (2004). Lakhas, an Arabic summarization system. Proceedings of DUC2004.
- Froud, H., Lachkar, A., & Ouatik, S. A. (2013). Arabic text summarization based on latent semantic analysis to enhance arabic documents clustering. arXiv preprint arXiv:1302.1612.
- Haboush, A., Al-Zoubi, M., Momani, A., & Tarazi, M. (2012). Arabic text summarization model using clustering techniques. World of Computer Science and Information Technology Journal (WCSIT) ISSN, 2221-0741.
- Steinberger, J., & Ježek, K. (2012). Evaluation measures for text summarization. Computing and Informatics, 28(2), 251-275.
- Torres-Moreno, J. M. (Ed.). (2014). Automatic text summarization. John Wiley & Sons.

The Acquisition of Verbal-Derivational Patterns by Moroccan HS in France

Amal El Haimeur, University of Kansas

amalelhaimeur@ku.edu

This study focuses on Moroccan Arabic as a heritage language (henceforth HL) in France and aims at contributing to the understanding of the linguistic outcomes of the acquisition of Arabic as a HL in an immigrant context. This study investigates the acquisition of verbal-derivational patterns by Moroccan HS. Verbal derivational processes are characterized by the use of non-concatenative morphological processes, and semantic notions such as causativeness and reciprocity are lexicalized within the MA verb-pattern system. The studied patterns are the basic (P1, CCEC), causative (P2, CeCCeC), medio-passive (P5, tCeCCeC), and reciprocal (P6, t-CaCeC). 15 Moroccan-French participants took part in this study. The data were gathered through a production experiment.

The main finding of this study is that semantic distinction realized by pattern alternation is neutralized in the HL. Specifically, the basic pattern (P1) and periphrastic constructions were used predominantly in P5 and P6 targeted data. It is likely that the basic pattern is used as a default morphological device because it unmarked. By using P1 in P5 and P6 targeted data, morpho-semantic distinction is neutralized. The results also reveal that the basic pattern (P1) was acquired by all the participants and 40% of the participants acquired the causative (P2). The phonological form of P2 was also modified in the HL. The medio-passive pattern (P5) and the reciprocal patterns (P6) were not acquired. There were statistically significant differences among the use of the four patterns. Additionally, the findings of the experiment propose implicational hierarchies for the acquisition of verb patterns.

Less complex and less marked morphological structures characterize the HL. Adopting neutralization in verb patterns showed that HS speak a variety that is reanalyzed. Accordingly, HS in France have a distinct variety that was shaped by their linguistic experience. Implications of this study informed on a theoretical background for SLA.

The Morphosyntax of Exclamatives in Arabic: Against a clausal cartographic analysis

Hussein Al-Bataineh

Memorial University of Newfoundland

hhalbataineh@mun.ca

Saleem Abdelhady

Memorial University of Newfoundland

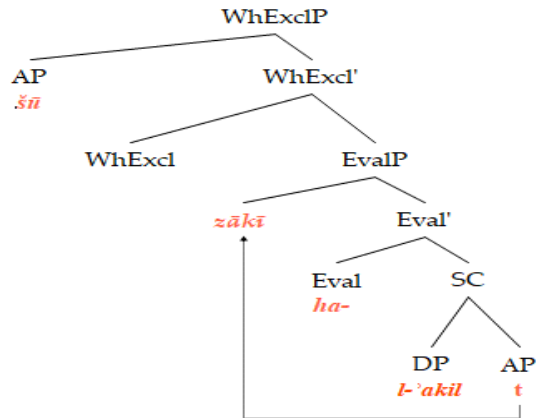
smabelhady@mun.ca

This paper investigates the morphosyntactic structure of Arabic exclamatives (Excls) and, by providing a critical assessment of debatable issues related to the clausal type, defining features, and projections at the left periphery of Excls in previous studies (Ambar, 2002; Cruschina et al., 2015; Gutiérrez-Rexach, 2001, 2008; Honda, 2011; Jónsson, 2010; Zanuttini & Portner, 2000), it claims that cartographic clausal analyses are inadequate to account for the peculiarities involved in Arabic Excls. Thus, the paper argues that an Arabic Excl, which is a non-TP construction, starts the derivation as a symmetric small clause (SC) formed from DP (the referent) and AP (the exclamative property). The SC is c-commanded by Eval(uative) head which carries [+EVALUATIVE] and [+DEGREE] features which can be checked only by the AP which moves to spec-EvalP. Depending on the type of Excl,

EvalP is c-commanded by a functional head that carries, in addition to [+EXCL], [+WH] in Wh-ExclPs in (1), [+VOC] in vocative Excls in (2), and [+V] in verbal Excls in (3).

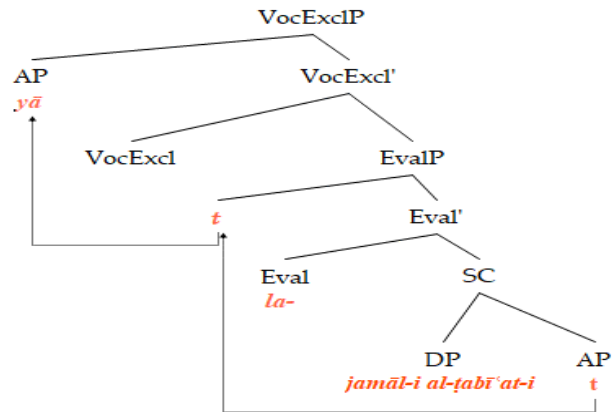
1)

*šū zākī (*ha)-l-'akil*
how delicious *(this)-the-food
'How delicious (the food is)!'



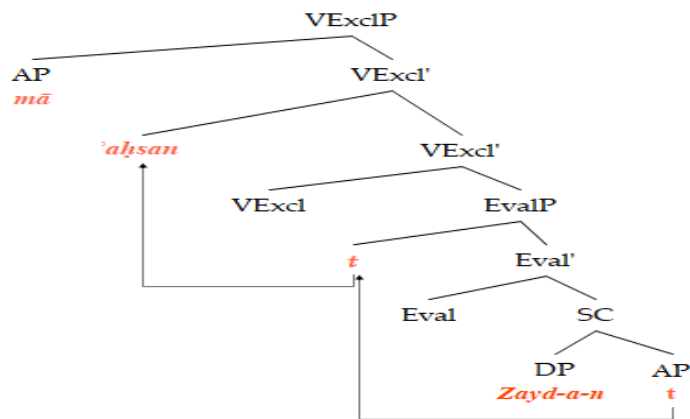
2)

yā la-jamāl-i al-ṭabī'at-i
O PREP-beauty-GEN the-nature-GEN
'How captivating the beauty of nature is!'



3)

mā 'aḥsana Zayd-a-n
PTCL excellent.EXCL Zayd-ACC-N



This approach is argued to account for the peculiarities in Arabic ExclPs such as (i) their inflexible word order, (ii) case alternation on the referent, (iii) the presence of spurious prepositions obligatorily or optionally, and (vi) the obligatory presence of some constituents which are not semantically required. This approach is supported by crosslinguistic data and claimed to have theoretical consequences that clarify peculiarities of exclamatives in other languages.

References

- Ambar, M. (2002). Wh-Questions and Wh-Exclamatives Unifying Mirror Effects. In C. Beyssade, R. Bok-Bennema, F. Drijkoningen, & P. Monachesi (Eds.), *Romance Languages and Linguistic Theory 2000: Selected Papers from "Going Romance" 2000, Utrecht, 30 November-2 December* (pp. 15–40). Amsterdam ; Philadelphia: John Benjamins Publishing.
- Cruschina, S., Giurgea, I., & Remberger, E.-M. (2015). Focus fronting between declaratives and exclamatives. *Revue Roumaine de Linguistique*, 60, 257–275.

- Gutiérrez-Rexach, J. (2001). Spanish exclamatives and the interpretation of the left periphery. In Y. D'hulst, J. Rooryck, & J. Schroten (Eds.), *Romance Languages and Linguistic Theory 1999: Selected papers from "Going Romance" 1999, Leiden, 9-11 December 1999* (pp. 167–194). Amsterdam ; Philadelphia: John Benjamins Publishing.
- Gutiérrez-Rexach, J. (2008). Spanish Root Exclamatives at the Syntax/Semantics Interface. *Catalan Journal of Linguistics*, 7(1), 117–133.
- Honda, M. (2011). Wh-exclamatives, factivity, and topicalization revisited. *Studies in Language Sciences*, (17), 87–117.
- Jónsson, J. G. (2010). Icelandic Exclamatives and the Structure of the CP Layer. *Studia Linguistica: A Journal of General Linguistics*, 64(1), 37–54.
- Zanuttini, R., & Portner, P. (2000). The Characterization of Exclamative Clauses in Paduan. *Language*, 76(1), 123–132.

The Semantic Structure of the Spatial Arabic Prepositions, fī and ‘ala

Abdessamad M'barki

University of Arizona

mbarki@email.arizona.edu

Ahmed Berrissoul

University of Mohammed V

ah.berrissoul@gmail.com

This paper approaches the polysemy of the spatial prepositions *fī* (in) and *‘ala* (on) based on principled polysemy proposed by Tyler and Evans (2003), Evans (2004, 2005, 2010). *fī* and *‘ala* play an important role in expressing how humans conceive the world they live in and how they locate objects in it, and there is a lack of application of the principled polysemy approach for Arabic prepositions.

Monosemic approaches to the semantics of prepositions assume that each lexical item is associated with a single abstract meaning (Bennett 1976, Miller and Johnson-Laird 1976). This approach fails to distinguish adequately between contextual interpretation created online and conventionalized meanings, that is, distinct meanings represented in memory. Tyler and Evans established criteria for determining distinct senses to separate between what counts as a distinct sense or contextual inference (Tyler and Evans, 2003). Monosemic approaches are too limited to convey the full scope of meaning; this is why I focus on principled polysemy.

According to the principled polysemy approach, prepositions have multiple senses based on the context they appear in. Each preposition has a network of senses, including a central sense and peripheral senses that are derived from the central sense (spatial sense). Using the principled polysemy approach, it is seen that *fī* (in) has different senses, as seen in the examples below:

1. محمد في الغرفة
Muḥammad fī al-ghorfati
Mohammad in the room
Mohammad is in the room (spatial sense)
2. يعمل محمد في الصباح
Ya ‘mal Muḥammad fī ṣ-ṣabāḥi
Works Mohammad in the morning
Mohammad works in the morning (temporal sense)
3. محمد في ورطة
Muḥammad fī warṭatin
Mohammad in trouble
Mohammad is in trouble (state sense)

The analysis in this article is based on two novels, “qiş-şato ħubin majūsiyya” and “Mawt şaghīr”. I analyzed the meanings of the prepositions *fī* and ‘ala in many contexts found in these two texts.

Arabic prepositions haven’t been studied thoroughly enough in order to allow Arabic learners to use them in an accurate and precise way to express concepts. There is a lack of clear and well principled semantic approaches that analyze the prepositions’ polysemy that will help learners use prepositions correctly. This paper will hopefully kickstart the analysis of Arabic prepositions and will help teachers, lexicographers, and learners better understand the multiple meanings of prepositions.

Bennet, D. C. (1976). *Spatial and Temporal Uses of English Prepositions: an Essay in Stratificational Semantics*. London: Logman.

Evans, V. (2004). *The structure of time: language, meaning and temporal cognition*. Amsterdam: Benjamins.

Evans, V. (2005). The meaning of time: polysemy, the lexicon and conceptual structure. *Journal of Linguistics*, 41(1), 33–75. doi: 10.1017/s0022226704003056

Evans, V. (2010). *How words mean: lexical concepts, cognitive models, and meaning construction*. Oxford: Oxford University Press.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Cambridge University Press.

Tyler, A., & Evans, V. (2003). *The semantics of English prepositions: spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press.

The Syntactic Properties of Negative Concord in Iraqi Arabic

Saja Albuarabi, University of Wisconsin – Milwaukee

saja@uwm.edu

The study investigates the complex distribution of negative concord items (NCIs) (i.e. *wəlaa* ‘no’, *ʔəbəd/nihaaʔiən* ‘at all/ never’, *bəʕd/lissəh* ‘yet’) in Iraqi Arabic (IA) with the primary goal being to provide an analysis for the licensing condition within that dialect. This study is the first such work to explicitly examine NCIs in IA. It contributes to a better understanding of the NCIs licensing in IA, and Arabic in general. Previous studies (Benmamoun, 2000, 2013; Alsarayreh, 2012) treat specific NCIs as heads project their own clausal projection; here, following Alqassas, I argue that they are Adverb Phrases (AdvPs) in specifier (Spec)-XP which can be licensed only by c-command and Spec-head relation as the head analysis fails to explain the fact where the NCIs can occur preverbally and postverbally, among other issues. The AdvPs analysis, in contrast, offers an explanation for the different locus of the NCIs. IA, in this study, is divided into two groups: *maa*-group which uses the free morpheme *maa*- and *maa-f* group which uses the enclitic *maa*- and the proclitic *-f* when negating sentences. Data from both groups show that IA is the second language that exhibits both types of NCIs beside Jordanian Arabic: strict negative concord when the NCIs must always co-occur with a negative marker postverbally or preverbally (1):

1. a. (*lissəh*) ʕəli maa saafər (*lissəh*).
(yet) Ali neg traveled._{3MS} (yet)
‘Ali has not traveled yet.’
- b. (*ʔilhissəh*) ʕəli *(maa) saafər-*if* (*ʔilhissəh*)
(yet) Ali neg traveled._{3MS}-neg (yet)
‘Ali has not traveled yet.’

And non-strict negative concord when the NCI *wəlaa* must only co-occur with the negative marker postverbally:

2. a. *(maa) idʕəh wəlaa waəhd. (*maa* group)
neg came._{3MS} no one
‘No one came.’
- b. *(maa) jaa-*if* wəlaa waəhd. (*maa-f* group)
neg came._{3MS}-neg no one
‘No one came.’

The NCI *wəlaa* cannot occur preverbally with the presence of negation as it will yield a double negation reading, and not a concordant reading:

3. a. *wəlaa waəhd maa saafər*
*maa*_{[iNeg]...}
no one neg traveled._{3MS}
‘No one did not travel.’ = ‘Everyone traveled.’
- b. [Op[¬] [_{iNeg}][NCI *wəlaa* *waəhd*_[uNeg]]] [Neg

The analysis of NCI licensing in the two groups shows that none of the previous theories of NCI licensing proposed in the literature extend to IA. For example, the syntactic agreement analysis, which states that NCIs require an agreement relation between the Neg head which is specified for an uninterpretable negation feature [iNeg] and the NCI which is specified for an uninterpretable negation feature [uNeg], does not provide an explanation for why the preverbal NCI *wəlaa* cannot occur with a negative marker in IA. Therefore, I propose an alternative account to NCIs licensing. Following Zeijlstra (2004), I argue that the preverbal NCI *wəlaa* in both groups is licensed by an abstract negative operator Op^- which occurs higher in the structure and c-command the NCI *wəlaa* (2). Thus, when the preverbal NCI *wəlaa* co-occurs with the negative marker *maa*, the result is a double negation reading (3.a). This is because both the abstract negative operator Op^- and the negative marker *maa* contain two semantics negation as shown in (3.b). I further argue that the abstract negative operator Op^- allow the NCI *wəlaa* to co-occur with a Negative Polarity Item (NPI) as it can license both the NCI and the NPI, as shown in (4):

4. a. *wəlaa waaħd ʔəkəl ʔəj ʃi*
 NCI no one ate._{3MS} NPI any thing
 ‘No one ate anything.’
 b. [Op^- [_{iNeg}] [NCI *wəlaa* *waaħd* [_{uNeg}] [VP ʔəkəl [NPI ʔəj ʃi [_{uNeg}]]]]]

Finally, the data from the *maa-ʃ* group shows that the enclitic *-ʃ* is not in a complementary distribution with NCIs (1.b) and (2.b). Unlike other Arabic dialects (i.e. Egyptian Arabic, Levantine Arabic), the enclitic *-ʃ*, in *maa-ʃ* group, does not occur with the quantifier *waaħd* when forming the negative quantifier *məħəd* as the ungrammaticality of (5) or with the NPI *ʕomr* ‘ever’ as the ungrammaticality of (6):

- | | |
|---|---|
| 5. <i>məħəd-ʃ saafər li-l-musʕl</i>
no one travelled. _{3MS} to-the-Mosul
Mosul
‘No one travelled to Mosul.’ | 6. * <i>wəlaa waaħd ma-ʕomru-ʃ saafər li-l-musʕl</i>
no one neg-ever-neg travelled. _{3MS} to-the-
Mosul
‘No one ever travelled to Mosul.’ |
|---|---|



NOTES

[illegible]