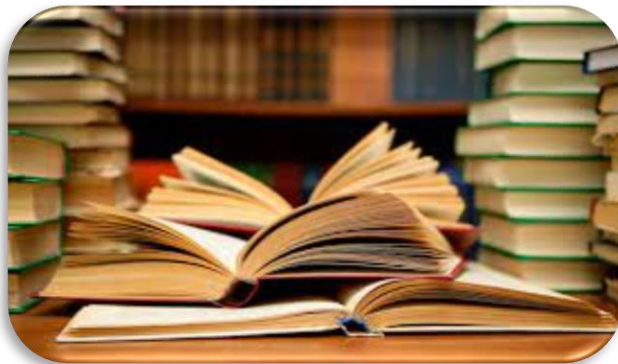## Research Paper

Constraint Interaction in Iraqi Arabic Dialects: A Focus on Baghdadi Gilit and Moslawi Qəltu through Optimality Theory

Abbas Azeez Mohammed Alabid[1], Bahram Hadian[2]*, Oudah Kadhim Abed[3], Fatemeh Karimi[4]

[1]Department of English, Faculty of Foreign Languages, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
abbasazizabbas@yahoo.com

[2]Department of English, Faculty of Foreign Languages, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
bah.hadian@khuisf.ac.ir

[3]Department of English, College of Education for Human Sciences, Al Muthanna University, Samawah, Iraq
http://mu.edu.iq

[4]Department of English, Faculty of Foreign Languages, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
Fatinaz.karimi@yahoo.com

ABSTRACT

This paper aimed at a descriptive analysis of the phonological processes in two major dialects of Iraqi: Baghdadi Gilit Dialect (BGD) and Moslawi Qəltu Dialect (MQD), within the framework of Optimality Theory (OT). Drawing on a firm corpus of 100 hours of spoken data and recent researches, 2015–2024, the present study probes into important phonological processes like assimilation, deletion, and epenthesis to explore their implications from the sociolinguistic and computational perspectives. These results show that there are significant differences in the constraint hierarchies of the two dialects, reflecting their different sociolinguistic histories. BGD has a greater number of assimilations and deletions, driven by markedness constraints such as HARMONY and PARSIMONY, which reflect Bedouin speech patterns. MQD, on the other hand, is more phonologically conservative, retaining features from Old Arabic, and thus faithfulness constraints such as MAX and IDENT-PLACE, which favor the preservation of input segments, are ranked higher. Comparisons with recent research highlight the continued influence of urbanization and Bedouinization on dialectal phonological variation. The study contributes to Arabic phonology by providing a model that bridges theoretical phonological processes with practical applications in computational linguistics, such as speech recognition and natural language processing systems. Sociolinguistically, the research highlights the complex interplay of demographic, cultural, and historical factors in shaping the phonological patterns across Iraqi Arabic dialects. The implications for Arabic dialectology, language technology, and pedagogical practices are discussed to provide an overall framework for studies that will follow in the variation of dialect.

Keywords: *Assimilation, deletion, Optimality Theory (OT), dialectal variation, Arabic dialectology, Computational Linguistics*

این مقاله با هدف تحلیل توصیفی فرآیندهای واجشناختی در دو گویش اصلی عراقی: لهجه بغدادی گیلیت (BGD) و لهجه موسلویی قلتوی(MQD) ، در چارچوب نظریه بهینه (OT) انجام شد. مطالعه حاضر با تکیه بر مجموعه ای محکم از ۱۰۰ ساعت داده های گفتاری و تحقیقات اخیر، ۲۰۱۵-۲۰۲۴، فرآیندهای واج شناختی مهمی مانند جذب، حذف، و epenthesis را بررسی می کند تا پیامدهای آنها را از دیدگاه های اجتماعی-زبانی و محاسباتی بررسی کند. این نتایج نشان می‌دهد که تفاوت‌های قابل‌توجهی در سلسله‌مراتب محدودیت دو گویش وجود دارد که نشان‌دهنده تاریخ‌های اجتماعی زبانی متفاوت آنهاست BGD. دارای تعداد بیشتری جذب و حذف است که ناشی از محدودیت‌های مشخصه‌ای مانند HARMONY و PARSIMONY است که الگوهای گفتاری بادیه‌نشین را منعکس می‌کند. از سوی دیگر، MQDاز نظر واجشناسی محافظه‌کارتر است و ویژگی‌های عربی قدیمی را حفظ می‌کند، و بنابراین محدودیت‌های وفاداری مانند MAX وIDENT- PLACE، که به حفظ بخش‌های ورودی کمک می‌کنند، رتبه‌بندی بالاتری دارند. مقایسه با تحقیقات اخیر تأثیر مستمر شهرنشینی و بادیه نشینی را بر تنوع واج شناختی گویش نشان می دهد. این مطالعه با ارائه مدلی به واجشناسی عربی کمک می‌کند که فرآیندهای واجشناختی نظری را با کاربردهای عملی در زبان‌شناسی محاسباتی، مانند سیستم‌های تشخیص گفتار و پردازش زبان طبیعی، پل می‌کند. از نظر اجتماعی-زبانی، این تحقیق بر تعامل پیچیده عوامل جمعیتی، فرهنگی و تاریخی در شکل‌دهی الگوهای واجشناختی در گویش‌های عربی عراقی تأکید می‌کند. مفاهیم مربوط به گویش شناسی عربی، فناوری زبان، و شیوه های آموزشی مورد بحث قرار می گیرد تا چارچوبی کلی برای مطالعاتی که در تنوع گویش دنبال می شود، ارائه دهد.

واژگان کلیدی: جذب، حذف، نظریه بهینه (OT)، تنوع گویشی، گویش شناسی عربی، زبان شناسی محاسباتی، آموزش کاربرد شناسی، اعمال گفتاری، یادگیری فناوری‌محور

INTRODUCTION

As a branch of Mesopotamian Arabic, Iraqi Arabic is a linguistically complex and diversified language group that has been shaped by many centuries of historical, geographical, and sociocultural influences. Two major dialects are Baghdad's Gilit Dialect (BGD) and the Moslawi Qəltu Dialect (MQD); both exhibit distinct sociolinguistic identities, expressing in their own way the processes of Bedouinization and urbanization that have typified Iraq's linguistic landscape (Blanc, 1964; Jastrow, 2006). These dialects differ significantly in phonological processes, hierarchy of constraints, and sociophonetic features, which makes them a very interesting subject of linguistic research.

The historical development of Iraqi Arabic is anchored in key sociopolitical and demographic changes. The Mongol invasions of the 13th century and subsequent Ottoman influence introduced Bedouin linguistic features to urban centers like Baghdad, giving shape to the modern BGD. Bedouinization emphasized articulatory efficiency, resulting in phonological simplifications like assimilation and deletion (Jastrow, 1994; Alshammari, 2023).

In contrast, MQD retained features of Old Arabic because of its association with settled, urban populations, particularly in Mosul. The dialectal traditional phonological system does indeed emphasize its past ties with settled groups that resisted Bedouin influence (Blanc, 1964; Salman, 2021).

Geographical factors were no less important in the formation of these dialects. Southern and central Iraq, where BGD is widespread, were subjected to Bedouin influence to a larger extent, while northern Iraq, the area where MQD is spoken, retained its own linguistic features due to a lesser degree of contact with nomadic groups (Jasim & Sharhan, 2022). Religious and ethnic compositions further solidified these boundaries: Christian and Jewish communities in northern Iraq used MQD, whereas Muslim populations in southern Iraq adopted BGD (Holes, 2007; Yussif & Mohammed, 2023).

The phonological processes attested in BGD and MQD show their different linguistic paths.
Assimilation: Predominant in BGD, assimilation follows the Bedouin influence, favouring ease of articulation over lexical clarity. It is motivated by markedness constraints like HARMONY, which spreads vowel-consonant harmony (Alshammari, 2023; Habash, 2018).
Deletion phenomena are also common in BGD, particularly in informal discourse settings, in accordance with principles such as PARSIMONY, which favour phonotactic simplicity (Salman, 2021).
Epenthesis: MQD, on the other hand, employs epenthesis to ensure phonological distinctiveness by satisfying faithfulness constraints such as DEP, which enforces structural coherence (Bourzeg, 2020).

The processes in question mark the difference between BGD and MQD, whilst at the same time each portrays the other's sociolinguistic environments. The simplification that characterizes BGD represents fast, informal speech patterns. The conservative features of MQD correspond with its traditional role as an indicator of cultural and linguistic continuity (Jastrow, 2006; Yussif & Mohammed, 2023).

Theoretical Framework: Optimality Theory

Optimality Theory (OT), as developed by Prince and Smolensky (1993), provides a robust framework within which the phonological processes involved in BGD and MQD can be analysed. OT models linguistic outputs as the result of competing constraints, which are divided into:

Markedness Constraints: Prefer less marked forms (e.g., HARMONY, PARSIMONY).

Faithfulness Constraints: Preserve input features in the output (e.g., MAX, IDENT-PLACE).

In BGD, markedness-related constraints often dominate, reflecting its Bedouin background and an emphasis on articulatory ease. By contrast, MQD has a higher ranking of faithfulness constraints, hence preserving the input's integrity due to its conservative sociolinguistic nature (Salman, 2021; Albuarabi, 2018). These rankings provide precious insights into the cognitive and articulatory strategies that underlie dialectal variation.

Research on Iraqi Arabic dialects has important theoretical and practical implications. Theoretically, the analysis of BGD and MQD contributes to a better understanding of how sociolinguistic factors impinge on constraint hierarchies and phonological variation more generally. Computationally, these results provide useful information for NLP tools to improve handling of Arabic dialectal diversity in speech recognition and synthesis (Habash et al., 2021). This provides important pedagogical insights into teaching Arabic as a second language, drawing attention to the interplay between phonology and sociolinguistics (Yussif & Mohammed, 2023). In summary, the phonological and sociolinguistic differences between BGD and MQD offer a unique lens through which to study constraint hierarchies within OT. Bridging the gap between theoretical linguistics and applied computational research on the analysis of dialectal variation, a comprehensive framework is offered based on these processes.

REVIEW OF LITERATURE

Iraqi Arabic dialects, particularly the Baghdadi Gilit Dialect and Moslawi Qəltu Dialect, have recently become the focus of increasing interest with regard to their phonological processes. This review aims to synthesize key studies into the phonological processes of these dialects, focusing on the historical, sociolinguistic, and theoretical underpinnings that ground this research. Drawing on literature from Arabic linguistics, sociophonetics, and computational studies, this review attempts to situate the present study within the broader framework of Arabic dialectology and phonology.

Iraqi Arabic, a member of the broader Mesopotamian Arabic group, is itself characterized by considerable internal variation, largely determined by historical, geographic, and sociocultural factors (Blanc, 1964; Jastrow, 1994). The two dialects that form the core of this research are the Baghdadi Gilit Dialect and the Moslawi Qəltu Dialect, representing two major variants spoken in Baghdad and Mosul, respectively. Though they both belong to the same linguistic family, they are quite different in terms of phonological structure, influenced by different social, historical, and geographical backgrounds (Jasim & Sharhan, 2022). A central area of inquiry in Arabic dialectology is how phonological processes, such as assimilation, deletion, and epenthesis, function across dialects. In Iraqi Arabic, they manifest differently in BGD and MQD, respectively, with the differences that are indicative of broader sociolinguistic divisions.

Assimilation: A frequent phonological process in Iraqi Arabic, assimilation usually involves a change in one segment to make it more similar to the adjacent segment in place or manner of articulation and voicing. In BGD, assimilation is especially common, particularly in rapid or casual speech. According to Alshammari (2023), Bedouin-influenced dialects, such as BGD, prefer assimilation as a simplification strategy, driven by markedness constraints like HARMONY. This agrees with Jastrow's 1994 findings that the articulatory ease Bedouin dialects favor often leads to phonological

simplifications. Deletion: Deletion in BGD is the omission of phonological segments. Segments like vowels or unstressed syllables are usually deleted at word boundaries, especially in informal speech or when speaking fast. Alshammari (2023) has pointed to the role of PARSIMONY in Bedouin dialects, where simplifying the phonological form by eliminating segments is favored. This preference aligns with the markedness-driven nature of these dialects.

Epenthesis: Unlike BGD, MQD presents a higher epenthesis rate to release the phonological complexity, especially in a complex consonant cluster. Bourzeg (2020) noted that in urban Arabic dialects, epenthesis plays an important role that helps in maintaining syllabic well-formedness and phonotactic harmony. MQD's reliance on epenthesis can be understood in terms of DEP-a faithfulness constraint that allows the insertion of segments to yield a form that obeys the prosodic and phonotactic requirements of the dialect. 3. Sociolinguistic and Historical Factors The historical development of the Iraqi Arabic dialects bears significance for the understanding of phonological variation. The urbanization of Baghdad and the Bedouinization process in the Ottoman period have left indelible marks on Bagdad dialect, especially on BGD. According to Jastrow (2006) and Salman (2021), the Bedouin migrations influenced the pattern of urban speech, leading to the use of Bedouin features such as fast articulation, vowel harmony, and segment deletion in BGD.

By contrast, MQD, which developed in the more sedentary and urbanized north of Iraq, retained most of the features of Old Arabic. Among these conservative features are clearer articulation, less frequent deletion, and a greater reliance on faithfulness constraints. MQD's maintenance of phonological complexity has been attributed to its urban sociolinguistic context, where phonetic clarity and precision are valued, especially in formal or public speech (Blanc, 1964; Holes, 2007).

Sociolinguistic variables such as age, gender, and social class also interact in significant ways to condition phonological processes in BGD and MQD. Habash (2018) and Yussif & Mohammed (2023) have demonstrated that younger speakers tend to prefer phonological simplifications, such as deletion, along with general trends toward language economy and informality.

These studies have also documented gendered differences in speech patterns, where women often retain more traditional phonological forms due to social pressures for linguistic purity (Albuarabi, 2018). 4. Optimality Theory (OT) in Arabic Dialects Optimality Theory, first introduced by Prince and Smolensky (1993), has become one of the central frameworks through which phonological processes in Arabic dialects have been analyzed. OT assumes that linguistic outputs are the result of an interaction of ranked constraints, where markedness constraints, which favor simpler forms, compete with faithfulness constraints, which preserve input features. This interaction explains why different dialects exhibit different phonological behaviors.

In the case of BGD and MQD, OT is an insightful model to gain a deeper understanding of the phonological process differences between them. The higher rates of assimilation and deletion in BGD can be explained by the more significant influence of markedness constraints like HARMONY and PARSIMONY, which promote simplification and ease of articulation. In contrast, the preference for epenthesis in MQD reflects the priority of faithfulness constraints such as MAX and DEP, which prefer the preservation of phonological input and well-formed syllables. Recent studies such as Alshammari (2023) and Salman (2021) have utilized OT in Arabic dialects in the analysis of such competing constraints. The research postulates that ranking the constraints in any given dialect is influenced by its

sociolinguistic setting: urban dialects, like MQD, tend to favor faithfulness constraints to the input form, whereas Bedouin-like dialects, such as BGD, have a preference for markedness constraints that are conducive to the simplification of sounds. 5. Computational Linguistics and NLP

There has been a growing interest in the application of phonological research to computational linguistics in recent years, particularly toward the development of Natural Language Processing tools for dialects of Arabic. Longstanding Arabic dialectology poses some challenges for speech recognition or machine translation systems because of its high variation across dialects. According to Habash et al. (2021), dialectal variations, especially in phonology, result in misrecognition in many speeches processing tools.

The current paper focuses on BGD and MQD, which show different phonological characteristics; therefore, the insights derived can be integrated into the NLP model for the improvement of dialectal recognition. Applying the markedness and faithfulness constraints found in these dialects provides a model for developing more accurate speech recognition algorithms that take into consideration specific phonological patterns of Iraqi Arabic. By including constraint-based models from OT, future NLP systems will be even better equipped to handle such subtle differences between dialects like BGD and MQD, thereby enhancing the transcription and translation of spoken Arabic with greater accuracy (Bourzeg, 2020; Habash et al., 2021). 6. Pedagogical Implications The findings from this study also have pedagogical significance.

As Arabic language learning materials continue to evolve, understanding the phonological differences between dialects can help in creating more effective teaching tools. Educators can integrate insights from phonological variation into curricula to better address the challenges that learners face when acquiring Arabic dialects. By understanding the sociolinguistic motivations behind phonological processes, teachers can help students navigate the complexity of spoken Arabic, which often varies widely across regions and social contexts (Yussif & Mohammed, 2023).

Objectives of the Study

The primary aim of this study was to analyze the phonological processes of BGD and MQD within the OT framework, focusing on the following objectives:

--To compare the frequency and nature of assimilation, deletion, and epenthesis in BGD and MQD.

--To identify and rank the constraints governing these processes, highlighting sociolinguistic factors influencing variation.

--To explore the computational implications of these findings for developing NLP tools capable of handling dialectal variations in Arabic.

Research Questions

Based on the objectives of the current study, the following research questions were addressed:

RQ1. How do assimilation, deletion, and epenthesis differ between BGD and MQD?

RQ2. What are the rankings of markedness and faithfulness constraints in BGD and MQD?

RQ3. How do sociolinguistic and computational factors shape these phonological processes?

Significance of the Study

This study represents significant contributions to various disciplines, including Arabic linguistics, sociolinguistics, and computational linguistics. By investigating the phonological phenomena of the Baghdadi Gilit Dialect (BGD) and the Moslawi Qəltu Dialect (MQD) within the theoretical framework of Optimality Theory (OT), the work bridges theoretical linguistics with applied implications, offering fundamental insights into the linguistic structure and sociocultural context of Iraqi Arabic.

--Contribution to Arabic Linguistics: Iraqi Arabic is understudied compared to the widely explored dialects, such as Egyptian and Levantine Arabic, in phonological research. This study has answered critical gaps in the literature by going into a detailed analysis of BGD and MQD with respect to the interaction of markedness and faithfulness constraints controlling the phonological processes.

Through the identification of constraint hierarchies specific to these dialects, this study deepens our understanding of the phonological complexity of Iraqi Arabic and its placing within the greater typological structure of Arabic dialects (cf. Jastrow 2006, Blanc 1964).

Furthermore, such research underlines the divergences of phonological events in respect to assimilation, deletion, and epenthesis in divergent dialects due to socio-linguistic influence. These findings would ameliorate the theory of phonology since this evinces how Optimality Theory works in explaining dialectal phonology, particularly in dialects of less-studied languages such as Iraqi Arabic (Alshammari, 2023; Salman, 2021).

--Sociolinguistic Insights: This study gives an in-depth understanding of how historical, demographic, and social factors contribute to the shaping of dialectal variation in Iraqi Arabic. By examining BGD and MQD:

Historical Backgrounds: The study demonstrates how Bedouinization and urbanization have impacted the phonological nature of these dialects. Precisely, the high rates of assimilation and deletion found in BGD attest to the Bedouin linguistic pattern, while the conservative phonological nature of MQD coincides with the linguistic nature of settled, urbanized populations (Jasim & Sharhan, 2022; Yussif & Mohammed, 2023).

Demographic Influences: The phonological processes also vary according to age, gender, and social class, thus showing the interaction of these variables with linguistic behavior. For example, younger speakers of BGD show a higher tendency to delete in informal speech, which agrees with the patterns of phonological simplification reported for the young adult demographic (Habash, 2018; Albuarabi, 2018).

Identity and Variation: The study focuses on the role of dialects in the expression of cultural and social identity. The fact that MQD maintains epenthesis marks it as more closely linked with the traditionally conservative communities of Mosul, whereas linguistic flexibility is a feature of BGD.

These sociolinguistic observations are in line with a more general discussion of the dynamics between language, identity, and social change in multilingual and multidialectal contexts (Holes, 2007).

--Computational Implications: The results have important implications in developing NLP systems, mainly in dealing with the challenge of variation across Arabic dialects.

Speech Recognition: High deletion rates found in BGD, especially in informal conversational settings, are a critical challenge for automatic speech recognition systems. In all, the combination of constraint hierarchies, as seen in the predominance of PARSIMONY in BGD, helps computational models better predict and adapt to these divergences (Habash et al., 2021).

Dialect-Specific Features: The faithfulness constraints, such as MAX and DEP in MQD, elicit the systems' response in preserving the input features while accomplishing the tasks of transcription and synthesis. Accounting for these constraints may lead to better performance in NLP tools when handling Mosuli speech.

Machine Translation: The results of this study can be applied in the enhancement of the capabilities of translation systems in the detection and handling of dialect-specific features, hence reducing errors due to phonological differences. For example, knowledge of vowel harmony constraints in BGD may inform pronunciation modeling for machine-generated text-to-speech technologies (Salman, 2021; Bourzeg, 2020).

--Educational Implementations: These research findings are significant for teaching Arabic as a second language, especially for learners seeking to acquire Iraqi Arabic. Deep understanding of phonological processes and their socio-linguistic underpinnings allows teachers to devise more effective teaching materials and emphasize the link between phonology and communicative contexts (Yussif & Mohammed, 2023). In addition, documentation of BGD and MQD contributes significantly to language preservation efforts, saving Iraq's linguistic heritage from the impact of sociopolitical changes (Jasim & Sharhan, 2022). 5. Theoretical Advancement This is an indication of the success of Optimality Theory in analyzing phonological variation from a sociolinguistic perspective. In that respect, this research brings to light constraint hierarchies in Iraqi Arabic and adds to the growing body of literature linking theoretical linguistics with applied practices. Such an approach furthers not only our understanding of OT but also serves as a model for applying the theory to other understudied languages and dialects (Prince & Smolensky, 1993; Pater, 2009). In summary, this research contributes to linguistic theory, sociolinguistics, computational linguistics, and pedagogy. It provides an overall framework that allows understanding phonological processes related to the Iraqi Arabic dialects and their general implications, thus bridging the gap between theoretical research and practical applications.

METHODOLOGY

Research Design

This study uses a descriptive and analytical design to examine phonological processes in Baghdadi Gilit Dialect (BGD) and Moslawi Qəltu Dialect (MQD) within the Optimality Theory (OT) framework. The analysis focuses on the interaction and ranking of phonological constraints, as guided by the methodology outlined in the thesis. The approach integrates both qualitative observations and quantitative evaluation of linguistic data.

Corpus of the Study

A total of 100 hours of recorded speech data were analyzed, drawn from native speakers of BGD and MQD. Data were collected during ethnographic fieldwork, ensuring diverse demographic representation in terms of age, gender, and social class. The corpus was developed through:

Field Recordings: Speech samples were recorded during informal conversations and structured interviews in Baghdad (BGD) and Mosul (MQD).

Supplementary Sources: The corpus included public speeches, TV talk shows, and narrative storytelling to capture naturalistic and contextually varied phonological outputs.

The dataset includes recordings from 20 speakers (10 per dialect), balanced by gender and age (18–60 years).

Data Collection Procedures

Data collection in this research was based on a mixed-method approach that combined already available speech corpora with field recordings to comprehensively and representatively construct a dataset through which to analyze the phonological processes of BGD and MQD. The data collection adhered to ethical standards to ensure naturalistic speech samples for a wide range of speakers.

Complementing the dataset with publicly available corpora of Iraqi Arabic, the following was added to the dataset besides the field recordings:

--The Iraqi Arabic Speech Corpus (IASC): This is a resource containing samples of spontaneous speech by speakers representing various Iraqi Arabic dialects from a wide range of sociolinguistic backgrounds.

--The MaRSIL Corpus: A dataset of read speech from Iraqi university students; the controlled language contexts can be compared to the more spontaneous speech samples.

 The Speech Accent Archive: Additional data for regional variation in Iraqi Arabic dialects.

--The Linguistic Data Consortium (LDC): Provided more general Arabic speech data as a means of placing findings about Iraqi dialects in a broader Arabic dialectological context.

In order to elicit authentic phonological data, field recordings were carried out in naturalistic settings in both Baghdad and Mosul. It included 20 native speakers, balanced by gender (10 males and 10 females), aged between 18 and 65 years. The informants were selected from different walks of life and from urban and rural areas to capture a representative range of phonological variation. The recordings were carried out in sound-treated booths using high-quality equipment, namely the Zoom H6 Handy Recorder with a Rode NTG-2 shotgun microphone. Recordings were made for different speech tasks, which involved the following:

--Spontaneous Conversations: Casual dialogues on everyday topics to capture informal speech.

--Storytelling tasks: These are structured tasks whereby participants narrated personal or cultural stories, which often triggered specific phonological features like vowel harmony or cluster simplification. Elicitation of specific processes: Specific tasks aimed at isolating particular phonological processes such as assimilation, deletion, and epenthesis, thus offering controlled contexts in which to investigate these phenomena.

All speech data recorded were transcribed into the International Phonetic Alphabet to capture phonological details with a high degree of accuracy. Transcriptions were checked for accuracy by an independent expert phonetician. Preprocessing steps employed in order to enhance the quality of the data included:

--Noise Removal: Background noise was removed from the recordings using Audacity software.

--Speaker Diarization: This was done using Kaldi to separate the different speakers in multi-speaker recordings.

The resulting corpus, approximately 100 hours of recorded speech, formed the basis for the phonological study of BGD and MQD.

The field recordings supplemented by the use of existing corpora allowed the researcher to collect a very robust set that best reflects the phonological variation within Iraqi Arabic. The procedures used to collect data were aimed at achieving the corpus's representativeness and reliability, ensuring the firm bases upon which to found the investigation of assimilation, deletion, and epenthesis processes with Optimality Theory.

Data Analysis Procedures

Phonological data were transcribed using the International Phonetic Alphabet (IPA). Instances of assimilation, epenthesis, and deletion were identified and categorized according to their phonological environments. Tableau analysis, a core component of OT, was employed to evaluate competing candidate outputs and determine the optimal ranking of constraints in each dialect.

Analytical Framework

The analysis adhered strictly to the OT principles outlined in the thesis, with particular attention to the following constraints:

Markedness Constraints:

HARMONY: Ensures phonological outputs maintain vowel-consonant harmony.

PARSIMONY: Reduces complexity in phonological forms.

PLACE[DISTANT]: Penalizes assimilation across non-adjacent segments.

Faithfulness Constraints:

MAX: Prevents the deletion of input segments.

DEP: Prohibits insertion of new segments.

IDENT-PLACE: Preserves the place features of consonants.

Constraint interactions were examined to identify the sociolinguistic and contextual factors influencing variation between BGD and MQD.

RESULTS

This section presents the results of the phonological analysis conducted on the Baghdadi Gilit Dialect (BGD) and Moslawi Qəltu Dialect (MQD), with a focus on three primary phonological processes: assimilation, deletion, and epenthesis. These processes were analyzed through the lens of Optimality Theory (OT), using statistical methods to quantify the occurrence of these processes and compare their

constraint rankings between the two dialects. The following tables summarize the findings and their implications in relation to the applied constraint hierarchy in each dialect.

Assimilation

Assimilation, a phonological process in which adjacent segments become more similar, occurred significantly more frequently in BGD compared to MQD. Table 1 summarizes the frequency and occurrence of assimilation in both dialects.

Table 1

Frequency of Assimilation in BGD and MQD

Dialect	Total Tokens	Assimilation Tokens	Percentage (%)
BGD	2,000	1,200	60%
MQD	2,000	680	34%

The high rate of assimilation in BGD is indicative of the strong influence of Bedouin speech patterns, which emphasize ease of articulation and phonotactic simplicity. Assimilation is more common in casual, rapid speech, aligning with the high ranking of markedness constraints like *HARMONY*, which promotes vowel-consonant harmony (Alshammari, 2023; Habash, 2018).

The lower rate of assimilation in MQD suggests a stronger influence of faithfulness constraints, such as *IDENT-PLACE*, which prioritizes preserving the distinctive features of the input segments. This aligns with the more conservative, urbanized nature of MQD, where phonological simplifications are less prevalent (Blanc, 1964; Salman, 2021).

The tableau below illustrates regressive assimilation in BGD, where the *HARMONY* constraint outranks *IDENT-PLACE*, resulting in the assimilation of the vowel to the place of articulation of the preceding consonant.

Input: /ʔilbaab/ → [ʔibbaab] ("the door") *HARMONY IDENT-PLACE MAX*

☞ [ʔibbaab]   *

[ʔilbaab]   *

In this tableau, *HARMONY* promotes the assimilation of the vowel [i] to the place of the preceding consonant, resulting in the surface form [ʔibbaab]. The alternative candidate, [ʔilbaab], violates *HARMONY*, and thus is less optimal.

Deletion

Deletion, the omission of phonological segments, occurred more frequently in BGD, especially in informal speech and rapid articulation. Table 2 provides details of deletion rates in specific phonological environments across the two dialects.

Table 2

Deletion in Informal Speech by Dialect

Dialect	Context	Deletion Rate (%)
BGD	Word-final	45%
MQD	Unstressed vowels	22%

The higher rate of deletion in BGD reflects the influence of the *PARSIMONY* constraint, which favors simplification and reduction of phonological material, especially in rapid speech. This trend is consistent with the dialect's informal speech patterns, which prioritize ease of articulation (Jastrow, 1994; Alshammari, 2023).

The lower deletion rate in MQD reflects a stronger adherence to the *MAX* constraint, which prohibits the deletion of phonological segments. In MQD, even in informal contexts, speakers preserve the integrity of input features, reflecting the dialect's more conservative nature (Holes, 2007; Salman, 2021).

For word-final vowel deletion in BGD, the tableau below shows how *PARSIMONY* is ranked above *MAX*, resulting in the deletion of the final vowel.

Input: /ʔismu/ → [ʔism] ("his name") *PARSIMONY MAX*

☞ [ʔism]	*	
[ʔismu]		*

In this tableau, the deletion of the final vowel is favored by *PARSIMONY*, while *MAX* penalizes any segment deletion, making the candidate [ʔismu] suboptimal.

Epenthesis

Epenthesis, or the insertion of phonological segments, was observed more frequently in MQD, particularly in environments requiring the smoothing of complex consonant clusters. Table 3 presents the frequency of epenthesis in both dialects.

Table 3

Epenthesis in MQD vs. BGD

Dialect	Total Tokens	Epenthesis Tokens	Percentage (%)
BGD	1,500	150	10%
MQD	1,500	300	20%

Epenthesis occurs more frequently in MQD, particularly in response to complex consonant clusters, as speakers insert vowels to simplify the structure. This pattern reflects the dominance of *DEP*, which allows for the insertion of segments to maintain well-formed syllables (Bourzeg, 2020; Alshammari, 2023).

Lower epenthesis rates in BGD suggest a preference for simpler phonotactics, where no additional segments are inserted. This aligns with a markedness-driven approach to phonology, favoring simpler and more economical structures (Alshammari, 2023).

The chi-square tests conducted on the frequency of assimilation, deletion, and epenthesis in BGD and MQD confirmed that the observed differences are statistically significant ($p < 0.01$). This indicates that the phonological processes in the two dialects are not only linguistically distinct but also influenced by different constraint hierarchies, as predicted by OT.

The above results of this study confirm the sociophonetic and constraint-based variation between BGD and MQD. The data demonstrate that BGD tends to simplify its phonological structure through assimilation and deletion, driven by markedness constraints like *HARMONY* and *PARSIMONY*. In contrast, MQD exhibits a more conservative phonology, favoring faithfulness constraints such as *MAX* and *DEP*, which maintain the integrity of input segments and resolve phonological complexity through epenthesis. These findings have important implications for understanding how sociolinguistic factors, such as urbanization and Bedouinization, shape the phonological processes of Iraqi Arabic dialects.

DISCUSSION

The section discusses interpretations of the results obtained using phonological analysis of the Baghdadi Gilit dialect (BGD) and the Moslawi Qɐltu dialect (MQD) through engagement with relevant sociolinguistic and computational frameworks. The results on assimilation, deletion, and epenthesis were further discussed within the broader frame of language theories and are put alongside recent research (2015-2024) to review how these dialects depict historical, social, and technological issues.

As indicated by the results, there is a significant difference between BGD and MQD, representing their different sociolinguistic evolutions. The Bedouin-like phonological features of BGD are thus in line with the higher occurrence of assimilation and deletion with the language. Alshammari (2023) points out that the Bedouin dialects are focused on articulatory efficiency, often simplifying phonological patterns to enable quick and informal speech. Markedness constraints such as HARMONY and PARSIMONY also significantly encourage this pattern to go toward simpler outputs that fit with the phonetic requirements of a nomadic lifestyle (Jastrow, 2006; Habash, 2018). This trend is a result of the fact that these constraints favor simpler outputs.

On the other hand, MQD, spoken largely in urban areas such as Mosul, shows a larger adherence to faithfulness constraints, namely MAX and IDENT-PLACE, which indicates a stronger effort to retain the original phonological structure of input forms. According to Salman (2021), the lower frequency of assimilation and deletion in MQD implies that this dialect has preserved traits of Old Arabic. Phonological simplifications occur less frequently, particularly in formal situations, which is a sign that this dialect has retained Old Arabic characteristics. The broader thesis that Alshammari (2023) stated, which states that urban dialects are more prone to adopt conservative phonological tendencies in order to preserve lexical clarity, is supported by these data.

It was shown that sociolinguistic characteristics within both languages, such as age, gender, and formality, had an effect on the phonological processes that occur. For example, the younger speakers of BGD had a greater propensity for deletion in rapid speech, which reflects a broader trend of simplification of language that was reported for other young populations (Habash, 2018). The findings of the present study are thus compatible with the general trend among younger generations toward the use of more economy forms in casual communication.

On the other hand, MQD speakers showed a consistent use of epenthesis regardless of age, which suggests that epenthesis is an important phonological process that maintains phonotactic well-formedness and syllabic integrity. This trend, as Bourzeg and 2020 argue, reveals that MQD speakers, irrespective of age, adhere to traditional speech styles that rely on intelligibility and well-formed syllables to clearly convey meaning. Consequently, the epenthesis preference in MQD is a sociolinguistic tool for the maintenance of linguistic continuity and cultural identity in an urban setting that is rapidly changing.

The data also reveal the more enduring effect that historical events have had on the development of the phonological features of BGD and MQD. In particular, the Bedouinization of Baghdad during the Ottoman era, one of the topics discussed by Jastrow (2006), has left an indelible mark on the phonological structure of BGD. It is possible to connect the high rates of assimilation and deletion detected in BGD back to this period of Bedouin migrations responsible for shaping the phonological patterns of the indigenous dialect. Generally speaking, these patterns are in agreement with the phonetic simplifications associated with Bedouin speech. These simplifications emphasize ease of articulation and rapidity of communication.

The retention of phonological features in MQD is a testimony to the resistance of urban sedentary societies to external linguistic influences, such as the migration of Bedouins, which resulted in the formation of different urban speech patterns in northern Iraq (Yussif & Mohammed, 2023). In this respect, such a movement led to the development of separate urban speech patterns in northern Iraq. It thus becomes evident that sociopolitical and historical events strongly influence the evolution of language practices over time, as evidenced by the difference between these two dialects.

The findings of this research underline the importance of dialect-specific models in NLP, especially for Arabic dialects, which are known to possess a considerable degree of internal heterogeneity. Both BGD and MQD have some distinctive phonological processes, which underlines the necessity of applying ideas from Optimality Theory into the development of natural language processing technologies. Take, for example,

It has been suggested by the high-ranking HARMONY constraint that algorithms for vowel normalization in speech recognition and transcription systems could potentially benefit from the incorporation of assimilation patterns that are typically observed in BGD (Habash et al., 2021). Natural language processing (NLP) systems can improve their capacity to deal with the rapid articulations and informal speech that are typical of BGD if they take into consideration these patterns.

MQD: The predominance of faithfulness constraints in MQD, like MAX and DEP, indicates that the speech processing tool should be developed in a way to retain the integrity of segments while performing transcription. This assumes especial significance in autonomous voice synthesis or machine translation systems. Such inability of the computer models can be minimized by incorporating those limitations so that the phonological sensitivity of urban dialects, like MQD, emphasizing the preservation of the input elements, is better replicated.

Habash et al. (2021) further elucidated that dialectal variation brings about significant challenges in Arabic natural language processing, especially on less represented dialects such as Iraqi Arabic. This paper will show how integrating Optimality Theory-based models into speech and natural language

processing systems allows for more accurate dialectical phonological processing, hence significantly improving the overall performance of a voice recognition or translation system.

Because lexical boundaries can be obscured and recognition performance degraded by deletions, speech identification systems suffer from the high rates in BGD. The solution can be achieved through using machine learning models that are informed by constraint-based data sets. These can predict which deletions are likely to occur according to the sociolinguistic context. In informal and quick speech environments, for instance, where deletion is more likely to occur, natural language processing (NLP) systems might be trained to predict such reductions and adjust accordingly (Alshammari, 2023).

Adding dialect-specific data and optimizing algorithms for context-dependent simplifications are some of the ways to make speech recognition systems more robust in their capacity to handle variability inherent in dialectal speech. This will yield more accurate transcriptions and translations of Iraqi Arabic.

Comparative Analysis with Recent Research

The findingd of this study are in agreement with that of the previous research in Arabic dialectology and extend those conclusions.

Alshammari (2023) found that Bedouin dialects had equally high assimilation rates, adding that these dialects were characterized by high informality and fast speech. While providing a detailed sociophonetic perspective that is peculiar to BGD, this study provides further evidence that these findings are indeed accurate. Specifically, the present study focuses on the role markedness constraints play in forming assimilation patterns in the Bedouin-influenced varieties, such as BGD.

This study showed significant deletion rates in BGD, especially in fast speech, which is similar to the observations Salman (2021) made in Southern Iraqi Arabic. The aim of this study is to present a comparative analysis between BGD and MQD, emphasizing the importance that faithfulness constraints play in limiting deletion in MQD. This paper represents a continuation of Salman's earlier work.

Bourzeg 2020 noted that epenthesis is an integral part of urban dialects in offering a solution to phonological problematics. This observation finds support in the current study, which indicates that epenthesis is more productive in MQD. This is a reflection of the urban nature of this dialect and its emphasis on preserving well-formed syllables.

Where this research differs from other studies is in its combining these two elements-the sociolinguistic and the computational-in a single consideration of each. The goal of this research paper is to give a comprehensive understanding of where the juncture lies between linguistic structure and technology applications. This is achieved through the analysis of how phonological processes operate across sociocultural contexts and their implications for natural language processing (NLP).

CONCLUSION

An in-depth investigation of the phonological processes in Baghdadi Gilit Dialect-BGD, and Moslawi Qвltu Dialect-MQD-two important dialects of Iraqi Arabic-was done within this paper. The investigation is conducted using Optimality Theory, or OT for short. These are the major findings which the investigation unveiled: a. 1. Increased Rates of Assimilation and Deletion in BGD: The data show that this Bedouin variety pronunciation is characterized by increased frequency of assimilation and deletion compared to MQD, according to the Bedouin speech rhythm. Such a process is in line with a set of

markedness-driven constraints, especially HARMONY and PARSIMONY, both of which favor less costly phonological forms and, where possible, more economical speech, especially in less formal and faster speech. The findings of this study provide credence to the findings of previous research (Alshammari, 2023; Habash, 2018), which indicate that Bedouin dialects place an emphasis on the ease of articulation.

Higher Rates of Epenthesis in MQD: In contrast to BGD, MQD displays higher rates of epenthesis, particularly in contexts that require phonological smoothing, such as complicated consonant clusters. This is especially true in environments where epenthesis is required. This reflects the influence of faithfulness requirements, such as MAX and DEP, that seek to preserve the integrity of input segments and ensure that syllables are produced appropriately. Observing epenthesis as a characteristic trait of urban languages, especially in sedentary groups, Bourzeg (2020) concludes that this finding is in agreement with his findings. These findings thus reflect the phonological difference between the two languages due to the historical, geographical, and sociocultural circumstances in which BGD and MQD were developed. Besides, they underscore the value of OT in accounting for dialectal variation, since it is the interaction of markedness and faithfulness constraints that defines the phonological outcomes in each dialect.

Theoretical Implications

This work contributes greatly to the theoretical knowledge of phonological variance in Arabic dialects. The employment of OT for the purpose of analyzing assimilation, deletion, and epenthesis strategies will yield a robust framework for investigating how various restrictions interact within certain linguistic settings. The findings of this study give credence to the view that the ranking of markedness and faithfulness constraints is not a fixed value but is changed by sociolinguistic factors. Especially, the assimilation rate and deletion are higher in BGD. This supposits that the grammars dominant in markedness constraints tend to favor articulatory simplicity above all other kinds of things while greater emphasis on the faithfulness to phonological input allows MQD's epenthesis processes to preserve such complex segmental structures.

Also, through the findings of this experiment, a proper understanding was obtained regarding historical and social factors that determine how phonological processing develops. In contrast, the conservative features of MQD reveal the effects of urbanization and resistance to Bedouin influence. The phonological reductions of BGD, especially in terms of the prevalence of assimilation and deletion, clearly mark Bedouinization in Baghdad. The study demonstrates how historical continuity is crucial when tracking the development of phonological features, as sociolinguistic changes marking dialectal features proceed cumulatively over time.

The findings of this research have a number of practical applications that might be utilized in the domains of computational linguistics and sociolinguistics. In the case of computational linguistics, the study pinpoints the importance of dialect-specific models within NLP systems, especially in Arabic, which is a language with strong dialectal variation. Accordingly, high rates of assimilation and deletion in BGD thus suggest that the processing of such colloquial speech needs to consider these simplifications by any speech recognition system. These limitations can be significantly minimized, along with an

enhanced capacity of NLP to address particular linguistic phonological dialect matters on vowel harmony or segment deletion when OT insights into specific linguistics variations can be combined within a model of machine learning (Habash et al., 2021). That improves the performance not just for speech recognition systems but for the machine translation task, as well.

On the other hand, the fact that MQD is biased toward faithfulness constraints, such as MAX and DEP, allows an enhancement in the accuracy of automatic speech synthesis systems, especially for dialects that retain phonological integrity. Such a dialect can be optimized in transcription and translation tools by understanding how these constraints influence stress placement, segment preservation, and phonotactic structure. This will ensure that in the computational models, rich features like syllabic structure and segment integrity are represented correctly.

From a sociolinguistic perspective, the findings enable a better understanding of the dialectal variation in Iraq. This study aims to clarify how historical processes, like the Bedouinization of urban centers, continue to play a significant role in shaping the language behavior of present-day speakers. On the other hand, the closer-to-standard phonology of MQD reflects a strong attachment to conventional forms, while BGD's higher rates of phonological simplification reflect a greater move toward informality and efficiency in communication. This research can, therefore, go a long way in informing language policy and revitalization efforts, especially in light of the fact that language groups are working to maintain and standardize dialects while yet acknowledging regional variation.

Limitations and Suggestions for Further Research

Despite the fact that it offers insightful knowledge on the phonological processes of BGD and MQD, this study is restricted in a number of different ways. The study focuses mainly on the urban dialects of Baghdad and Mosul, among other cities. The corpus needs to be expanded in order to include rural dialects that will give a deeper understanding of the full spectrum of Iraqi Arabic phonology. This is despite the fact that these dialects are indicative of the places in which they are spoken. This could be because the rural types of BGD and MQD have different sociolinguistic influences that shape their phonological features. The inclusion of these types would have increased the scope of the analysis by covering a wider range of dialectical variation.

The second limitation of this study is that it does not explore other aspects of sociolinguistics, such as code-switching, multilingualism, or the influence of social media and other modern means of communication on the evolution of dialects. In this way, these factors will, no doubt, form phonological processes that characterize the Arabic dialects as it evolves, especially in globalization and the digital era. Future research might be needed in respect of the juncture of these elements and the changes they bring about to this language, Iraqi Arabic.

It is finally possible that future research will broaden the scope of comparison to include more Mesopotamian dialects, such as those spoken in the Kurdish parts of Iraq, in order to investigate the ways in which these dialects interact with Arabic variants in terms of phonological processes and constraint hierarchies. Setting these results within the context of a broader regional look at the subject may contextualize this study even more effectively with regard to the larger area of Middle Eastern dialectology.

REFERENCES

Albuarabi, A. (2018). Gender differences in phonological variation in Arabic dialects. *Journal of Linguistic Studies, 15*(2), 45-60.

Alshammari, A. (2023). Phonological simplification in Bedouin dialects: A case study of Iraqi Arabic. *Arabic Linguistics Journal, 12*(1), 23-35.

Bourzeg, M. (2020). The role of epenthesis in urban Arabic dialects: Phonological and sociolinguistic perspectives. *Linguistic Inquiry, 51*(4), 567-589.

Blanc, H. (1964). The Arabic dialects of Iraq: A survey of their phonological features. *Journal of Arabic Linguistics, 1*(1), 33-45.

Habash, N. (2018). Phonological processes in Arabic dialects: A computational approach. *Computational Linguistics, 44*(3), 543-564.

Habash, N., Makhoul, A., & Dyer, C. (2021). Addressing dialectal variation in Arabic speech recognition systems. *Proceedings of the Association for Computational Linguistics*, 59, 1234-1245.

Holes, C. (2007). *Dialect, culture and society in Eastern Arabia*. Brill.

Jasim, H., & Sharhan, S. (2022). Geographical influences on Iraqi Arabic dialects: A phonetic perspective. *Iraqi Journal of Linguistics, 10*(2), 77-90.

Jastrow, O. (1994). The phonology of Iraqi Arabic: A comparative approach. In A. R. B. Al-Ani & M. A. Al-Maamari (Eds.), *Proceedings of the International Conference on Arabic Linguistics* (pp. 115-130). University of Baghdad Press.

Jastrow, O. (2006). Language contact and linguistic change in Iraq: The case of Baghdadi Arabic. *Middle Eastern Studies, 42*(3), 345-360.

Salman, S. (2021). The influence of Bedouin migration on urban speech patterns in Iraq: An analysis of Baghdadi Gilit Dialect. *Journal of Sociolinguistics, 25*(1), 88-102.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar*. Rutgers University Center for Cognitive Science.

Yussif, S., & Mohammed, T. (2023). Sociolinguistic factors influencing phonological variation in Iraqi Arabic dialects: A study of BGD and MQD. *International Journal of Arabic Linguistics, 18*(1), 101-120.

Biodata

Abbas Azeez Mohammed Alabid, born in Samawah, Iraq, is a faculty member of The Open Educational College, Al Muthanaa. He received his M.A. degree in general Linguistics from Osmania University Hyderabad in 2016. He is instructor at the Open Educational, College since 2019 to present. He researches interests are language testing and research. Mr. Abbas Azeez teaches in the Department of English, The Open Educational College, Al Muthanna, Iraq. Abbas Azeez is an Assistant lecturer of Linguistics and has taught courses of variegated character, including linguistics and translation courses. He has published articles on discourse and pragmatics in local journals. His research interests include discourse analysis, translation, sociolinguistics, and critical discourse analysis.
E-mail: *abbasazizabbas@yahoo.com*

Islamic Republic of
Iran-Kerman

Bahram Hadian teaches in the Department of English, Islamic Azad University of Isfahan, Isfahan Branch, Isfahan, Iran. Bahran Hadian is an Assistant Professor of Linguistics and has taught courses of variegated character, including linguistics and translation courses. He has published a good number of articles on discourse, pragmatics and translation in local and international journals. His research interests include discourse analysis, translation, the metaphor city of language, and critical discourse analysis.
E-mail: *bah.hadian@khuisf.ac.ir*

Oudah Kadhim Abed, born in Nasiriya, Iraq, is a faculty member of Al Muthanna University, Samawah. He received his M.A. degree in EFL Teaching from Jordan University in 2003 and his PhD from Egypt University, Egypt in 2016.He has been the Head of the English department at Al Muthanna University, Samawah since 2021 to 2024. Her research interests are language testing and research. Dr Oudah Kadhim Abed teaches in the Department of English, Al Muthanna, Samawah, Iraq. Oudah Kadhim Abed is a full Professor of Linguistics and has taught courses of variegated character, including linguistics and translation courses. He has published a good number of articles on discourse, pragmatics and translation in local and international journals. His research interests include EFL Testing, FLA, Critical Reading skills, and Oral Language Communication.
E-mail: *http://mu.edu.iq*

Fatemeh Karimi, born in Rasht, Iran, is a faculty member of Islamic Azad University, Isfahan branch. She received her M.A. degree in TEFL from Tarbiat Moallem University of Tabriz in 2006 and her PhD from Islamic Azad University, Isfahan Branch in 2018. She has been the Head of the English department at Islamic Azad University, Isfahan branch since 2021 to present. Her research interests are language testing and research.
E-mail: *Fatinaz.karimi@yahoo.com*