Jordan Journal of Modern Languages and Literatures Vol.17, No. 2, 2025, pp385-403

JJMLL

A Variationist Analysis of the Complementizer *ʔinnu* in Ammani Arabic

Dala Abu-Shreea
Department of English Language and Linguistics, Jordan University of Science and Technology, Jordan

Marwan Jarrah
Department of English Language and Literature, The University of Jordan, Jordan

Yazeed Hammouri *
European Languages, TheUniversity of Jordan, Amman, Jordan

Ekab Al-Shawashreh
Department of English Language & Literature, Yarmouk University, Jordan

Received on: 16-10-2023 **Accepted on: 30-5-2024**

Abstract

This study explores the realization of the complementizer *ʔinnu* in Ammani Arabic employing modern quantitative methods of the variationist approach. It examines the effect of some extra-linguistic and linguistic factors (age, gender, education, type of matrix, definiteness, specificity, word order, subject type and factivity) on the realization or omission of the complementizer in AA. The results reveal that the variable realization of the complementizer *ʔinnu* in AA is linguistically conditioned. The omission of the complementizer *ʔinnu* is subject to some linguistic constraints including the type of the preceding matrix and the word order of the embedded clause within *ʔinnu.* These factors play a statistically significant role in its realization/omission. Moreover, the results show that the social factors turned out to be statistically insignificant in constraining overt and null complementizer in AA.

Keywords: Variation sociolinguistics; Complementizer Omission; Ammani Arabic; Social and linguistic constraints; Variant choice.

1. Introduction

In spontaneous speech data, syntactic variables are less common than phonological variables (Cheshire 1999; Hudson 1996).[1] This is why syntactic variables seem likely to be difficult to study and recognize. However, this type of variation is still worthy of examination as they help in unraveling essential aspects of language properties. This study focuses on syntactic alternation between null and overt complementizer *ʔinnu* in Ammani Arabic (hereafter AA).

This study is a variationist investigation of the variable realization of the complementizer *ʔinnu* in AA. It investigates the correlation between certain linguistic and extra-linguistic factors and the realization or omission of *ʔinnu*. This phenomenon has been widely investigated by many variationists (e.g., Ferreire and Del 2000 for *that* omission in English; Liang et al. 2022 for the complementizer *que*

* Doi: https://doi.org/ 10.47012/jjmll.17.2.1
* Corresponding Author: **hyazeed@ju.edu.jo**

omission in French). Our investigation allows us to examine how some extra-linguistic and linguistic factors might condition the variable realization of the complementizer *ʔinnu* in AA, a variety that is understudied in terms of variationist analysis.

Through extensive corpus-based studies, *ʔinnu* has been described as a complementizer as well as a discourse marker (Cowell, 1964; Bloch, 1986; Germanos, 2013). Therefore, providing a description of the complementizer *ʔinnu* by applying quantitative variationist methods to examine how extra-linguistic and linguistic factors constrain its realization in AA constitutes the main motivation to conduct this study.

In Modern Standard Arabic (MSA), 'ʔinna', which means 'that', is one of a group of complementizers known as 'ʔinna and its sisters'. Its equivalent in JA is *ʔinnu* that naturally comes in multiple forms: *ʔinha* (feminine) or *ʔinhum* (plural) as illustrated in the examples (1, 2, and 3) below:

(1)

a. *bagūl* ***ʔinnu*** *l-walad* *ʔiʤa*
Say.1SG *that* *the-boy* *came*

b. *bagūl* *øC* *l-walad* *ʔiʤa*
Say.1SG *ø* *the-boy* *came*

(2)

a. *ga:l-at* ***ʔinha*** *l-miss* *ʔiʤat*
Said-3SG.F *that* *the-teacher* *came*

b. *ga:lat* *øC* *l-miss* *ʔiʤat*
Said-3SG.F *ø* *the-teacher* *came*

(3)

a. *gal-l-i* ***ʔinhum*** *l-awlad* *ʔiʤu:*
He said (to me) *that* *the-boys* *came*

b. *gal-l-i* *Øc* *l-wala:d* *ʔiʤu:*
said-to-me *ø* *the-boys* *came*

As shown in the examples above, the same sentence is realized in two different ways (overt complementizers in (a) and null complementizers in (b)) and still gives the same meaning. This variable realization of *ʔinnu* constitutes an ideal linguistic variable to be investigated in that it is a case of having two ways of saying the same thing. According to Labov (1972), variation becomes definite when there are "two or more ways of saying the same thing" (271). Therefore, this study addresses the following research questions:

1. What is the distribution of null/overt complementizer in AA?
2. What is the correlation between some linguistic factors (type of matrix, definiteness, specificity, word order, subject type and factivity) and the choice of null/overt complemetizer in AA?

3. What is the correlation between some extra-linguistic factors (age, gender, level of education and region) and the choice of null/overt complemetizer in AA?

To answer these questions, sociolinguistic interviews with 32 native speakers of AA were conducted and then analyzed following the variationist approach (Labov, 1972). The following sub-section provides a description of the key tenets of this approach.

1.1 Theoretical Framework

Al-Wer (2009, 1) states that "variation is an inherent characteristic of every living human language. This means that in every language, there is more than one way of saying a same notion, and no individual speaks in exactly the same manner all the time and in all situations". Similarly, Wolfram (2006, 333) further adds that "if structure is at the heart of language then variation defines its soul". Accordingly, variationists seek to understand how language functions by describing systems in a quantitative approach as a method to provide a precise explanation of variation and the potential means for language development and change.

Variation can be found within the same speech community at different linguistic levels, including pronunciation, word choice, and grammatical structures. Different variants of any of the preceding linguistic components do not necessarily express different propositional meanings; rather, they may differ only in terms of their social meanings (Al-Wer 2009). Variation is never arbitrary or chaotic, but rather comes as a result of the influence of certain extra-linguistic (e.g., age, gender, education, social status, origin, residential area, etc.) and/or linguistic factors (e.g., animacy, transitivity, definiteness, specificity, type of clause, etc.). According to Labov (2001), it is rather an implicit aspect of a linguistic structure that indicates how language is used to express specific social attributes. Similarly, Al-Wer et al. (2020, 1) argues that "linguistic and social factors go hand in hand in structuring variability in language and any consequent instances of language change".

Another major property of the variationist approach is its reliance on the vernacular as it is "the most systematic data for linguistic analysis" (Labov 1984, 29). That is why we resorted in our study to sociolinguistic interviews in order to obtain the vernacular (spontaneous) speech data. Another important property of the variationist approach is the principle of accountability that "requires that all the relevant forms in the subsystem of grammar that you have targeted for investigation, not simply the variant of interest, are included in the analysis" (Tagliamonte 2012, 10). Respecting this principle, we extracted all occurrences of overt complemetizers in the data as well as all occurrences where the complemetizer could have occurred but did not.

After providing a description of the theoretical framework, we present a survey of the pertinent literature in the following section.

2. Literature Review

Research on Arabic linguistics has mainly focused on the Jordanian dialects (Cleveland 1963), Ammani dialect (Abdel-Jawad 1981; Al-Wer 1987, 1991 1997, 2003, 2007: 7), rudimentary leveling (e.g., Trudgill 2004, 89-93; Palva 1994), the impact of gender and status on emphasis in JA (Omari &

Jaber 2019), sociolinguistic variation in the Jordanian dialects (Al-Hawamdeh 2016), phonological variation in main Jordanian cities (Al-Tamimi 2001), and the variation in socio-political development (El-Salman 2003). Most of these studies examined variation in JA according to different aspects (see also Al-Shawashreh et al. 2024). However, the use of *ʔinnu* has not yet been explored from a variationist perspective. This section reviews existing studies on sociolinguistic variation in Jordanian Arabic.

Germanos (2010) examined *ʔinnu* as a complementizer and a discourse marker (DM). The study found that *ʔinnu* as a DM "signals an elaboration - an illustration or a further explanation, for example - of what has preceded it in the discourse" (145). In complement clauses introduced by *ʔinnu*, as a complementizer, it can either serve as the subject, the object or the predicative object of the main clause. It is worth mentioning that Germanos investigated *ʔinnu* by focusing on its discourse functions rather than on the correlation between the impact of some extra-linguistic and linguistic factors and its different realizations. This means that Germanos studied the overt occurrences of *ʔinnu* and the discourse functions they display.

Al-Shawashreh (2016) analyzed two syntactic variables, word order, and pro-drop in Jordanian Arabic. The study concluded that these two variables are constrained by social and/or linguistic factors. It was also found that the dominant word order in vernacular JA is SV(O) which occurs more frequently than VS(O) and that dropping the subject pronoun in JA is more frequent than keeping it as it can be easily understood due to the rich inflections on the verb.

Following the same variationist sociolinguistic framework, Khater (2021) investigated the realization of personal pronouns in AA. The statistical analysis revealed that age and sex play a key role in subject pronoun variation in (AA). Furthermore, the study showed that the grammatical person, polarity, tense and clause type I are statistically significant in constraining the variant choice.

Abdel-Aziz et al. (2024) investigated sociolinguistic variation in Determiner Phrase (DP) genitive alternatives in AA. To that end, she collected the spontaneous speech of 32 speakers from Amman and examined the extent to which the choice of free state nominals (FSN) or construct state nominals (CS) is constrained by social factors like age, sex, level of education and region, along with linguistic factors, such as definiteness, animacy, complexity and alienability. The results showed that CS is more frequent than FSN. Social factors (except for region) play a minimal role in comparison with linguistic factors in governing variant choice.

Another study focusing on variation in AA, Alshaboul et al. (2022) tackled how intensifiers in AA are governed by extra-linguistic factors (age, gender, and education) along with some linguistic factors (i.e., the position of intensifiers and the semantic class, function,, polarity and emotionality of adjectives). Adopting the variationist approach introduced by Labov (1972), the researchers analyzed more than 15 hours of audio-recordings obtained from 32 speakers of AA and extracted all tokens that included (ʔikθi:r) and its variants ([ʔikθi::r],[ʔikθi:r ʔikθi:r], and [dʒidan]). The findings show that the extra-linguistic factors are statistically nonsignificant in the distribution of intensifiers in AA. By contrast, linguistic factors like the semantic class and function of adjectives have statistically significant effects on the choice of the variants of (ʔikθi:r), indicating that the use of (ʔikθi:r) is linguistically constrained.

Inspection of previous variationist studies on syntactic variation in vernacular Arabic reveals that syntactic variation in spoken Arabic has been widely investigated. Yet, the literature on the realization/omission of the complementizer *ʔinnu* in Arabic varieties is very limited (if any). Thus, the present study fills a gap in the previous variationist literature on vernacular Arabic.

3. Methods

This section provides a description of the corpus of the study and the procedures of collecting and analyzing the data. It also presents the independent variables/factor groups operationalized in this study and the motivations for choosing them.

3.1 The Corpus

The corpus of the present study included 32 native speakers of AA; they were all born and raised in Amman. As shown in Table (1) below, the participants were classified according to their age, sex, and level of education. As indicated in previous studies in sociolinguistics, each of these factors has been reported to constrain the probability of variant choice. Age cohorts included 18-35, 35-50 and more than 50 years old for younger, middle-aged and older speakers. The reason behind this stratification is to assess the existence and directionality of any possible language change in progress. This also allows us to compare speech patterns in real time (Boberg 2004; Labov 1994). The participants were also stratified according to their sex (16 females & 16 males). Males and females are known to be different in their linguistic behavior. One of the most common differences is the presupposition that females favor prestigious variants than males. In the same manner as age, the linguistic distinctions between males and females help in predicting any possible linguistic change in progress in AA.

Table 1: The stratification of the participants according their age, sex and education

Level of education	**Male**			**Female**			**Total**
	Young	Middle-aged	Old	Young	Middle-aged	Old	
High	4	4	2	4	4	2	20
Low	2	2	2	2	2	2	12
Total	6	6	4	6	6	4	32

In other similar studies, one's level of education was reported to have an effect on the speech pattern of the speakers. Al-Wer (2000, 3) draws attention on education as a factor and suggests that "it is not level of education per se which correlates with linguistic usage; rather that level of education is actually an indicator of the nature and extent of the speakers' social contacts." She further clarifies that education is a 'proxy variable', one that is not directly relevant in itself, but serves in place of an unobservable variable. The participants were considered highly-educated if they obtain a bachelor degree and above. Otherwise, they are considered low educated.

3.2 Data collection

A mixed-method approach is adopted by applying qualitative and quantitative tools to best achieve the objectives of the study. Sociolinguistic interviews were conducted to collect data. The audio-recorded interviews lasted approximately 15 hours, 25-40 minutes each. This approach is utilized for the purpose

of collecting the vernacular, which is the most systematic data for linguistic analysis (Labov 1972). It best shows the effect of sociolinguistic factors on the choice of the variant of *ʔinnu* in the natural speeches of the native speakers.

Following the snowball technique (Milroy & Gordon, 2008), the data was collected from 32 participants who were selected to test the extent to which extra-linguistic and linguistic factors have effects on the targeted variable of this research, syntactic alternation between covert and overt *ʔinnu* in AA. A sequence of communicative modules - dreams, traditions, danger-of-death, personal experiences, career choices, pandemic related topics- was carefully prepared to encourage interviewees to be emotionally engaged in the interviews and thus focus on the narratives and events more than on the way they speak (see Appendix B). Doing so made the interviewees pay minimal attention to their speech and thus speak comfortably and spontaneously, leading to the desired speech style, i.e., the vernacular. This method is believed to be an effective tool for such kind of research and was described as one of the most influential means as it attenuates the effects of observer's paradox (Labov 1984). With the aim of diminishing any kind of formality during the interview, the questions were "formulated to be as colloquial as possible, avoiding any 'bookishness' of syntax and lexicon" (Milroy and Gordon 2003, 60).

3.3 Data Analysis

After data collection, all occurrences of overt complemenizer were extracted from the data. Moreover, in correspondence with the principle of accountability, a key tenet of the variationist approach (Labov 1972), we extracted all tokens where the complementizer *ʔinnu* could have occurred but did not. The principle of accountability "requires that all the relevant forms in the subsystem of grammar that you have targeted for investigation, not simply the variant of interest, are included in the analysis" (Tagliamonte 2012, 10). The extracted tokens (947) were then transferred into an excel sheet. After that, the tokens were coded according to a number of extra-linguistic and linguistic factors that are hypothesized to constrain variation. After the coding stage, all coded tokens were analyzed using Goldvarb X, which is a computer program that is capable of providing distributional and multivariate analyses in terms of the correlation between the effects an array of extra-linguistic and linguistic factors and the choice of variant (Sankoff, Tagliamonte and Smith 2005). In the present study, using this program can help us determine the extent to which the chosen linguistic and extra-linguistic independent variables affect the realization/omission of the dependent variable *ʔinnu*.

3.4 Factor Groups

This study explores the roles of some linguistic (specificity, definiteness, subject type, type of matrix, word order and factivity) and extra-linguistic (age, gender, and education) factors that may condition the choice of overt or null *ʔinnu* in AA. In this section, we explain why we chose these extra-linguistic and linguistic factors in particular.

3.4.1. Age

Age is considered as one of the most important extra-linguistic factors that limit linguistic change (Labov 1994). Age allows analysts to compare/contrast speech patterns in real time to examine any possible linguistic change in progress. According to previous research, younger speakers are more open to adopting new prestigious variants. Al-Wer (1991) asserts that among Arabs, the pressure exerted on old people is significantly larger than that exerted on young people.

3.4.2 Gender

Another decisive social factor in linguistic evolution is gender; its impact on phonological variables in JA and other variants of colloquial Arabic has been proven (Abdel-Jawad 1981, 1987; Al-Khatib 1988; Al-Wer 1991). The present data are analyzed for any potential gender impacts on conditioning the variable realization of *ʔinnu* in AA. If present, an impact may be caused because of the reported role that women play in leading language change (Labov 1990). Bassiouney (2009) argues that before studying the relationship between language and gender, one should begin with the presumption that gender as a social factor is rarely independent. Rather, it interacts with other extra-linguistic factors such as class and education. Furthermore, Bassiouney (2009, 213) concludes that analysts should assume that "the range of behaviors engaged in by women and men are not independent- no more Mars and Venus- but overlap and are highly contextualized".

3.4.3 Education

Level of education in Arabic speaking countries has been reported to play a significant role in predicting variant preference. According to Owens (2001, 435) and Al-Khatib (1988, 60), educated speakers use SA forms much more than illiterates do. Contrariwise, Al-Wer (2002) challenges this notion by stating that one's level of education is not necessarily associated with their choice of standard variants. To address the influence of education on *ʔinnu* complementizer drop in AA, participants in the present study were stratified as low-educated if they received primary/elementary/secondary education and highly-educated if they completed their higher education, i.e., they obtained diploma, BA, MA, PhD.

3.4.4 Definiteness

Definiteness is frequently viewed as the demarcation line between noun phrases (NPs) that introduce new referents and those that relate to already familiar referents. According to Abbott (2006), definiteness is a feature that draws a distinction between identifiable referents (definite NPs) and unidentifiable referents (indefinite NPs) in a given context. This linguistic notion is anticipated to have an effect on variant choice. Thus, the present study aims to test this anticipation in an attempt to investigate whether definite and indefinite possessors affect null or overt *ʔinnu* in AA.

3.4.5 Specificity

Various assumptions have been made about specificity. Enç (1991) concludes that some analysists characterize a NP as specific, when the speaker has a particular referent in mind. Meanwhile, others suggest collapsing the specific use of indefinites with the referential use of definites and the nonspecific use of indefinites with the attributive use of definites. However, the most wide-spread view on specificity seems to be the one where a NP is considered specific if it has a "wide scope over an operator" (Karimi 2008).

3.4.6 Word Order

Previous research indicated that word order rules significantly constrain linguistic behavior. Despite their importance, little is known about how these rules work. Most studies on word order in Arabic highlight it as a dependent variable (Al-Shawashreh 2016), only few studies focus on the effect of word order rules as an independent variable governing variant choice.

3.4.7 Subject Type

Another factor linked with complementizer dropping is subject type. Considerable attention has been paid to the pro/lexical subject in the previous literature in the syntactic theory. It has received much attention in languages that are morphologically rich in terms of its agreement inflections (Chomsky 1981). The current study considers this linguistic factor to investigate whether lexical or pro subjects prefer overt or null *ʔinnu* in AA.

3.4.8 Type of Matrix

Most of the previous studies on the variable realizations of the English complementizer ‘that’, which is the approximate equivalent to the Arabic complementizer *ʔinnu*, "assume that overt and null *that* clauses have the same underlying structure, and predict that these clauses show (nearly) the same syntactic distribution (Shim & Ihsane 2017, 515). Shim & Ihsane (2017) also argue that the type of matrix preceding the complementizer can play a great role in determining its distribution.

3.4.9 Factivity of the Preceding Verb

Jarrah (2019) investigated the clausal complements of factive vs. non-factive verbs in terms of their syntactic and semantic behavior. He argues that the verbs which take that-complements are divided into two classes, namely: factive verbs and non-factive verbs. The former class of verbs (such as *know*, *realize* and *regret*, among others) carries along the speaker's presupposition that the complement sentence represents a true proposition. On the other hand, non-factive verbs (such as *believe*, *think* and *assume*) propose that such verbs are not accompanied by a similar presupposition. Syntactically, these two classes of verbs are different in that complements of factive verbs are weak islands for extraction, whereas complements of non-factive verbs are not. The present study aims to study the difference between factive and non-factive verbs as a linguistic factor governing the distribution of null/overt complementizer *ʔinnu* in AA.

After having presented the motivations for operationalizing the aforementioned extra-linguistic and linguistic factors on the variable realization of *ʔinnu* in AA, we turn now to present the distributional and multivariate analyses.

4. Results

This section provides the results and interprets them in light of both findings of previous studies and theoretical groundings of sociolinguistic variation. As noted above, the main aim of the present study is to demonstrate how linguistic and extra-linguistic factors influence the choice of the variable *ʔinnu* in AA. More specifically, the results of the distributional and multivariate analyses as well as the roles of the factor groups in determining the realization of *ʔinnu* in AA are represented below.

4.1 Overall Distribution

The results in Table (2) below present the overall distribution of the variants in the corpus, i.e., overt and null *ʔinnu* in AA. The results show that overt *ʔinnu* is more frequent than null *ʔinnu* in AA, (57.9 % and 42.1%, respectively).

Table 2: Overall distribution of variants

Variant	Number of Occurrences	%
Overt	548	57.9
Null	399	42.1
Total	947	**100**

4.2 Extra-linguistic and Linguistic Factors: A Multivariate Analysis

With the help of multivariate analysis, three pieces of evidence can be gathered. It provides the statistical significance of the independent factors (less than 0.05) when they are run together concurrently. Another advantage of using multivariate analysis is the magnitude of effect that is shown by the range value of each statistically significant factor group. The range value is calculated by comparing the highest and the lowest factor weights within each factor group. This magnitude of effect indicates the strength of the factor group compared to the other factor groups, i.e., the higher the range value the stronger the factor group. In addition, the constraints hierarchy within each factor group indicates whether it favors or disfavors the choice of the variant. Factor weights above 0.50 denote that a factor has a favoring effect on variant selection while those below 0.50 indicate that a factor has a disfavoring effect.

The results in Table (3) below present the results of the multivariate analysis with regard to the contribution of extra-linguistic and linguistic factors to the probability that overt *ʔinnu* will be selected. The results in Table (3) show that among the variables considered only the type of matrix and word order are statistically significant in the selection of overt *ʔinnu*. Conversely, factors such as gender, education, age, subject definiteness, subject specificity, factivity of the verb, and subject type do not have statistical significance on the choice of overt *ʔinnu*. The range values of the statistically significant linguistic factors spot that the type of matrix has a stronger effect on variant choice than that of word order (54 & 49 respectively). Within type of matrix, the constraints hierarchies show that prepositions favor overt *ʔinnu* the most (.91), followed by other categories (.69) and nouns (.64). The constraints hierarchies further

show that verbs and adjectives disfavor overt *ʔinnu* (.42 and .37). Regarding word order, the constraints hierarchies show that SVO and verbless clauses favor overt *ʔinnu* (.58 & .57) while VSO and OVS disfavor it (.34 & .09).

Table 3: Multivariate analysis of the contribution of extra-linguistic and linguistic factors to the probability that overt *ʔinnu* will be selected

Corrected mean			0.592
Log likelihood			-599.455
Significance			0.085
Total number			548/947
Type of Matrix	Factor weight	%	N
Preposition	.91	92	23/25
Other Categories	.69	76	96/125
Noun	.64	72	108/150
Verb	.42	51	239/461
Adjective	.37	44	82/186
Range	*54*		
Word Order			
SVO	.58	65	104/160
Verbless	.57	65	217/333
VSO	.43	50	226/450
OVS	.09	25	1/4
Range	*49*		
Age			
Young (18-35)	[.52]	59	222/378
Middle-Aged (35-50)	[.50]	59	169/285
Old (50<)	[.47]	55	157/284
Gender			
Female	[.51]	58	366/629
Male	[.48]	57	182/318
Education			N
Low	[.52]	61	289/477
High	[.48]	55	259/470
Factivity			
Non-Factive	[.52]	54	178/329
Factive	[.45]	47	62/132
Definiteness			
Definite	[.52]	57	295/520
Indefinite	[.46]	58	183/314
Specificity			
Non-Specific	[.57]	58	148/245
Specific	[.46]	56	217/389
Type of Subject			
PRO	[.53]	57	272/480
Lexical	[.41]	55	94/172

The following section presents a discussion of the results. It provides answers to the research questions and integrates the results within the wider existing literature.

5. Discussion

As revealed by the multivariate analysis above, the results indicate that the type of the matrix predicate and the word order are the only factors that statistically determine the realization of *ʔinnu* in AA, while the factors of sex, age, education, specificity, definiteness, subject type and factivity were found to be statistically nonsignificant. This finding is similar to the numerous variationist studies that conclude that an array of linguistic factors does constraint variant choice (Abdel-Jawad 1981, Al-Khatib 1988, Al-Wer 1991, El-Salman 2003, Al-Shawashreh 2016). Embarking upon matrix type, as it has the strongest effect of variant choice; we can conclude that the local syntactic relationship between the complementizer *ʔinnu* with nouns and prepositions is stronger than that with verbs and adjectives as a result of linguistic case assignment. Likewise, it was observed that word order, more specifically when connected with categorical and thetic sentences, conditions the variant choice proven statistically by the fact that SVO and verbless sentences favor overt over null *ʔinnu*.

This result could be explained in light of the fact that prepositions in Arabic are considered case assigners. That is to say prepositions assign case to their complement positions and are always followed by nouns, unlike adjectives and nouns which only assign cases in limited linguistic situations. Phrases that have an accusative case in Arabic are complementizer phrases (CPs), phrases where the Arabic complementizer *ʔinnu* is the head, and determiner phrases (DPs)/nominal phrases (NPs). This explains why the factor weight is (.91) when *ʔinnu* is overt and preceded by a preposition and (.64) when it is preceded by a noun. That is to say, in these two cases, *ʔinnu* is overt for the case assignment (see Jarrah et al. 2022 for a relevant discussion). On the other hand, when *ʔinnu* is preceded by a verb or an adjective, it is usually followed by an adjunct; a phrase or a clause that can be removed from a sentence without making the sentence ungrammatical. In other words, verbs and adjectives do not necessarily assign cases. The complementizer, therefore, does not need to be overt.

The multivariate results also indicate that word order comes second in light of its effect on variant choice with SVO and verbless sentences favoring overt *ʔinnu* at almost the exact rate (.58 and .57). This result can be interpreted using the term pair thetic vs. categorical, used in information structure studies for topic less, all-new sentences and sentences with a topic.

Based on the multivariate results, we can conclude that this variation is only linguistically conditioned, since the linguistic factors of type of matrix and word order are found to have an effect on the selection of *ʔinnu* in AA, while other factors seemed to have less or in some cases a non-exiting effect.

The first possible conclusion that can be deduced from these results is that social demographic factors like education, age and sex have no statistically significant effects on the variation of *ʔinnu* in AA. It has been previously argued that syntactic variation is constrained less by extra-linguistic factors than by situational and linguistic constraints and that morpho-syntactic variation may rarely, if ever, distinguish social groups the way that phonological and morphological variants do (Winford 1996, 188, Hudson

1996, 45). A possible reason for this phenomenon could be that morpho-syntactic variants are below the level of conscious awareness; they are less frequently heard, and they are presumably less likely to be associated with a specific social class. The lack of social patterning of syntactic variants in different social communities demonstrated above forms clear evidence that there is a very marginal relationship between morpho-syntactic variation and the social world.

Bassiouney (2009) argues that it should not be taken as given that males and females speak differently merely because of their sex differences. She assures that research on gender has exceeded this presupposition. Holmes and Meyerhoff (2003b, 9) argue that when linguists make generalizations about a community at large, they apply their generalizations to both men and women. Gender, according to Bassiouney (2009), is still an important factor in studying language variation and change, but, it is a factor that interacts with other independent variables in a community, i.e., it has to be "put into context" (2003b, 9).

6. Conclusion

The scarcity of studies on the variable realization of the Arabic complementizer *ʔinnu* in AA is the main motivation behind conducting this study. Based on the variationist approach, this study has provided a quantitative analysis of the overt and null *ʔinnu* in a corpus of AA. With the aim of revealing the effects of some social and linguistic factors on the choice of *ʔinnu*, 10 hours of audio-recordings from 32 socio-linguistic interviews were quantitatively analyzed.

The multivariate analysis of the apparent time data has shown that extra-linguistic factors, i.e. age, gender, and education have statistically no significant impact on the realization of *ʔinnu*. Previous research in the field of linguistic variation in Arabic suggests that the difference of speech between men and women is considered a given. While this might be the case in some variants, the results of this research prove that gender has no active role in governing the realization of *ʔinnu* and that is likely because it is not stigmatized or correlated with any type of prestigious behavior. This finding corresponds with other socio-linguistic studies that suggest that morpho-syntactic variation is generally less conditioned by extra-linguistic factors than by linguistic factors (Winford 1996, 188; Hudson 1996, 45).

In addition, the analysis has further shown that word order and type of matrix were the only statistically significant factors in predicting the choice of overt and null *ʔinnu* in AA. This result indicates that nouns, prepositions and formulaic expressions such as 'maʃa:llah' (God bless), 'ʔilħamdu lila:h' (thank God), and 'wallah' (I swear) favor overt *ʔinnu*. This preference can be explained in relation with linguistic case assignment. Likewise, word order is second in in terms of its strength on variant choice with SVO and verbless sentences favoring overt *ʔinnu*. This result can be interpreted using the term pair thetic vs. categorical, used in information structure studies for topic less, all-new sentences and sentences with a topic.

This study has contributed to the variationist literature on vernacular Arabic by providing quantitative rather than intuited results that highlight discrepancies between "the analyst's conception of the data and the data themselves" (Ernestus and Baayen 2011, 374). Therefore, it is recommended for

future research to use vernacular Arabic in investigating complementizer omission in other varieties of Arabic as well as examining other levels of linguistic (phonological, lexical morpho-syntactic & semantic) variation.

تحليل تنوعي للحرف المصدري "إنه" في العربيّة العمانيّة

داليا أبو شريعة
قسم اللغة الإنجليزيّة واللغويات، جامعة العلوم والتكنولوجيا الأردنيّة، الأردن

مروان جراح
قسم علوم التأهيل، جامعة العلوم والتكنولوجيا الأردنيّة، الأردن

يزيد الحموري
اللغات الأوروبيّة، الجامعة الأردنيّة، عمان، الأردن

عقاب الشواشرة
قسم اللغة الإنجليزيّة وآدابها، جامعة اليرموك، الأردن

الملخص

تهدف هذه الدراسة إلى استكشاف حذف المكمل "إنه" في العربيّة العمانيّة باستخدام الأساليب الحسابيّة الحديثة للمنهج التنوعي، وتركز هذه الدراسة على تأثير بعض العوامل الاجتماعيّة واللغويّة مثل العمر والجنس والتعليم ونوع الجملة والتعريف والتخصيص وترتيب الكلمات ونوع الفاعل وصحة المحتوى في ظهور أو حذف المكمل "إنه" في العربيّة العمانيّة. وأظهرت النتائج أن استخدام أو حذف "إنه" في العربيّة العمانيّة مقيد بعوامل لغويّة تشمل نوع الجملة السابقة وترتيب الكلمات للجملة المضمنة "إنه"، وتلعب هذه العوامل دورًا ذا دلالة إحصائيّة في استخدام أو حذف "إنه" في العربيّة العمانيّة. وعلاوة على ذلك، أظهرت النتائج أن العمر والجنس والتعليم ليست ذات دلالة إحصائيّة في تقييد "إنه" في العربيّة العمانيّة.

الكلمات المفتاحية: اللغويات الاجتماعيّة التنوعيّة، حذف "إنه"، العربيّة الأردنيّة، القيود الاجتماعيّة واللغويّة، اختيار المتغيّر.

Endnotes

[1] Our sincere gratitude goes to two anonymous reviewers for their constructive criticism and suggestions.

References

Abdel-Aziz, Haneen, Marwan Jarrah, Abdel Rahman Altakhaineh and Ekab Al-Shawashreh. 2024. A Variationist Approach to NP Genitive Alternatives In Arabic. *Kervan. International Journal of African and Asian Studies* 28: 196-201.

Abd-El-Jawad, H. 1981. Lexical and phonological variation in spoken Arabic in Amman. Unpublished PhD diss, University of Pennsylvania.

Al-Hawamdeh, A. 2016. A Sociolinguistic investigation of two Hōrāni features in Sūf, Jordan. Unpublished Phd diss, University of Essex.

Aljabali, W. 2020. The Variable realizations of /k/ sound in Jordanian Arabic in Ajloun. Unpublished master's Thesis, University of Yarmouk, Jordan.

Al-Khatib, M. 1988. Sociolinguistic change in an expanding urban context: A case study of Irbid city. Unpublished PhD diss, University of Durham.

Al-Shawashreh, E. 2016. Aspects of Grammatical variation in jordanian Arabic. Unpublished PhD diss, University of Ottawa, Canada.

Al-Shawashreh, Ekab, Marwan Jarrah, and Eman Al Khalaf. 2024. Variation in Verbal Negation in Jordanian Arabic: A Corpus-based Analysis. *Lingua* 297. 103644.

Alshaboul, Asia, Marwan Jarrah, Sharif Alghazo & Ekab Al-Shawashreh. 2022. A Sociolinguistic Analysis of Intensifiers in Ammani Arabic. *Ampersand* 9: 100086. 10.1016/j.amper. 2022. 100086.

Alshakhanbeh, Sumaya and Sharif Alghazo. 2022. A Pragmatic Analysis of Criticism Strategies against Government Policies on Social Media in Jordan: A Gender-based Investigation. *Jordan Journal of Modern Languages and Literatures* 14 (2): 263-286.

Al-Wer, Enam. 1991. Phonological Variation in the Speech of Women from Three Urban Areas in Jordan. Unpublished PhD diss, University of Essex.

Al-Wer, Enam. 2000. Jordanian and Palestinian Dialects in Contact: Vowel Raising in Amman. *Essex Research Report in Linguistics* 18: 26-49.

Al-Wer, Enam. 2007. "The formation of the dialect of Amman: From chaos to order". In Dominique Caubet, Catherine Miller, Janent Watson (eds.), *Arabic in the City: Issues in dialect contact and language variation* (pp. 69-90). London: Routledge.

Al-Wer, Enam. 2009. "Variation". *Encyclopedia of Arabic Language and Linguistics* IV, 627-639 1-20.

Al-Wer, Enam, Uri Horesh, Deema AlAmmar, Hind Alaodini, Aziza Al-Essa, Areej Al-Hawamdeh, Khairia Al-Qahtani and Abeer Ab Hussain. 2022. Probing Linguistic Change in Arabic Vernaculars: A Sociohistorical Perspective. *Language in Society* 51(1): 29-50.

Bassiouney, Reem. 2009. *Arabic Sociolinguistics: Topics in Diglossia, Gender, Identity, and Politics*. Georgetown University Press.

Bayley, Robert. 2013. "The quantitative paradigm". In Jack Chambers and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change* (pp. 117-142). Oxford, UK: Blackwell.

Bloch, Ariel. 1965. Die Hypotaxeim Damaszenich-Arabischen. MitVergleichenzurHypotaxeimKlassisch-Arabischen, Wiesbaden: Harrassowitz.

Boberg, Charles. 2004. Ethnic Patterns in the Phonetics of Montreal English. *Journal of Sociolinguistics* 8 (4): 538-568.

Cleveland, Ray. 1963. A Classification for the Arabic Dialects of Jordan. *Bulletin of the American Schools of Oriental Research* 171(1): 56-63.

Cowell, Mark. 1964. *A Reference Grammar of Syrian Arabic (based on the dialect of Damascus)*. Washington, DC: Georgetown University Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.

El Salman, M. 2003. Phonological and morphological variation in the speech of Fallahis in Karak (Jordan). Unpublished PhD diss, Durham University.

Enç, M. 1991. The Semantics of Specificity. *Linguistic Inquiry* 22: 1-26.

Ernestus, Mirjam and Baayen Harald. 2011. Corpora and exemplars in phonology. *The Handbook of Phonological Theory* (2nd Ed.) (pp. 374-400). Wiley-Blackwell.

Germanos, Marie Aimee. 2010. From complementizer to discourse marker: the functions of ʔənno in Lebanese Arabic. In Owens, Jonathan and Alaa Elgibali, (Eds.), *Information Structure in Spoken Arabic* (pp. 145-164). Routledge, New York.

Habib, Rania. 2021. The Use of the Discourse Markers YaʕNi and ʔInnu: "I Mean" in Syrian Arabic. *Journal of Pragmatics* 178: 245–257.

Hamdieh, Aya, Marwan Jarrah, Abdel Rahamn Altakhaineh and Ekab Al-Shawashreh. 2022. Variation in Negation Patterns of Verbless Clauses in Ammani Arabic. *Jordan Journal of Modern Languages and Literatures* 14 (4): 1039 -1058.

Hudson, Richard. 1996. *Sociolinguistics* (2nd.ed). Cambridge: Cambridge University Press.

Holmes, Janet and Miriam Meyerhoff. 2003. Different voices, different views: An introduction to current research in language and gender. In Janet Holmes and Miriam Meyerhoff (eds), *The Handbook of Language and Gender*. (pp. 1-17). Oxford: Blackwell.

Jarrah, Marwan. 2019. Factivity and Subject Extraction in Jordanian Arabic. *Lingua* 219: 106-126.

Jarrah, Marwan, Rasheed Al-Jarrah, and Ekab Al-Shawashreh. 2022. No Case Tampering Once Transfer Domain Is Formed! *The Linguistic Review* 39 (2): 203-232.

Karmi, Simin (ed). 2008. *Word Order and Scrambling*. Wiley & Sons.

Khater, R. 2021. The realization of personal pronouns in Ammani Arabic: A variationist approach. Unpublished master's Thesis, University of Jordan.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1984. "Field methods of the project on linguistic change and variation. In: John Baugh and Joel Sherzer (eds), *Language in Use: Readings in Sociolinguistics*. Englewood Cliffs (pp. 28–53). NJ: Printice-Hall.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell Publishers.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.

Liang,Yiming, Pascal Msili and Heather Burnet. 2021. New Ways of Analyzing Complementizer Drop In Montréal French: Exploration of Cognitive Factors. *Language Variation and Change* 33: 359-385.

Milroy, Lesley and Matthhew Gordon. 2003. *Sociolinguistics: Methods and Interpretation*. Oxford: Blackwell.

Milroy, James and Lesley Milroy. 1977. Belfast: Change and variation in an urban vernacular". In Peter Trudgill (ed), *Sociolinguistic Patterns in British English* (pp. 19–36). London: Edward Arnold.

Obeidat, Hussein and Khadidja Hammoudi. 2020. Child Dialect Variation and Change: Case of Amman speech community, Jordan. *Jordan Journal of Modern Languages and Literatures* 12 (1): 17-30.

Omari, Osama and Aziz Jaber. 2019. Variation in the Acoustic Correlates of Emphasis in Jordanian Arabic: Gender and Social Class. *Folia Linguistica* 53 (1): 169-200.

Owens, Jonathan. 2001. Arabic Sociolinguistics. *Arabica* 48 (4): 419–469. Retrieved November 23, 2020, from http://www.jstor.org/stable/4057667

Owens, Jonathan, Bill Young, Trent Rockwood, David Mehall and Robin Dodsworth. 2010. "Explaining Ø and overt subjects in spoken Arabic. In Jonathan Owens and Alaa Elgibali (eds), *Information Structure in Spoken Arabic* (pp. 20–60). London: Routledge.

Palva, Heikki. 1994. Bedouin and Sedentary Elements in the Dialect of es-Salṭ. *Actes des Premières Journées Internationales de Dialectologiearabe de Paris*: 459-469.

Poplack, Shana and Sali Tagliamonte. 2001. *African American English in the Diaspora*. Oxford: Blackwell.

Poplack, Shana, David Sankoff and Christopher Miller. 1988. The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation. *Linguistics* 26: 47–104. 10.1515/ling.1988.26.1.47.

Poplack, Shana. 1993. Variation theory and language contact. In Dennis Preston (ed), *American Dialect Research*. (pp. 251-286). Philadelphia: John Benjamins.

Sankoff, David, Sali Tagliamonte and Eric Smith. 2005. Goldvarb X: A multivariate analysis application. Available at: http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.html.

Shim, Jin and Tabea Ihsane. 2017. English Overt and Null Complementizers. *Studies in Generative Grammar* 27 (3): 515-533. doi:10.15860/sigg.27.3.201708.515

Tagliamonte, Sali. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.

Tagliamonte, Sali. 2013. *Roots of English: Exploring the History of Dialects*. Cambridge: Cambridge University Press

Tauli, Valter. 1968. *Introduction to a Theory of Language Planning*. Uppsala: Almqvist and Wiksell.

Trudgill, Peter. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Oxford: Oxford University Press.

Weinreich, Uriel, William Labov and Marvin Herzog. (968. "Empirical foundations for a theory of language change". In Winfred Lehmann and Yakov Malkiel (eds), *Directions for Historical Linguistics* (pp. 95–189). Austin, TX: University of Texas Press.

Winford, Donald. 1996. "The problem of syntactic variation". In Jenifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter and Julie Solomon (eds.), *Sociolinguistic Variation: Data, Theory and Analysis* (pp. 177-192). Stanford: CSLI Publications.

Wolfram, Walt. 2006. "Variation and language: Overview". In Keith Brown (ed.), *Encyclopedia of Language and Linguistics* (pp. 333–341). Amsterdam: Elsevier.

Appendix (1)
List of Phonetic Symbols
The International Phonetic Alphabet for MSA

Arabic Script	IPA symbol	Symbols used in IVAr
أ- الهمزة	ʔ	ʔ
ب	b	b
ت	t	t
ث	Θ	th
ج	dʒ	J
ح	ħ	h
خ	x	x
د	d	d
ذ	ð	dh
ر	r	r
ز	z	z
س	s	s
ش	ʃ	sh
ش	ʧ	ch
ص	$s^ʕ$	S
ض	$d^ʕ$	D
ط	$t^ʕ$	T
ظ	$ð^ʕ$, $z^ʕ$	Dh, Z
ع	ʕ	3
غ	ɣ	Gh

ف	f	f
ق	q	q
ق	ɡ	ɡ
ك	k	k
ل	l	l
م	m	m
ن	n	n
ه	h	h
و	w	w
ي	j	y
V	v	v

Appendix (2)
Coding Protocol of the Data

Realization	code
Null	N
Overt	O
FG 1 Speaker's age	Code
Old (>50)	O
Middle Aged (35-50)	M
Young (18-35)	Y
FG 2 Speaker's sex	Code
Male	M
Female	F
FG 4 Education	Code
High	H
Low	L
FG 5 Type of matrix	Code
Verb	V
Adjective	A
Preposition	P
Noun	N
Other categories	O
FG 6 Factivity	Code
Factive	F
Non-Factive	N
FG 7 Word Order	Code
SVO	1
VSO	2
OSV	3

Verbless sentences	4
FG 8 Definiteness	Code
Definite	D
Indefinite	I
FG 9 Specificity	Code
Specific	S
Non Specific	N
Type of Subject	Code
Lexical	L
PRO	P
Non applicable	/