A Horn of Pepper or a Head of Onion: An Analysis of Semantic Variation of Classifiers in Jordanian Spoken Arabic from a Cognitive Sociolinguistic Approach

Abdel Rahman Mitib Altakhaineh *, Aseel Zibin and Lama Ahmed Khalifah

Abstract: This study examines the semantic variation in fruit and vegetable classifier usage in Amman, Jordan, employing a cognitive sociolinguistic approach. The semantic variation revolves around using idiomatic classifiers, such as ra’s bas’al (“head of onion”), in contrast to neutral classifiers, i.e., habbet bas’al (“a piece of onion”) or numerals, such as bas’al:n (“two onions”). This study focuses on classifiers used with fruits and vegetables, which are particularly relevant due to their physical shapes often prompting metaphorical classifiers and their tendency to take Arabic collective forms that are grammatically singular but semantically plural, complicating the issue of counting and potentially leading to the innovation of novel classifiers. The sample comprised 50 individuals from Amman, stratified based on their gender, age, and education. Data were collected through semi-structured interviews. The findings reveal a statistically significant inclination among older, male, and less formally educated speakers towards favoring idiomatic classifiers over the neutral options. This preference suggests that the choice between idiomatic and neutral classifiers may be influenced by social factors. We categorized the metaphors underlying the idiomatic classifiers as entrenched, conventionalized, and transparent, based on Müller (2009). The context of conventional metaphors demonstrates that the source domains of these metaphors could be active for a speaker at a specific moment but may not be active for another speaker at another moment, proposing that metaphoricity is not only a property of a linguistic item but also the cognitive achievement of a certain speaker. The preference for idiomatic classifiers, we argue, may be associated with notions of lower refinement, traditionalism, or reduced prestige.

Keywords: cognitive sociolinguistics; semantic variation; metaphor; Jordan; classifiers

1. Introduction

Sociolinguistics, according to Weinreich et al. (1968), relies on the assumption that the variability of language reflects the social organization of a community. Key studies that have contributed to this field include Labov’s investigations in Martha’s Vineyard (Labov 1963) and New York (Labov 1966, 1972), as well as Trudgill’s (1972) study in Norwich. Consequently, the examination of structured variation has been predominantly focused on phonology and morpho-syntax, as summarized in Chambers et al. (2002). However, the analysis of the onomasiological and semasiological aspects of linguistic structure rarely forms a central theme in sociolinguistic research. While studies exploring the geographical distribution of words (e.g., Orton and Dieth 1962; Upton and Widdowson 1999) and lexical borrowings (Poplack et al. 1988) have been conducted, investigations into onomasiological variation and change in the “Labovian” sense have only recently emerged (Boberg 2004). In addition, recent studies have also shown an interest in socio-semasiology. Noteworthy attempts to examine variations in meaning have been carried out within functional paradigms (Hasan 2009) or frameworks of discourse analysis (Stenström 2000; Tagliamonte and D’Arcy 2004; Cheshire 2007; Wong 2008). The outcomes of these studies
indicate that the construction of meaning could be constrained by socio-demographic factors and that changes in semantics could be affected by speakers’ intention to convey different aspects of their identity. As a result, these studies suggest that further exploration of semantic variation is a worthwhile pursuit. We also hold the belief that societal structural changes call for increased focus on matters of sociolinguistic variation in discourse. This is because the characteristics of such variation may gain significance for users and potentially unveil an ongoing process of language change.

By observing the classifiers employed by native speakers of Jordanian Spoken Arabic (henceforth JSA), a dialect of spoken Arabic in Jordan, which is located in the Middle East, and by being native speakers of this dialect ourselves, we have noted a distinction in their preferences when classifying fruits and vegetables. Some speakers tend to use idiomatic expressions that include conventionalized metaphors that can either be activated or not (see Section 2.1), such as ras bas'al (“a head of onion”) or garni:n filfil (“two horns of pepper”), while others utilize neutral classifiers for all types of fruits and vegetables, i.e., habbeh (“one piece”), or numerals, e.g., habbe:n farawlalah (“two strawberries”) (see Jaradat and Jarrah 2022). In this context, a “neutral classifier” refers to a term whose usage has become highly generalized, transcending its original nominal meaning. In this study, we refer to the classifier “habit-” as “neutral” because it does not carry strong sociolinguistic connotations or associations with specific social groups. In other words, habbeh can likely refer to a “piece” of nearly anything, while the other terms are limited to certain fruit or vegetable nouns. It is also termed “modern” and “contemporary” to indicate its prevalent usage in present-day Jordanian Arabic, especially in urban contexts. Regarding variants, in this paper, we use the variant garin to reflect the use of the idiomatic classifier because it is the most commonly used variant among Jordanians; however, it is important to note that other variants, such as karin and karin, also exist. The aim of the paper is not to examine variations in dialects; whether the respondent used garin, karin, or karin, our results would still be the same because the respondent has used the idiomatic classifier. These variations in dialects did not have a bearing on the results associated with the type of classifier used in our data, and thus, dialect variation was not a variable.

Through examining these different uses, we have identified potential influences from social factors such as age, gender, and education. We chose these factors based on our own observations and on previous studies that have demonstrated their influence on the linguistic expressions employed by Jordanians (refer to Al-Hloul et al. 2023; Zibin et al. 2024a). As the first option for classifiers involves the use of idiomatic expressions that may have emerged through metaphors, our study aims to explore the semantic variation in the utilization of these classifiers by native speakers of JSA from a cognitive sociolinguistic perspective. Specifically, this study aims to provide answers to the following research questions:

- To what extent is the use of idiomatic versus neutral classifiers for fruits and vegetables in Amman socially influenced by age, gender, and education?
- What are the social meanings associated with using idiomatic classifiers and neutral classifiers in Amman?
- For a better understanding of the Arabic fruits, vegetables, and classifiers, below is a table that shows each Arabic term used throughout the study and its equivalent:

<table>
<thead>
<tr>
<th>Fruit/vegetable name</th>
<th>English equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>bas'al</td>
<td>Onion</td>
</tr>
<tr>
<td>mo:z</td>
<td>Banana</td>
</tr>
<tr>
<td>filfil</td>
<td>Pepper</td>
</tr>
<tr>
<td>ðu:m</td>
<td>Garlic</td>
</tr>
<tr>
<td>bando:rah</td>
<td>Tomato</td>
</tr>
<tr>
<td>baf'ti:x</td>
<td>Watermelon</td>
</tr>
<tr>
<td>ðurah</td>
<td>Corn</td>
</tr>
<tr>
<td>farawlalah</td>
<td>Strawberries</td>
</tr>
</tbody>
</table>
2. Literature Review

2.1. Theoretical Framework

Meaning has been recognized as the most crucial aspect of linguistic structure within cognitive linguistics (Geeraerts and Cuyckens 2007, p. 14), as it views semantic structure as a flexible reflection of speakers’ perceptions and their adaptation to the physical and cultural reality that they interact with. Consequently, one would expect that variation in meaning is regularly addressed in cognitive research. Indeed, lexical meaning variation has been extensively studied within this framework, particularly in the context of corpus research (e.g., Beeching 2005; Divjak 2006; Gries and Divjak 2009). Conversely, the relationship between lexical variation and external sociolinguistic factors has been relatively underexplored. Geeraerts et al.’s (1994) study on onomasiological variation in the context of dialect development is a seminal work in this area. Another important study is that of Robinson (2010), who examined the semantic variation in the use of the adjective awesome in English from both a cognitive semantic and sociolinguistic perspective. Regarding semasiological variation, results also suggest that the use of different concepts can be explained by various interacting factors, with dialect being one of the most significant.

While both sociolinguistics and cognitive linguistics theoretically agree on the variable nature of meaning, the exploration of semantics has been approached to different extents within these frameworks. In cognitive semantics, arguments are often made at an abstract level, while in sociolinguistics, these arguments are usage-based. This study adopts a theoretical framework comprising two parts: cognitive semantic analysis and language variation. First, a cognitive semantic analysis of fruit and vegetable classifiers is conducted following Müller (2009) and a sociolinguistic analysis is carried out based on Labov’s variationist approach. With regard to the former, the fundamental concept in cognitive metaphor theory revolves around the distinction between dead and alive metaphors, and therefore, a thorough examination of this differentiation is pertinent to any analysis of metaphors (Müller 2009, p. 10). Müller’s (2009) classification is adopted in this study because he introduced a new classification that takes into account the usage aspect, establishing a tripolar classification of “dead”, “novel”, and “entrenched” verbal metaphors based on the criteria of conventionalization, novelty, and transparency. According to this classification, dead metaphors are highly conventionalized and opaque, novel metaphors are not yet conventionalized and transparent, and entrenched metaphors are both conventionalized and transparent (Müller 2009, p. 11). This new classification depends on the criterion of cognitive activation within a specific speaker and moment in time (Müller 2009, p. 11).

The advantage of this classification lies in its scalar nature, allowing for the recognition of various degrees of metaphoricity instead of viewing metaphors as strictly falling into two rigid categories, i.e., dead and alive (ibid). For instance, when examining conventional verbal metaphors, it becomes clear that the source domains of these metaphors may be mentally accessible for one speaker at a particular time but not for another speaker at a different time. This suggests that the quality of being metaphorical is not solely inherent to a linguistic expression, but, rather, is dependent on the cognitive processes of individual speakers (Müller 2009). Thus, following the dynamic view of Müller, we explain the metaphoricity of the target idiomatic classifiers, which we argue are entrenched metaphors, and then analyze them not as either dead or alive metaphors, but sleeping [metaphors that have a low or no degree of activation] or waking [metaphors that have a high degree of activation]. We hypothesize, based on our experience, that elderly males and females who are not formally educated are more likely to activate conventionalized metaphors such as ra:s basāl (“a head of onion”), garin fifif (“a horn of pepper”), and the like, while, for young

<table>
<thead>
<tr>
<th>• Idiomatic classifier</th>
<th>• English equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>• ra:s</td>
<td>• Head</td>
</tr>
<tr>
<td>• garin</td>
<td>• Horn</td>
</tr>
<tr>
<td>• sinn</td>
<td>• Tooth</td>
</tr>
<tr>
<td>• zirr</td>
<td>• Button</td>
</tr>
<tr>
<td>• ḍarnuːs</td>
<td>• Cob</td>
</tr>
</tbody>
</table>

*Correspondence: a.altakhaineh@ju.edu.jo*
male and female educated participants, these metaphors are either sleeping or not used at all.

Regarding the second part of our theoretical framework, Labov (1972, p. 271), often considered the founder of variationist sociolinguistics, argues that the study of language variation becomes essential when there are “two or more ways of saying the same thing”. Al-Wer (2009, p. 1) asserts that “variation is an inherent characteristic of every human language”, indicating that in any language, multiple ways exist to express the same idea, and individuals do not speak in the same manner consistently across all situations. Additionally, Wolfram (2006, p. 333) suggests that if structure is the core of language, then variation defines its essence. Variationists aim to quantitatively describe linguistic patterns as a key component in providing accurate explanations of language variations and predicting language change. The study of variation in language encompasses everyday vernaculars, examining how language varies among different groups of speakers and elucidating the underlying relationship between this variation and social factors. Tagliamonte (2011, p. 20) describes variationist sociolinguistics as the branch of linguistics that examines the fundamental characteristics of language, taking into account both linguistic structure and social structure, as well as grammatical and social meanings. It acknowledges that explanations for these language properties require reference to both external (social) and internal (systemic) factors. Variations are not limited to sounds alone; they can be observed in pronunciation, word choice, and grammatical structures within the same speech community. While numerous studies on phonological and syntactic variation have been conducted on JSA (see Al-Wer 2009; Omari and Jaber 2019; Al-Rawafi et al. 2020; Al-Hloul et al. 2023), a study that examines potential semantic variation in this Arabic variety is non-existent to the best of our knowledge, which is the main motivation for this study.

Thus, this study adopts a cognitive–sociolinguistic framework, which can be regarded as a relatively new field that examines language use through drawing on the convergence of methods and theoretical frameworks that are typically linked to the two fields of study (Pütz et al. 2014). In this framework, researchers aim to provide a more comprehensive understanding of the formation and variation of meaning and also explain the meaning variation that is observable across languages and cultures (Pütz et al. 2014). That is, in cognitive sociolinguistics, it is argued that a usage-based analysis of language can be performed only if social and cultural factors that shape usage events “are systematically considered alongside the cognitive ones” (Pütz et al. 2014, p. 6). In addition, cognitive sociolinguistics also explores speakers’ implicit and explicit perception of and attitudes to linguistic variation. It also investigates how cultural norms appear in individual usage events and studies the extent to which cultural, social, and cognitive factors are related (Pütz et al. 2014).

Adopting a cognitive sociolinguistic approach not only facilitates the analysis of language change in progress but also moves beyond the scrutiny of changes that have already transpired. This affords us a deeper comprehension of the underlying mechanisms propelling language change and the ability to forecast forthcoming changes. The findings of this study emphasize the influential role of social variables, encompassing age, gender, and education, in shaping language variation and change. Through investigating whether the selection of classifiers in Jordanian Spoken Arabic (JSA) is socially constrained, while also considering the potential influence of activation of entrenched metaphors on the choice between idiomatic and neutral classifiers, we can obtain a more comprehensive understanding of the motivations behind language change and its societal ramifications.

2.2. Previous Studies

In the field of sociolinguistics and cognitive linguistics, some studies have been conducted to investigate the phenomenon of semantic variation. These investigations generally aim to reveal the complexities of word meaning and how it can vary across different contexts and variables. For instance, Schmid et al. (2021) examined the variation in the usage of specific lexico-grammatical patterns, focusing on the pattern “that’s adj” (e.g., “that’s right”),
“that’s nice”, or “that’s okay”), and they found that adjective choices are influenced by both social and individual factors, emphasizing the importance of considering individual routines and entrenchment in language use. This study highlights the interplay between social and individual variation, suggesting that individual variation is fundamental to understanding broader social linguistic trends and language change. Similarly, 

Stratton and Sundquist (2022) employed variationist sociolinguistic methods to investigate the intensifier system in Oslo Norwegian, revealing that both linguistic and social factors influence intensifier usage. Their findings showed that predicative adjectives are intensified more often than attributive ones, and social factors such as gender and age significantly affect intensifier usage. The study illustrated how changes in language use, like the increasing use of “skikkelig” (meaning “proper”) among young women, can reflect broader sociolinguistic patterns and ongoing language change.

In the context of JSA, Al-Tamimi (2001) explored phonetic and phonological variation among Fallahi migrants in Irbid, focusing on variables such as social class, gender, education, and age. The study found that non-local urban features are spreading among rural migrants, with social class and gender playing significant roles in language variation. Women were observed to be more innovative in adopting non-local features despite cultural constraints, while men tended to retain local indigenous features as a marker of identity. The research underscored the dynamic interaction between social factors and language use, particularly in diglossic communities where multiple prestigious varieties compete.

Although variation among classifiers has not been investigated, the previous studies suggested that social class, origin, age, and gender play a key role in variation. This is also demonstrated by Al-Shawashreh (2016), who examined grammatical variation in vernacular Jordanian Arabic (JA), focusing on word order and pro-drop variation. Using a variationist sociolinguistic framework, his study analyzed a corpus recorded in the Irbid metropolitan area, considering factors such as age, sex, education, and urban/rural origin. The findings revealed that SV(O) word order is predominant, with VS(O) order less frequent, and that null subject pronouns are common in vernacular JA. Key predictors of word order and pro-drop choices include transitivity, definite subject pronouns, switch references, and the person and number of subjects. Social factors also played a significant role, with age, sex, and urban/rural origin influencing variation. The study highlighted the value of empirically analyzing spontaneous speech to uncover syntactic variation in modern spoken Arabic, offering new insights beyond those obtained from intuited, elicited, or written data.

Similarly, Omari and Jaber (2019) explored the impact of gender and social class on the acoustic characteristics of emphasis in Jordanian Arabic. By recording the reading of minimal pairs and analyzing various acoustic measurements, they found significant differences between genders and social classes. Gender differences were evident in F1 and F2 at onset and midpoint positions, while social class differences were noted in F1 and F2 at onset, in vowel duration, and in the post-release duration of voiced emphatic stops. Lower-middle-class speakers preferred stronger emphasis cues. These findings showed relationships between gender, social class, and the acoustic correlates of emphasis, contributing to our understanding of the interplay between sociolinguistic factors and linguistic variation in emphasis expression.

The reviewed studies collectively support the current research by demonstrating the multifaceted nature of language variation, influenced by both individual and social factors. They highlight the importance of examining how social variables affect language use in order to understand language change and variation. By situating the current study within this framework, we can highlight its contribution to the broader understanding of how social, cognitive, and other factors interact to shape language use.

Robinson (2012) argued that sociolinguistic investigations of semantics have been limited and that this gap persists. While many studies on JSA have focused on phonological or syntactic aspects, there has been no exploration of semantic variation in JSA. Addressing this gap, this study examines semantic variation in classifiers related to fruits and vegetables.
among native Jordanian speakers, analyzing how social factors such as age, gender, and education influence classifier use. By investigating these demographic variables, this study aims to contribute to a more comprehensive understanding of the factors shaping language use and variation in this specific domain.

3. Methodology

3.1. Sample

The sample in this study consisted of 50 participants, including 25 males and 25 females, residing in Amman, the capital city of Jordan. The participants were stratified according to age based on Lawton et al.’s (1992) classification: young adults (18–29), middle-aged (30–59), and older adults (60+) (see Al-Hloul et al. 2023). Of these, 17 participants were young, 17 were middle-aged, and 16 were older. The participants were also stratified according to their educational level. In this study, we define “educated” individuals as those who have completed at least a secondary education (12 years of schooling) and have graduated from a university or other higher education institution. “Non-educated” individuals are those who have not completed secondary education and therefore have not pursued higher education. This classification is chosen based on research indicating that formal education significantly affects linguistic usage. While Al-Wer (2013) highlighted that education serves as a proxy variable, reflecting the breadth and nature of an individual’s social contacts, in this study, we show that education level is in fact a factor in sociolinguistic analysis. Specifically, 25 participants were classified as educated, and 25 were considered non-educated. The population of Amman is approximately 4,000,000 people, and the sample was selected from that population. Table 1 below, displays the sample’s characteristics.

Table 1. Characteristics of the sample.

<table>
<thead>
<tr>
<th>Level of Education</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y</td>
<td>M</td>
<td>O</td>
</tr>
<tr>
<td>Educated</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Non-educated</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>

The sampling procedure we followed is stratified sampling, which is a sampling technique used by researchers to ensure that a representative sample is drawn from a population with distinct subgroups or strata (Rahman et al. 2022). In our case, the subgroups are based on social factors, namely, age, gender, and education. The goal of stratified sampling is to reduce potential bias and increase the precision of estimates by ensuring that each subgroup’s characteristics are properly represented in the sample. By using stratified sampling, the aim was to address potential issues that might arise if a random sample technique was used without considering the distribution of age, gender, and education. This method increases the likelihood of the sample being representative of the entire population and can lead to more accurate and reliable results.

3.2. Data Collection

Data were collected through semi-structured interviews conducted with 50 participants. After dividing the population into the age-based strata, data were collected by approaching individuals in public places and asking them to take part in an interview which lasted for approximately 20 min. That is, two researchers (one male and one female) spent 4 weeks going to different public places in Amman, including malls, shopping centers, restaurants, and universities, to collect data from participants that fitted into our strata or subgroups. This approach helped ensure that the sample included individuals from each stratum, which allowed us to make more accurate and meaningful inferences about the entire population. Drawing our sample from a diverse range of locations such as malls, restaurants, universities, and shopping centers does not automatically introduce
bias towards middle- and upper-class participants. For instance, individuals from low-income backgrounds may also go to affordable restaurants, while public universities may cater to students supported by scholarships based on academic merit, a military service background, or being from underprivileged areas. Additionally, shopping centers and malls attract a wide spectrum of visitors, including those who may simply browse without making purchases or opt for lower-cost items. Therefore, by considering these various factors, we aim to capture a more comprehensive representation of the linguistic diversity across different socio-economic strata in Jordanian Spoken Arabic.

The participants were informed that the study was about language use but were not informed about the specific aim of the study, so that they would not be self-conscious about their language and would thus hopefully produce naturalistic data (see Al-Hloul et al. 2023). The interviewers started by asking the participant questions about their age and education and some other questions to break the ice (e.g., are you enjoying your time at the mall?), then moved to the interview questions. The interviews consisted of two types of questions. The first type involved participants filling in the blanks with a classifier, specifying the number of the fruit or vegetable. The number of items was seven, namely, bas‘al (“onion”), filfil (“pepper”), ẓo:m (“garlic”), bandō:rah (“tomato”), ẓurah (“corn”), mo:z (“banana”), and bat‘i:x (“watermelon”). Note that the last two items were included as distractors, because they are typically used with the classifiers garin and ras, which are already associated with bas‘al (“onion”) and filfil (“pepper”), respectively. The first type of question included the following sentences:

\[\text{flārē}:t \ \text{ālījō}:m \ \text{ba}s‘al \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 2) \ “\text{Today I bought} \ \text{onions from} \ \text{the market}” \ (\text{amount} \ 2).\]
\[\text{flārē}:t \ \text{ālījō}:m \ \text{mō}:z \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 5) \ “\text{Today I bought} \ \text{from} \ \text{the market}” \ (\text{amount} \ 5).\]
\[\text{flārē}:t \ \text{ālījō}:m \ \text{fi}lfil \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 3) \ “\text{Today I bought} \ \text{from} \ \text{the market}” \ (\text{amount} \ 3).\]
\[\text{flārē}:t \ \text{ālījō}:m \ \text{ẓo}:m \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 4) \ “\text{Today I bought} \ \text{from} \ \text{the market}” \ (\text{amount} \ 4).\]
\[\text{flārē}:t \ \text{ālījō}:m \ \text{bandō}:rah \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 6) \ “\text{Today I bought} \ \text{from} \ \text{the market}” \ (\text{amount} \ 6).\]
\[\text{flārē}:t \ \text{ālījō}:m \ \text{bat‘i}:x \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 1) \ “\text{Today I bought} \ \text{from} \ \text{the market}” \ (\text{amount} \ 1).\]
\[\text{flārē}:t \ \text{ālījō}:m \ \text{ẓurah} \ \text{mīn} \ \text{su}:g \ (ṭīlkmījeh \ 3) \ “\text{Today I bought} \ \text{from} \ \text{the market}” \ (\text{amount} \ 3).\]

This type of question ensured that the participants produced the target expressions that we were interested in. The second part of the interviews included some qualitative, open-ended questions answered after the fill-in-the-blank question. These questions aimed to explore the reasons behind the use of specific classifiers, why certain groups of people use one type of classifier more than others, and what this choice reflects about their characteristics such as gender, age, and education level. Examples of the questions asked in this part are as follows:

1. When you go to buy vegetables or fruit from the grocer, do you usually say biddī garin filfil, ḥarīn filfil, karīn filfil (“I want a horn of pepper”), whatever your dialect? If yes, why?
2. If you do not use words like garin (“horn”), ẓarnu:s (“cob”), zirr (“button”), ra:s (“head”), etc., to order vegetables and fruit, then what words do you use? Why?
3. Who do you think would use ra:s bas‘al (“head of onion”), garin mo:z (“a horn of banana”), and garin filfil (“a horn of pepper”) more? Older speakers or younger ones? Why?
4. Who do you think would use ra:s bas‘al (“head of onion”), garin mo:z (“a horn of banana”), or garin filfil (“a horn of pepper”) more? Males or females? Why?
5. Who do you think would use ra:s bas‘al (“head of onion”), garin mo:z (“a horn of banana”), or garin filfil (“a horn of pepper”) more? Educated or non-educated speakers? Why?
6. Where do you think these words, e.g., garin, ẓarnu:s, zirr, etc., come from?
7. The interviews were tape-recorded and transcribed to be analyzed later. The procedure followed to analyze the data is provided in the following section.

3.3. Data Analysis

The first part of the interview (fill in the blank) was analyzed quantitatively using descriptive analysis (means and standard deviation), and to decide whether the three variables (age, gender, and education) have a statistically significant effect on the use of the two types of classifiers (idiomatic or neutral), Chi-square was used. With regard to the second part of the interview, thematic analysis was used. Thematic analysis is a qualitative research method used to analyze and identify patterns, themes, and insights within textual data, such as responses to open-ended questions in interviews or surveys. It is commonly used to uncover underlying meanings, attitudes, and experiences in participants’ responses. Thematic analysis allows researchers to explore the richness and complexity of participants’ responses, which offers insights that might not be captured by quantitative methods alone (Nowell et al. 2017; Altakhaine et al. 2024). We read the transcripts of the interviews carefully, familiarized ourselves with the data, categorized data into preliminary themes, reviewed and refined the themes, and wrote a description of each theme supported by excerpts from the participants. To avoid data analysis bias, we used Cohen Kappa to measure the inter-rater reliability between us (three researchers); the score was 0.81, which is excellent. With regard to the cognitive linguistic analysis, we examined the metaphors underlying the idiomatic classifiers using Müller’s (2009) classification. To ensure the reliability of our analyses, we employed Cohen’s Kappa, yielding an excellent score of 0.82.

3.4. Ethical Considerations

The participants were provided with an explanation of the research’s general objectives, procedures, and benefits. They were given the opportunity to ask questions and were fully informed before granting their voluntary consent to participate. The informed consent process emphasized the voluntary nature of participation and their right to withdraw at any point without any repercussions. To protect participants’ privacy, all collected data were coded and stored on secure, password-protected servers. Participants’ identities were kept strictly confidential, and their responses were reported in aggregate form to ensure anonymity (see Appendix A).

4. Results

The interviews yielded idiomatic classifiers such as garin, ras, zirr, and others, a neutral classifier which can be used to quantify all types of fruits and vegetables, i.e., habbeh (“piece”), and numerals such as basal’ta:n (“two onions”). The last category was not used frequently, since it is usually used when the order is just one or two pieces of fruit or vegetables, and even when the quantity is 1 or 2, it is marked for use with some fruit and vegetables, such as ?filfilte:n (“two peppers”). To provide an answer to the first research question, Table 2, below, provides the results of the Chi-square test in relation to the use of idiomatic vs. neutral classifiers by males and females.

<table>
<thead>
<tr>
<th>Classifiers</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
<th>Chi-Square Value</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiomatic classifiers</td>
<td>78</td>
<td>22</td>
<td>100</td>
<td>52.27</td>
<td>0.000</td>
</tr>
<tr>
<td>Neutral classifiers</td>
<td>47</td>
<td>103</td>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>125</td>
<td>125</td>
<td>250</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows that there are differences in classifier choice attributed to gender, as evidenced by the Chi-square value (52.27), which is statistically significant at 0.05. Cross-tabulations reveal that males tend to use the idiomatic classifiers, while females tend to
use the neutral ones. Table 3 below shows the results of the Chi-square test with regard to classifier choice based on age.

Table 3. Chi-square results in relation to the choice of classifier based on age.

| Idiomatic Classifiers | Younger | Middle-Aged | Older | Total | Chi-Square Value | Sig.  
|-----------------------|---------|-------------|-------|-------|-----------------|------
|                       | 13      | 33          | 56    | 102   |                 |      
| Neutral Classifiers   | 72      | 52          | 24    | 148   | 51.27           | 0.000 
| Total                 | 85      | 85          | 80    | 250   |                 |      

Table 3 indicates that there are differences in classifier choice attributed to age, as shown by the Chi-square value (51.27), which is statistically significant at 0.05. Cross-tabulations reveal that individuals in the younger and middle-aged age groups tend to use the neutral classifiers, whereas older individuals tend to use the idiomatic classifiers. Table 4, below, shows the results of the Chi-square test with regard to classifier choice based on education.

Table 4. Chi-square results in relation to the choice of classifier based on educational level.

| Classifiers       | Educated | Non-Educated | Total | Chi-Square Value | Sig.  
|-------------------|----------|--------------|-------|-----------------|------
| Idiomatic Classifiers | 23       | 94           | 117   | 80.99           | 0.000 
| Neutral Classifiers   | 102      | 31           | 133   |                 |      
| Total               | 125      | 125          | 250   |                 |      

Table 4 demonstrates that there are differences in classifier choice attributed to educational level, as evidenced by the Chi-square value (80.99), which is statistically significant at 0.05. Cross-tabulations reveal that educated individuals tend to use the neutral classifiers, while those who are not formally educated tend to use the idiomatic ones.

The second part of the interviews conducted in this study focused on the qualitative aspect of classifier use, aiming to unravel the underlying reasons for the selection of specific classifiers and the social meanings associated with the use of these classifiers, providing an answer to the second research question. Furthermore, the study sought to uncover the correlation between these choices and the sociodemographic characteristics of individuals, i.e., gender, age, and education. During the interviews, participants provided valuable insights into their rationale for choosing specific classifiers. The decision to use a particular classifier often depends on a complex interaction of factors. These factors could include familiarity, accuracy, interpretability, prestige, and age. Younger participants highlighted their preference for simpler, more neutral classifiers due to their ease of implementation and interpretation, especially when they deal with the public. For example, one young female participant stated the following:

أنا ما بحكي رأس ووزر، بينين قدعيين وشعيين

ًانا ما: buḥki ra:s wu zirr, bibajnu: ًادي:mi:n wu ُجك:جي:n

“I do not say ra:s or zirr they sound old and vulgar.”

Another young female, when asked about garin when talking about bananas, said the following:

هة فرن، زي سي! وستي ما بتعرف تقرأ وتكتب

haha garin, zajj sitti! wu sitti ma bti'ruf titra: wu tuktub
“haha garin, this sounds like my grandma! I mean my grandmother is illiterate”

Another middle-aged male commented that he uses both types of classifiers, moving lately towards the more neutral expressions because the other ones are “old-fashioned”; he also added the following:

لَا أُشُوف بصلة بندكرني بالروس وعشان هياد إذان بنحكي راس بصل

lamma ʔaʃːuː basːaleh bitːakkirmi bikrːuː s wu ʔaʃːan heː:k ʔana bahki raːs basːal

“When I see an onion, it reminds me of heads, and that is why I say raːs basːal”. He also added the following:

طيب، لما بدأ أكثر من حبيت، شو بدك مكحك؟ بدي 3 فلفلات؟ معقول؟

tˤajib, lamma biddak ʔakˤar min ḥabbten, ʕuː biddak tehːki? biddi ʔalaʔ filfilːat? maʔguːl?

“Besides, if you want more than two pieces, what would you say? biddi 3 filfilːat (I want 3 peppers)!! This sounds ridiculous!”

Therefore, younger participants often favored modern expressions; meanwhile, older participants demonstrated a predilection for traditional (idiomatic) classifiers, which could be attributed to familiarity.

Furthermore, the influence of gender on language choices emerged as a significant aspect of the study. Among the participants, it became evident that gender exerted a notable sway over the selection of idiomatic expressions. Older males showed a strong tendency to use idiomatic classifiers, often choosing phrases that are connected to their culture or area. After carefully looking at the interviews, it was evident that older male participants preferred using idiomatic expressions that felt traditional and familiar. However, younger males are less inclined to use these expressions, similar to younger females.

Contrastingly, the speech of female participants appeared to be marked by a preference for more neutral and contemporary expressions. This preference was highlighted by one female participant, who articulated the following:

انا ما بحب أحكي راس بصل ورس زهرة جنس الزلام محكمهم وبيبنوا دفشي وقدم

ʔana maː bahibb ʔahki raːs basːal wu raːs zahra, bahiss ʔizzłaːm biːkuːhum. wu bibajnu difːjːn wu ʔdaːm

“I tend to avoid phrases like raːs basːal ‘head of onion’ or raːs zahra ‘a head of cauliflower’; they seem to be said by older males and carry a sense of masculinity and antiquity.”

This sentiment was echoed by several female participants, who shared a similar sentiment, attributing these expressions to an older generation.

The gendered linguistic divergence extended beyond mere phrase preference. Participants provided details on their family dynamics, offering insights into how these expressions were distributed across generations, educational levels, and genders. A participant observed the following:
A Horn of Pepper or a...
between the pointed, elongated shape of a pepper and the form of a horn. *sinn tho:m* draws on the similarity in shape between a garlic clove and a tooth, with the metaphor found in the resemblance between the shape and size of a clove of garlic and a tooth (“GARLIC IS A TOOTH”). *zirr bandorah* metaphorically likens a tomato to a button, reflecting its small, round form (“A TOMATO IS A BUTTON”). Finally, *sarnu s duraḥ* (“vertical woody piece of corn”) also metaphorically likens the shape of corn to a vertical piece of wood on which the pieces of corn are found. Hence, the polysemy observed in the idiomatic classifiers reflects the nature of language, where words and expressions evolve through metaphorical extensions, acquiring new meanings.

The metaphors underpinning classifiers such as *ras basal*, *garin filfil*, *sinn tho:m*, *sarnu s duraḥ*, and *zirr bandorah*, as delineated by Müller (2009), are deeply entrenched in the linguistic system of certain speakers, with their activation influenced by factors such as gender, age, and education. The conventionalization stems from the widespread acceptance and standardized usage within the linguistic community (cf. Pütz et al. 2014). The transparency, on the other hand, emerges from the visual resemblance between the shapes of these fruits or vegetables and the objects that they metaphorically represent. It can also be observed that a metonymy is at play here, where the shape of the fruit or vegetable is used to refer to it, giving rise to the metonymy (SHAPE FOR FRUIT/VEGETABLE).

In addition, the results of the study reveal distinctions in the use of the idiomatic classifiers vs. the neutral classifier *habbeh* (“piece”) as in *habbet (bas'al, filfil, tho:m, bandorah)* (“a piece of onion, pepper, banana, garlic, tomato”), as well as the numerals, among different demographic groups in Amman. These distinctions were statistically significant and shed light on the influence of age, gender, and education on the use of classifiers. With regard to age, the preference for entrenched metaphors such as *ras basal* (“head of onion”), *garin mo:z* (“a horn of banana”), and similar expressions among older participants highlights the generational aspect of linguistic choices. These deeply ingrained and highly conventionalized metaphors reflect a linguistic continuity with traditional language usage. This tendency may be preferred by older speakers, aligning with findings from studies like Stratton and Sundquist (2022) that suggest that age can have an impact on the choice of intensifiers.

Concerning gender, the observed variations in classifier usage point to the complex sociolinguistic landscape within the Ammani community (see Altakhaineh et al. 2024). The preference for specific metaphors can be attributed to cultural and societal factors that shape language use. For instance, the inclination of males towards employing entrenched metaphors may be reflective of their exposure to specific linguistic patterns in their social environments. This aligns with research by Trudgill (1986), Al-Tamimi (2001), Cheshire (2002), and Eckert and Rickford (2002), as well as Al-Hloul et al. (2023), which suggests that gender can have an effect on linguistic choices and that females may lean towards using words associated with prestige and sophistication. This also supports Trudgill’s (2000) idea that females are more innovative, as they have abandoned the “old” terms while males continue to perpetuate them.

In terms of education, its influence on the use of classifiers is particularly evident in this study’s findings. Educated participants displayed a greater propensity to employ the neutral classifier *habbeh* (“piece”). This inclination may stem from a desire to distance themselves from terms associated with lower education levels or a perceived lack of sophistication. This resonates with the observations made by Omari and Jaber (2019), further emphasizing the link between linguistic choices and education and social class.

Regarding the neutral classifier *habbeh*, it initially referred to a concrete, countable entity—“seed” or “grain”. This specific meaning is deeply rooted in the physical and agricultural contexts of early Arabic-speaking communities. Over time, *habbeh* extended metaphorically to denote “a small thing”, broadening its application from tangible seeds to any small item. This shift demonstrates a common linguistic process where concrete nouns acquire more abstract, generalized meanings. In contemporary Jordanian Spoken Arabic, *habbeh* functions as a classifier, indicating “a piece” of something. This usage reflects further
semantic bleaching, where the term loses its original specific reference to seeds and gains a broader, more generalized use.

For the analysis, the quantitative data (from Chi-square tests) and qualitative data (from interviews) complement each other, as the qualitative data contextualize statistical findings by uncovering the socio-cultural meanings attached to language choices. Together, these data types provide a comprehensive understanding of how demographic factors influence language choice. For example, the quantitative analysis reveals that older participants tend to use idiomatic classifiers more frequently, while the qualitative interviews explain this preference as being rooted in tradition and familiarity. Additionally, while Table 2 shows a significant gender difference in classifier usage, qualitative insights explain why females might prefer neutral classifiers due to perceptions of modernity and sophistication, contrasting with older males, who favor traditional idiomatic expressions.

These results align with the framework adopted in this study, namely, cognitive sociolinguistics, supporting both Müller’s cognitive semantic analysis and Labov’s variationist approach. The dynamic view of metaphoricity, proposed by Müller, emphasizes that the activation of metaphors is context-dependent and can vary within individual speakers and moments in time. In line with Labov’s variationist approach, it was suggested that the use of idiomatic vs. neutral classifiers in Amman is socially constrained by factors, including age, gender, and education (see Pütz et al. 2014). Younger, female, and educated speakers may be less inclined to activate idiomatic classifiers, possibly due to considerations of social prestige and a preference for more neutral classifiers. The avoidance of these classifiers by these speakers may stem from an association with older, less refined speech, while educated individuals might opt for more neutral linguistic forms.

Such language attitudes offer insights into how speakers view different linguistic forms, yet these attitudes do not always translate into actual linguistic behavior (see Lægsgaard 2000; Garrett 2010; Dragojevic 2018). This discrepancy highlights that while speakers may express certain preferences towards particular fruit and vegetable classifiers, these preferences do not necessarily reflect their choices in real-world usage. Thus, the qualitative data in this paper focus on capturing these perceptions, emphasizing that language attitudes are complex and multifaceted, often existing independently of behavioral correlations. The tendency towards the use of neutral classifiers may hint at the possibility of an ongoing language change, indicating that Jordanian Arabic speakers could be gravitating towards adopting neutral alternatives, while gradually moving away from idiomatic classifiers.

6. Conclusions and Recommendations

This study has investigated semantic variation in the use of classifiers in Amman. It argues that the choice between idiomatic and neutral classifiers for fruits and vegetables in Amman may be socially constrained by gender, age, and education. This research significantly advances our understanding of the interaction between cognitive linguistics and sociolinguistics in shaping linguistic preferences. It emphasizes the importance of demographic factors in influencing classifier selection in Jordanian Spoken Arabic. Müller’s (2009) scalar classification framework proves influential for analyzing the metaphorical nature of these classifiers, categorizing them as entrenched metaphors that speakers can activate but are restricted by specific speaker-related factors, namely, gender, age, and education. The preference of idiomatic classifiers, which may come with associations of being less refined, potentially vulgar, or associated with a lower social class, contrasts with the neutral and straightforward neutral options. These preferences are clearly influenced by social factors, and the choice between classifiers can be regarded as a language change in progress. These findings highlight the dynamic and context-dependent nature of language usage, emphasizing the necessity of a comprehensive approach that considers both cognitive and sociolinguistic dimensions in relation to semantic choices. Moreover, this research paves the way for further exploration into the relationship between language, cognition, and sociocultural context, aligning with the principles established by Labov.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by Research and Ethical Committee of The School of Foreign Languages at The University of Jordan (4/2023/2024).

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: We wish to thank the three anonymous reviewers and the special issue editors for their useful comments on an earlier draft of the paper. We would also like to thank the participants for taking part in this study. All remaining errors and inaccuracies are, of course, ours alone. We would also like to acknowledge that we used AI tools to proofread some sentences, as we are not native speakers of English.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A
Consent for Taking Part in the Interviews
You are invited to participate in a research investigation. Before deciding to join this study, it is crucial to comprehend the purpose and nature of the research. The objective of this study is to conduct a cognitive sociolinguistic analysis of Jordanian Spoken Arabic. If you choose to participate, you will be involved in an interview where questions about the expressions you use in specific contexts will be posed. The interview is expected to last around 20 min and will be recorded. Your involvement in this study is entirely voluntary, and you retain the right to refuse to answer any or all questions. You may also opt to end your participation at any point.

All participant data will be handled confidentially. For any inquiries about the study, your rights as a research participant, or any other concerns related to your involvement, please reach out to the primary author at: a.altakhaineh@ju.edu.jo.

Having read and comprehended the provided information and having had the opportunity to seek clarification, you acknowledge that your participation is voluntary, and you may withdraw at any time without providing a reason and without incurring any costs. You understand that you will receive a copy of this consent form. With full awareness, you willingly agree to participate in this study.

Participant’s signature ______________________________ Date __________

Notes
1 Note that JSA contains three main sub-dialects, including Urban JSA, Bedouin JSA, and Rural JSA (see Zibin and Al-Tkhayneh 2019); our investigation revealed that variation in the use of these classifiers can occur in these three sub-dialects.
2 “Usage-based language” is rooted in the assumption that language does not only include a set of form-meaning units to be used in everyday life, but is also itself the product of language use grounded in language users’ experience (Pütz et al. 2014).
3 In the fill-in-the-blanks question, we used the same sentence throughout to neutralize the effect of context on classifier choice.
References


Cheshire, Jenny. 2007. Discourse variation, grammaticalisation and stuff like that. Journal of Sociolinguistics 11: 155–93. [CrossRef]


Poplack, Shana, David Sankoff, and Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. Linguistics 26: 47–104. [CrossRef]


Robinson, Justyna A. 2010. Awesome insights into semantic variation. *Advances in Cognitive Sociolinguistics* 45: 85. [CrossRef]


Tagliamonte, Sali A., and Alex D’Arcy. 2004. He’s like, she’s like: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8: 493–514. [CrossRef]

Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1: 179–95. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.