

Revisiting definiteness: Insights from two types of definite articles in Tihami Arabic

Word count: 11934

Abstract Recent studies have looked into the semantic typology of definiteness, focusing on languages that exhibit two kinds of definites in typologically unrelated languages (e.g., Schwarz 2009; 2013; 2019; Jenks 2015). In these languages, the two definite articles systematically correspond to two prominent theories of definiteness, namely uniqueness (where the definite article is licensed by the uniqueness of the definite description in a particular domain) and familiarity (where the article is licensed by a linguistic antecedent in the text or the conversation). This study presents data from Tihami Arabic which shows two definite articles, *im-* and *al-*, challenging this conventional dichotomy of definiteness. Instead, the two definites of Tihami align with Löbner's (1985; 2011) classification of uniqueness, which distinguishes between inherent and context-derived uniqueness. While *al-* occurs with nouns that are inherently unique, *im-* appears with nouns that derive their uniqueness from the discourse. Our findings reinforce Löbner's approach, and extend our cross-linguistics perspective on languages with two types of definite articles.

Keywords: Two types of definites, definiteness, uniqueness, Tihami Arabic

1 Introduction

Recent studies have looked into the semantic typology of definiteness, focusing on languages that exhibit two kinds of definites in typologically unrelated languages (e.g., Schwarz 2009; 2013; 2019; Jenks 2015; 2018). In these languages, the two definite articles systematically correspond to two prominent theories of definiteness: uniqueness and familiarity. On the one hand, the uniqueness approach is built on the understanding that we utilize a definite description to refer to entities that have a unique role or property within a given contextual domain (e.g., Russell 1905; Strawson 1950). For example, *the table* in (1) refers to a unique table in the office context. On the other hand, the familiarity approach is broadly about the idea that a definite description picks out a referent that is familiar to the discourse participants (e.g., Heim 1982; Kamp 1984). Although there is a discrepancy in the literature on the precise definition of familiarity, the focus is on what Roberts (2003) calls “strong familiarity” in which a definite description is licensed by a linguistic antecedent in the text or the conversation. The definite description *the book* in (2), for instance, refers back to the same book introduced by the indefinite description *a book* in the prior sentence.

- (1) **Context:** *The speaker is standing in an office with exactly one table.*
The table is covered with books. Uniqueness (Schwarz 2009: 79)
- (2) John bought **a book** and a magazine. **The book** was expensive. Familiarity (Schwarz 2013: 537)

The uniqueness-familiarity distinction has gained support from cross-linguistics studies, where it has been argued that languages morphologically distinguish between the two

types of definiteness (Schwarz 2009; 2013; 2019; Arkoh & Matthewson 2013; Ortmann 2014; Jenks 2015; 2018; Ingason 2016; Cisneros 2019; Šereikaitė 2019). In light of these studies, we examine definiteness in Tihami (a dialect of Arabic spoken along the southern Red Sea coast of Saudi Arabia and Yemen by about 5,344,000 speakers, Joshua Project 2023). Tihami exhibits two distinct definite articles in its definiteness paradigm: *im-* and *al-* (3).^{1, 2}

- (3) a. **Im-bayt** lu-hu ʔaʕmidah.
 IM-house for-it pillars
 ‘The house has pillars.’
 b. **As-sama:** ma:l-ha ʔaʕmidah.
 AL-sky NEG-for-it pillars
 ‘The sky has no pillars.’

As we discuss in depth in section 3, the uniqueness-familiarity distinction falls short in accounting for the two types of definites in Tihami, where both *im-* and *al-* can be used in both contexts. As a preview, the example in (4) shows that, in a context that is known to support familiar definiteness, we find that both *al-* and *im-* are possible.

- (4) ʔali dʒa:ba ʕahi: wa gahwah. **Im/Al-gahwah** amar
 Ali bring.PFV.3SG.M tea and coffee. im/al-coffee bitter.COMPARA
 ‘Ali brought tea and coffee. The coffee is more bitter.’

However, the definites *al-* and *im-* are not simply in free variation in Tihami. In Section 5, we argue that the Tihami data are best understood in terms of the approach to definiteness in Löbner (1985; 2011), which is laid out in Section 4. On this view, definiteness is viewed in terms of inherent uniqueness or contextual uniqueness.³ We argue that *im-* attaches to nouns that are inherently non-unique; the uniqueness of such nouns is determined by the context. In contrast, *al-*, with its phonological variants, appears with nouns that are inherently unique, which makes it context-independent. Following Šimík (2018), we utilize situation semantics to account for this distinction, where the referent of *im-* is evaluated solely relative to the topic situation while the referent of *al-* is evaluated relative to all situations that are similar to the topic situation.

Before proceeding, a note is in order about the terminologies used in this study. The two types of definites are often labeled as *weak* (which often corresponds to uniqueness) versus *strong* (which often corresponds to familiarity) in many languages (e.g., Schwarz 2009; 2013; 2019; Jenks 2015). To avoid confusion, we will deviate from the mainstream labeling of the two definites cross-linguistically and we will henceforth use IM-versus AL- labelings for the two definites in Tihami. This is partly because the *weak-strong* labeling involves an analysis that does not quite capture Tihami data, as we shall argue shortly.

¹ The two definites show different phonological variations; while *al-* assimilates with its following coronal sounds (Alahmari 2015), *im-* does not show assimilation, at least in the Tihami variety under discussion. Note that despite the fact that Tihami shares the definite *al-* with other Arabic varieties, the definite *im-* is considered one of the distinctive features of the dialect which sets it apart from other Arabic varieties (Turner 2018; 2021).

² The way Tihami data is collected is by consulting my own judgment (being a native speaker of Tihami Arabic), and the judgments of some family members and friends who are themselves native speakers of Tihami. The available publications of the dialect were also consulted, such as Rabin (1951); Greenman (1978); Prochazka (1988); Alahmari (2015); Alqahtani (2015); Behnstedt & Goldbloom (2016); Watson (2018); and Alfaifi & Davis (2021).

³ In fact, this line of analysis is not completely new and was previously used to capture the two definites in some other languages (cf. Ortmann 2014).

The remainder of this study is structured in the following manner. In Section 2, we lay out the background about the two types of definites cross-linguistically, focusing on the uniqueness-familiarity distinction. Then, Section 3 tests Tihami's two definites in light of the uniqueness-familiarity distinction, showing that the distinction fails to adequately capture the distribution of the two definites. An alternative model, Löbner's (1985; 2011) approach, is introduced in Section 4. Section 5 shows the analysis of the two definites in Tihami. Before concluding in Section 7, Section 6 situates the two definites in Tihami within the broader typology of definiteness.

2 Two definites across languages

There have been many attempts to account for definiteness uses in natural language (see Abbott 2006; Heim 2012 for review), which are mostly based on the English definite article *the*. Languages that morphologically exhibit two definite articles have allowed us to further investigate the nature of definiteness. One line of analysis—which was first proposed by Schwarz (2009)—suggests that languages with two distinct definite markers explicitly exemplify the two prominent theories, namely uniqueness and familiarity (cf. Schwarz 2009; 2013; 2019; Jenks 2015; Cisneros 2019). In this section, we will provide a sketch of the two theories here and a selected definition of each theory, followed by illustrative examples from Standard German.

Starting with the uniqueness theory, the core concept of this theory is that the definite description picks out an individual that uniquely fits the description provided by the definite description. It is based on Russell's (1905) tradition of definiteness which suggests that the definite article has two effects, namely the uniqueness and the existence of the referent. Some examples of uniqueness are something like *the king of Jordan* and *the sun*, where the definite description picks out a unique individual that fits the descriptions of the king and the sun, respectively. The reference quality, however, doesn't need to be universal, as it can be evaluated relative to a contextual or situational domain. For instance, the sentence in (5), repeated from (1) above, is understood to mean that there is only one table in the given context that satisfies the definite description *the NP*. That is, the definite description denotes the uniqueness of the table with respect to the context, even though there is actually more than one table in the world.

- (5) **Context:** *The speaker is standing in an office with exactly one table.*
The table is covered with books. (Schwarz 2009: 79)

The second theory that characterizes definiteness is familiarity. The idea of familiarity is that the definite description picks out a referent that is familiar to the discourse participants (Heim 1982; Kamp 1984; Roberts 2003). Although there is debate about what counts toward familiarity in the literature, Roberts (2003) distinguishes between two types of familiarity: *weak* and *strong* familiarity. The weak familiarity is the role of identifying the referent by different mechanisms, such as discourse salience; for example, a referent can be identified by the discourse participants if that referent is salient enough in the discourse. The other type of familiarity is the strong familiarity, where the definite is used anaphorically, in that it refers to a previous linguistic expression in the utterance or the discourse. For example, the definite description *the book* in (6), repeated from (2) above, anaphorically refers to its indefinite counterpart that was already introduced in the utterance.

- (6) John bought a **book** and a magazine. **The book** was expensive. (Schwarz 2013: 537)

Because the antecedent plays a key role, Schwarz (2009) adopts the anaphoric use of the definite (i.e., the strong familiarity in the sense of Roberts 2003) as a definition for familiarity as, again, demonstrated in (6); the definite description picks out the very same book introduced in the earlier utterance.

Among the languages that embody the uniqueness-familiarity distinction in their definiteness paradigm is Standard German (Schwarz 2009). The language distinguishes two definite markers—“strong” versus “weak” definites—that appear in a very limited context, typically after prepositions. The two definites of German can be distinguished morphologically by whether or not the definite article fuses with a preceding preposition; the strong article doesn’t show fusion with a preceding preposition, as in (7a), compared to the weak definite article that shows fusion, as demonstrated in (7b).

- (7) a. Hans ging **zu dem** Haus.
Hans went to the_{strong} house
‘Hans went to the house.’
b. Hans ging **zum** Haus.
Hans went to.the_{weak} house
‘Hans went to the house.’ Standard German (Schwarz 2009: 14)

According to Schwarz (2009; 2013), each of the two definites represents a theory that characterizes definiteness; the weak definite, on the one hand, embodies the uniqueness theory which claims that the definite description refers to a unique entity in a particular context. This is shown in example (8), where the weak definite refers to an entity that is unique in the global context—namely, the moon.

- (8) Armstrong flog als erster **zum** Mond
Armstrong flew as first.one to.the_{weak} moon
‘Armstrong was the first one to fly to the moon.’ Standard German (Schwarz 2009: 40)

The strong definite, on the other hand, corresponds to the familiarity theory, according to Schwarz’s definition, which claims that the definite description refers back to an entity already established in the context. As exemplified in (9), the definite description *the book* selects the same book introduced earlier in the discourse by the indefinite *a book*.

- (9) In der New Yorker Bibliothek gibt es ein Buch über Topinambur. Neulich
in the New York library exists EXPL a book about topinambur recently
war ich dort und habe #im/ **in dem** Buch nach einer Antwort auf die
was I there and have in.the_{weak}/ in the_{strong} book for an answer to the
Frage gesucht, ob man Topinambur grillen kann.
question searched whether one topinambur grill can
‘In the New York Public Library, there is a book about topinambur. Recently I
was there and searched in the book for an answer to the question of whether one
can grill topinambur.’ Standard German (Schwarz 2009: 30)

As a crucial consequence of this distinction between the strong and the weak definites, the weak definite is odd in anaphoric environments, as shown in (9). Contrastively, the strong definite is infelicitous in uniqueness environments, where the definite description

is mentioned for the first time, as shown in (10), where the mayor is a unique mayor in the utterance context.

- (10) Der Empfang wurde **vom** /#von dem Bürgermeister eröffnet.
 the reception was by.the_{weak}/ by the_{strong} mayor opened
 ‘The reception was opened by the mayor.’ Standard German (Schwarz 2009: 40)

However, Schwarz reports some discourse environments where the preference between the two definites is not consistent among speakers. One of these examples is associative anaphora (or *Bridging* in the sense of Hawkins 1978), where the definite description is licensed by virtue of some link between the definite description and its antecedent. However, there are sub-cases of associative anaphora where preference is consistent, depending on the type of relationship between the definite description and its antecedent. One of these associative anaphora sub-cases is where the definite description is construed as part of its antecedent—a *part-whole* relationship. These contexts show a preference for the weak definite; this type of relationship is often known as the *part-whole* relationship. An example is given in (11), where the crisper is part of the fridge, and the definiteness of *the crisper* is licensed by the part-whole relationship between the crisper and the fridge.

- (11) Der Kühlschrank war so groß, dass der Kürbis problemlos **im**
 the fridge was so big that the pumpkin without.problem in.the_{weak}
 /#in **dem** Gemüsefach untergebracht werden konnte.
 /in the_{strong} crisper stowed be could
 ‘The fridge was so big that the pumpkin could easily be stowed in the crisper.’
 Standard German (Schwarz 2009: 52)

Another sub-case of associative anaphora that shows consistency in its choice of the article is what Schwarz (2009) refers to as the *producer-product* relationship, where the definite description is the producer of its antecedent. This type of associative anaphora demonstrates a clear preference for the strong definite article. It is given in (12), where *the author* is the producer of *the play*, and definiteness emerges by virtue of this kind of relationship between the two elements.

- (12) Das Theaterstück missfiel dem Kritiker so sehr, dass er in seiner Besprechung
 the play displeased the critic so much that he in his review
 kein gutes Haar #**am** /an **dem** Autor ließ.
 no good hair on.the_{weak} /on the_{strong} author left
 ‘The play displeased the critic so much that he tore the author to pieces in his
 review.’ Standard German (Schwarz 2009: 53)

In short, the two definites of Standard German embody the uniqueness-familiarity characterization of definiteness; while the weak article represents uniqueness, the strong article is for familiarity environments, although there are some exceptions, as the case with associative anaphora. Overall, this is just one example of many unrelated languages that embody this distinction, such as Fering, Dutch, Upper Sorbian, Upper Silesian, Thai, Mandarin Chinese, and Korean (Ortmann 2014; Jenks 2015; Schwarz 2019).

3 Tihami's two definites in light of uniqueness-familiarity distinction

Because the uniqueness-familiarity distinction has gained substantial support across various linguistics studies, with each definite article reflecting a theoretical framework, it is natural to ask whether Tihami's two definites fit this pattern as well. We show however that this is not a good fit. The two definites in Tihami deviate from this distinction, as both definite articles can occur in familiarity and uniqueness environments.

In Tihami, both definite articles—*im-* and *al-*—are compatible with nouns in uniqueness contexts. For example, (13a) demonstrates that both articles can refer uniquely to the noun *tʰarwilah* 'table'. Similarly, (13b) shows that the noun *gamar* 'moon' is compatible with both definites and it refers uniquely.

- (13) a. ʕali: yastaxdim **im/atʰ-tʰa:rilah.**
 Ali use.IPFV.3SG.M IM/AL-table
 'Ali uses the table.'
 Context: *There is only one table in the room.*
- b. ʕali: ʃa:fa **im/al-gamar.**
 Ali see.PFV.3SG.M IM/AL-moon
 'Ali saw the moon.'

Both definites are also compatible with (strong) familiarity contexts, where the referent is already established in the discourse. In (14), the definite descriptions *Baloot* and *chess* in (14a) are introduced by the indefinite counterparts in the preceding sentence. This pattern continues in (14b), where both definite articles show up on the noun *ħadi:d* 'iron', which has been previously mentioned in the preceding sentence.

- (14) a. ʕali laʔba balaut wa ʃitʰrandʒ. **Im/Al-balaut asʰʕab min im/aʃ-ʃitrاندʒ.**
 Ali play.PFV.3SG.M Baloot and Chess. IM/AL-Baloot harder than IM/AL-chess
 'Ali played Baloot and chess. The Baloot was harder than the chess.'
- b. ʕali ʃa:la ma:j-in wa ħadi:d-in. **Im/Al-ħadi:d θagi:l(-in).**
 Ali lift.PFV.3SG.M water-NUN CONJ iron-NUN. IM/AL-iron heavy(-NUN).
 'Ali lifted water and iron. The iron was heavy.'

The reader might conclude from these data that *im-* and *al-* are simply in free variation, being always interchangeable. In fact, there are contexts where one or the other is not felicitous. These contexts, however, do not align with the uniqueness-familiarity division. For instance, the noun *mother* always takes the definite *al-*, regardless of the type of context it occurs in (15).

- (15) a. ʕali: sʰa:fah ***im/al-umm** gabla walad-ha
 Ali shake.hand.PFV.3SG.M *IM/AL-mother before son.her
 'Ali shook hands with the mother before her son.'
- b. Umm-in wa bint-ha daxalu su:g-in. ***Im/Al-umm**
 mother-NUN CONJ daughter.her enter.PFV.3PL market-NUN. *IM/AL-mother
 kan-an la:bsah ahmar.
 COP-3SG.FEM wear.IPFV.3SG.FEM red
 'A mother and her daughter entered a market. The mother was wearing red.'

Moreover, although generics are not discussed in languages that correspond to the uniqueness-familiarity distinction, the definite generics in Tihami also present a challenge for the uniqueness-familiarity distinction. Definite generics consistently appear with the definite *al-* and are ruled out with *im-*. This is particularly the case with nouns that occur with generic-denoting predicates like *widespread* and *rare* (16).⁴

- (16) a. ***Im/Al-dʒamal** na:dir fi: ameri:ka.
 *IM/AL-camel rare in America
 ‘The camel is rare in America.’
 b. ***Im/Al-ruzz** mintaʃir fi a:siya.
 *IM/AL-rice widespread in Asia
 ‘Rice is widespread in Asia.’

To recapitulate, the challenge posed by Tihami’s two definites to the uniqueness-familiarity distinction lies in the fact that both definites can occur in familiarity and uniqueness contexts. However, this observation should not lead to the conclusion that these definites are interchangeable; rather, each definite conveys a distinct meaning as will be discussed in 5. In the next section, we will discuss an alternative account of definiteness that leads to a more accurate understanding of the distinction between *al-* and *im-* in Tihami.

4 Löbner’s approach to definiteness

While the uniqueness-familiarity distinction has a lot of crosslinguistic support, an alternative to this model has been proposed in Löbner (1985; 2011), where definiteness is not determined solely by the (linguistic) context, but also by the lexical semantics of the noun. Löbner defines definiteness in terms of two properties: (i) whether the noun is inherently unique (ii) whether the noun is relational. We introduce these ideas below.

Definiteness, in Löbner’s sense, is uniqueness, but his take on uniqueness is different from that of Russell (1905) and Strawson (1950). He perceives uniqueness as either inherently encoded in the lexical semantics of the definite description or derived from the context. This is where he divides definiteness into semantic and pragmatic definites. The core of semantic definiteness is concerned with inherentness, but that isn’t the only way to be lexically inherent; one can get the same effect through adnominal modification. By contrast, pragmatic uniqueness is concerned with nouns that are inherently non-unique and they derive their uniqueness from the context. The noun *sky*, for example, according to Löbner’s approach, is inherently unique because the referent of the sky is the same across multiple contexts. However, a noun like *girl* is a non-unique noun because its referent can change according to the situation in which the definite description is uttered (i.e., the referent of *the girl* is not necessarily the same across multiple contexts).

Löbner (1985; 2011) reaches this conclusion by appealing to two semantic properties of nouns. The first property is, as discussed immediately above, whether the NP is inherently unique or non-unique. Beyond (non-)inherent uniqueness, the second semantic property Löbner bases his approach on is relationality. In his sense, a noun can be either inherently

⁴ We note that nouns in Tihami are not morphologically required to appear with either definite determiner. Indefinite nouns are unmarked for either *al-* or *im-*, as in (i). See AUTHOR (2024) for in-depth discussion of the (in)definiteness paradigm in Tihami.

- (i) ʔali ya:kul tufa:h.
 Ali eat.IPFV.3.SG.M apples
 ‘Ali eats apples.’

relational or inherently nonrelational. This has to do with the question of whether a noun is a one-place or a more-than-one-place predicate term. For example, *woman* and *wife* differ in arguments. While *woman* is not inherently relational as it is a one-place predicate, *wife* is inherently relational since the concept *wife* is related to another noun—hence involves a further argument—in that, a wife cannot be a wife unless she is a wife of someone (e.g., *Bill's wife*).

Based on these two semantic properties—namely uniqueness and relationality, Löbner (2011), as an extension of Löbner (1985), proposes a classification of nouns. He distinguishes four types of concepts: sortal, relational, individual, and functional nouns; again, the four types of nouns are sensitive to uniqueness, shown at the top of Table 1, and relationality as shown in the left-hand side of the table.

	inherently non-unique	inherently unique
Inherently nonrelational	Sortal nouns (SN): <i>Dog, stone, book</i>	Individual nouns (IN): <i>Sun, weather, Mary</i>
Inherently relational	Relational nouns (RN): <i>Sister, finger, uncle</i>	Functional nouns (FN): <i>Father, head, age</i>

Table 1: The four noun types with respect to relationality and uniqueness (Löbner 2011: 307)

Explaining each type of nouns in Table 1, sortal nouns are inherently nonrelational due to the fact that they are one-place predicates. They are also inherently non-unique because nouns like *dog*, *stone*, or *book* can have different referents across different situations.

On the other hand, relational nouns share with sortal nouns the fact that they are inherently non-unique, but they diverge from them in that they are inherently relational by virtue of having an extra argument which is typically occupied by a possessor. The relationship between the relational noun and the noun it relates to is one-to-many in that the possessum might not be unique to the possessor. The relational noun *sister*, for instance, might not be unique for whom she is a sister of as in, say, *John's sister*, where John might have more than one sister.

In contrast, functional nouns are inherently unique and inherently relational. Although functional nouns share with relational nouns the fact that they are both relational, the relationship between the functional noun and the noun it relates to (i.e., its extra argument which is typically a possessor) is one-to-one. The functional noun *father*, for example, is a unique individual for whom he is a father of, as one can have only one father—a biological father, at least. Here the uniqueness relationship between the possessor and its possessum holds even across multiple situations.

Moreover, individual nouns are inherently unique, behaving similarly to functional nouns with respect to uniqueness in that they maintain the definite description referent across multiple situations. However, they diverge from functional nouns in that they are inherently nonrelational since they are typically one-place predicates. Individual nouns include multiple types of nouns; beside proper names and pronouns, Löbner (2011: 281) classifies individual nouns into subtypes: role terms (e.g., *US President*, *pope*), terms for institutions (e.g., *catholic church*), terms for unique objects (e.g., *sun*, *earth*, *moon*) or

singular events (e.g., *World War II*), and abstract terms (e.g., *world*, *date*, *temperature*) (Löbner 2011: 284).⁵

The key criterion of Löbner's approach is whether the referent of the definite description is determined by an inherent or accidental uniqueness. Based on this criterion, Löbner (2011) makes a distinction between semantic and pragmatic uniqueness.

Semantic uniqueness means that the uniqueness of the definite description is context-independent, and it is inherent to the meaning of the noun. This is the case with individual nouns and functional nouns. For example, *the weather*, *the time*, and *the air* are all individual nouns, where the uniqueness is encoded in the lexical semantics of the noun, in that each noun has only one referent. This is also the case with functional nouns, for example, *the Prime Minister* and *age*, where the referent is unique for the nouns it relates to. In these types of nouns, Löbner (1985: 311) argues that the appearance of the definite is semantically redundant because these nouns are semantically unique without the article; this is reflected in some individual nouns such as proper names and pronouns in English which do not take a definite article (e.g., *John*, *Bill*, *she*, *they*), for example; in fact, these cases are sensitive to the scale of uniqueness, introduced by Löbner (2011: 320) and refined by Ortmann (2014: 314), where proper names and pronouns occupy a higher position of semantic uniqueness which results in their lack of a definite article as we shall see in more detail in §6.

While individual nouns and functional nouns have been recognized as sources of inherent uniqueness and collectively labeled as semantic uniqueness, they are not the only contributors, as noted by Löbner (1985; 2011). There are instances where a noun itself does not possess inherent uniqueness; instead, its uniqueness comes from the modifying elements. Löbner (1985: 301) refers to these as *Complex Functional Concepts*. Examples include superlatives and ordinals, which modify nouns that are not inherently unique, such as sortal and relational nouns, as illustrated in (17). In this context, although *wrestler* and *daughter* are inherently non-unique nouns, their uniqueness is derived from the ordinal and superlative modifiers that are considered to be inherently unique.

- (17) a. the {next/ last/ third/ most successful} **wrestler**
 b. the {next/ last/ third/ most successful} **daughter** of Bill

In pragmatic uniqueness, by contrast, uniqueness comes from the linguistic or extra-linguistic context, because the noun itself is not inherently unique, as is the case with sortal nouns and relational nouns. In order for a sortal noun and a relational noun with a definite article to refer uniquely, contextual information is needed. This can be achieved by using various context-dependent environments such as referencing a specific entity within a particular domain or the use of definite descriptions that are licensed anaphorically by their antecedents as the case in familiarity and associative anaphora examples. Under Löbner's approach, uniqueness for non-inherent unique nouns might also be introduced by a uniqueness-establishing relative clause (typically a restrictive relative clause).

⁵ A note is in order about the four classifications of nouns and their relationship to culture and polysemy. Although the classification of the four types of nouns is fair to a large extent, it is however sensitive to cultural differences, particularly with nouns that are deemed to be inherently unique. For example, a word like *God* is an individual noun in cultures that believe in one god. However, it is a sortal noun in cultures that believe in more than one god. A similar consideration applies to a word like *head* which is considered a functional noun, but there exist animals, for example, with more than one head. However, it is better to keep in mind that such classifications are not absolute, and there are cultural differences to some extent. Secondly, as pointed out by Löbner (2011: 282), the noun classification is relativized by polysemy. A noun can represent more than one noun type depending on the different meanings the noun has. For example, *child* has a relational reading 'direct descendant of' and a sortal reading 'nonadult'.

In this example, *the woman* becomes unique by virtue of its following restrictive relative clause *he went out with last night*.⁶

- (18) What is wrong with Bill? Oh, **the woman** he went out with last night was nasty to him. (Hawkins 1978: 131)

To conclude this section, Löbner’s approach to definiteness claims that definiteness is uniqueness, distinguishing between two types of uniqueness (namely, semantic and pragmatic uniqueness). We have seen that semantic uniqueness definites are context-independent, because the nouns they occur with are inherently unique as is the case with individual nouns and functional nouns. On the other hand, pragmatic uniqueness is for nouns that are inherently not unique, and therefore their uniqueness comes from various shapes of the context or the discourse. We show in the following section the Tihami data and argue that it is best captured by appealing to Löbner’s (1985; 2011) approach to definiteness.

5 Tihami data and the analysis

In this section, we present the distribution of the two definites in Tihami Arabic in light of Löbner’s (1985; 2011) approach to definiteness. We show, in Section 5.1, that *al-* correlates with semantic uniqueness as it occurs with nouns that are inherently unique such as individual nouns, functional nouns, and generics. Because the nouns it occurs with are inherently unique, they require no context to satisfy the role of definiteness. On the other hand, *im-* occurs with nouns that are inherently non-unique such as sortal nouns and relational nouns, and consequently they require contexts to draw their uniqueness, as shown in Section 5.2. In section 5.3, we deal with “flexible” nouns. These nouns are inherently unique but can shift to being inherently non-unique in the right context. This section underscores the idea that contextual uniqueness is the key factor in Tihami’s definiteness system. In section 5.4, we offer a formal analysis of definiteness to account for Tihami, proposing that an inherently unique NP is an NP whose referent is the same referent across multiple situations, while an inherently non-unique NP is as an NP where its referent is different across multiple situations, following Šimík (2018).

5.1 *Al-* as a semantic uniqueness definite

The appearance of the definite *al-* is restricted to nouns that are characterized as inherently unique, in Löbner’s sense, such as individual nouns and functional nouns, due to the fact that they have one referent across multiple situations (see §5.1.1). This extends to include definite generic NPs, which refer uniquely to KIND across multiple situations/worlds (see §5.1.2). This leads us to the conclusion that *al-* is a semantic uniqueness definite.

⁶ We have to note that these pragmatic-uniqueness-providing environments are the most common environments, and they are by no means an exclusive list of all the environments.

5.1.1 *Al-* with individual and functional nouns

Starting with nouns that are inherently unique, individual nouns and functional nouns trigger the occurrence of the definite *al-*, but the degree to which these nouns are sensitive to pragmatic environments varies, a matter to be discussed further in §5.3.⁷

Initially, let us discuss how *al-* occurs with individual nouns. As discussed in §4, Löbner (2011: 284) classifies individual nouns into sub-types, namely *role terms*, *institution terms*, *unique objects* or *events*, and *abstract terms*. Interestingly, these nouns trigger the appearance of *al-* in Tihami, as demonstrated in (19). This can be explained by the fact that such nouns are inherently unique, and require no context for identifying their referents.

- (19) a. **Ar/*im-rasu:l** ma:ga:l kiða.
AL/*IM-prophet NEG-say.PFV.3SG.M like.that
'The prophet didn't say that.' (Role term)
- b. ʕali qid-hu fi **al/*im-dʒa:miʕah**.
Ali become-he in AL/*IM-university
'Ali went to college.' (Institution term)
- c. **As/*im-sama:** ma:l-ha ʔaʕmidah.
AL/*IM-sky NEG-for-it pillars
'The sky has no pillars.' (Unique object)
- d. ʕali ya-ðkir **al/*im-ħarb al/*im-ʕa:lamiyah aθ/*im-θaniyah**.
Ali 3SG.M-remember.IPFV.3SG.M AL/*IM-war AL/*IM-world AL/*IM-second
'Ali remembers World War II.' (Unique event)
- e. La: tiga:rin **al/*im-ħubb bi ad/*im-dʕulm**.
NEG compare.IPFV.2SG.M AL/*IM-love to AL/*IM-injustice
'Don't compare love to injustice.' (Abstract terms)

The reason these nouns are ruled out with the definite *im-* is that such terms are context-independent, and the referents of these nouns can be identified across multiple situations. For example, the role term *ar-rasu:l* 'the prophet' can be identified by a Tihami speaker without the need for context support. Similarly, the institution term *al-dʒamiʕah* 'the college' refers to the institutional concept of college, and the referent of this institution is context-independent.⁸ Further, a unique object or event is unique in the universal domain as shown in (19c)-(19d); a term like *as-sama:* 'the sky' refers to one object that cannot be mistaken with other things; similar considerations apply to the unique event *al-ħarb al-ʕa:lamiyah aθ-θaniyah* 'World War II' in (19d). Abstract terms, as in (19e), likewise, are inherently unique since they refer to one concept, and such concept can be uniquely referred to in multiple situations, independent of context.⁹

The other type of noun that triggers the appearance of *al-* are functional nouns. Again, these nouns are characterized by being inherently unique and inherently relational; a functional noun involves an extra argument—typically a possessor—with which the functional noun has a one-to-one relationship. Due to the inherent uniqueness of the posses-

⁷ As we shall see in §5.3, the degree to which the individual noun or the functional noun can be forced by the context varies; some of these nouns are resistant to *im-* as the case with *as/*im-sama:* 'the sky' while some others, for example, *af/im-fams* 'the sun' are more flexible, depending on the context, and therefore yield a meaning difference.

⁸ Note, however, that a term like *college* is polysemous, as it might be used as a sortal noun, and its referent needs a context to refer uniquely.

⁹ The apparent exception for individual nouns is proper names and pronouns which are considered types of individual nouns, but they are typically not marked with any definite marker. The fact that they are not marked with any definite article is attributed to their position within the uniqueness scale, initially proposed by Löbner (2011: 320). We will come back to this point as we discuss the uniqueness scale in §6.

sum to the possessee, functional nouns require *al-* as a definite marker. Take (20) as an example, where *the king* is in a unique possessive relationship to the speaker and the group for whom the king is a king of, which results in using the article *al-*, and rules out the use of *im-* in this context.¹⁰

- (20) **Al/*im-malik** bizur-na bukraḥ.
 AL/*IM-king visit.FUT-1PL tomorrow
 ‘The king will visit us tomorrow.’

In summary, the uniqueness of individual nouns and functional nouns is inherent to the lexicon, meaning they do not require context to establish their uniqueness. However, this is not the only source of inherent uniqueness; definite generic nouns also contribute to this concept, which we will discuss next.

5.1.2 *Al-* with generics

Although Löbner (2011: 279-80) abstracts away from discussing generics, inherent uniqueness has a clear extension to genericity. While individual nouns and functional nouns typically refer to unique entities or abstract concepts, generics refer to abstract objects called KINDS. This is evident in Tihami where *al-*, with its phonological variations, provides a generic interpretation of singular, plural, countable, and uncountable nouns, as shown in (21a)-(21d). Examples (21a)-(21d) show that *al-*'s semantics licenses its occurrence with nominals that are selected by kind-denoting predicates such as *rare*, *a type of*, and *widespread* (see Carlson 2019: 254-255 about generics).

- (21) a. **Al/*im-kalb** min anwa:f aθ-θadiyat.
 AL/*IM-dog.SG.GEN from types AL-mammal.PL
 ‘The dog is a type of mammal.’
 b. **Al/*im-kila:b** min anwa:f aθ-θadiyat.
 AL/*IM-dog.PL from types AL-mammal.PL
 ‘Dogs are a type of mammal.’
 c. **Ar/*im-ruz** mintafjir fi a:sya.
 AL/*IM-rice.PL widespread in Asia
 ‘Rice is widespread in Asia.’

¹⁰ The idea of inherent uniqueness becomes more complex when discussing Complex Functional Concepts. According to Löbner (1985; 2011), inherent uniqueness extends beyond nouns and can be encoded in modifying elements like superlatives and demonstratives. However, this doesn't quite work in Tihami; Complex Functional Concepts in Tihami are sensitive to syntactic factors that determine the type of definiteness marker realized on the noun. For instance, demonstratives in Tihami can flexibly occur before or after their head noun. When a demonstrative precedes its noun, the noun must include the definite marker *al-*, irrespective of whether the noun is inherently unique, as seen in example (1a). However, when the demonstrative follows the noun, the noun must take the definite marker *im-* instead, as illustrated in examples (1b). We will not discuss the interaction between the two definite markers and syntax in this work, leaving this for future study.

- (i) a. ḏa: **al/*im-walad** mumta:z.
 This AL/*IM-boy amazing
 ‘This boy is amazing.’
 b. **Im/*al-walad** ḏa:h mumta:z.
 IM/*AL-boy this amazing
 ‘This boy is amazing.’

- d. **Al/*im-miyah** na:dirah fi al-marri:x.
 AL/*IM-water.PL rare in AL-Mars
 ‘Water is rare on Mars.’

Assuming a generalizable process for treating properties as kinds (e.g., Chierchia’s 1998 down ‘ \cap ’ operator), Löbner’s inherent uniqueness extends easily to kinds. While an individual referent for *the dog* might vary across contexts, the taxonomic reference (as in, e.g., Dayal 2004) does not. This is why *al-* is required in Tihami.

5.2 *Im-* as a pragmatic uniqueness definite

Having shown the distribution of *al-*, the definite *im-*, by contrast, occurs with nouns that are inherently non-unique, such as sortal nouns and relational nouns. In this case, the definite article is required when the uniqueness of such nouns is determined solely by the context.

To remind the reader, nouns like *t’awilah* ‘table’, *bayt* ‘house’, *jarif* ‘street’, and *kitab* ‘book’ are all sortal nouns that are characterized as being inherently non-unique. The other type of nouns that are inherently non-unique are relational nouns like *kafarah* ‘tire’, *xal* ‘uncle’, *yadd* ‘hand’, and *hiḍyan* ‘shoes’; these nouns are different from sortal nouns in that they are in relation to an additional argument. Due to their inherent non-uniqueness, they occur with the definite marker *im-* whose role is to derive uniqueness from the context. In what follows, we list the most reliable contexts that provide uniqueness. One of the context forms is immediate situation uniqueness where the definite description refers to a particular entity in a particular domain immediately related to the discourse participants, as demonstrated in (22). In these examples, the definiteness of the relational noun *sa:hib* ‘friend’ and the sortal noun *mikayyif* ‘air conditioner’ are derived from the discourse, where the air conditioner and the friend are not ambiguous for the discourse participants, and as a result, they appear with the definite *im-*.

- (22) a. ṡali aflaha maʔa **im/*al-sa:hib**.
 Ali go.PFV.3SG.M with IM/*AL-friend
 ‘Ali went with his/the friend.’
 b. ṡayyil **im/*al-mikayyif!**
 turn.on.IPFV.2SG.M IM/*AL-air.conditioner
 ‘Turn on the air conditioner!’

Pragmatic uniqueness also arises in anaphoric contexts, where the definiteness of non-unique nouns is licensed by an indefinite already established in the context. This is the case with the sortal noun in (23a), where the uniqueness *im-bint* ‘the girl’ in the second sentence is licensed by the indefinite *bint-in* ‘girl’ in the preceding sentences. A similar consideration applies to the relational noun *ixt* ‘sister’ in (23b).

- (23) a. Walad-in wa bint-in daxalu im-qaṡah. ṡali ṡarafa
 boy-NUN and girl-NUN enter.PFV.3PL IM-hall. Ali recognize.PFV.3SG.M
im/*al-bint.
 IM/*AL-girl
 ‘A boy and a girl entered the hall. Ali recognized the girl.’
 b. ṡift ixtain. **im/*al-ixt** im-sayi:rah la:bs-ah aḥmar.
 see.PFV.1SG.M sister.DUAL. IM/*AL-sister small wear.IPFV.3SG-F red
 ‘I have seen two sisters. The younger sister was wearing red.’

The examples presented in (23) necessitate an explanation for similar cases examined previously, specifically those in (4) and (14). In (23), only *im-* is permitted, whereas both examples (4) and (14) allow both definite articles, despite the fact that all three examples involve anaphoricity. The difference is that the nouns featured in (4) and (14)—such as *gahwah* in (4) and *balaut* ‘Baloot,’ *fitrandʒ* ‘chess,’ and *ħadīd* ‘iron’ in (14)—can be understood either anaphorically or generically within anaphoric contexts, thus permitting the occurrence of both definites. In contrast, the nouns *bint* ‘girl’ and *ixt* ‘sister’ in (23) can only be interpreted anaphorically in the given context.

While these examples could all broadly be subsumed under a general version of “familiarity,” there are places where this explanation does not easily extend. For instance, uniqueness might also be coming from a relative clause that makes the definite unambiguous. Uniqueness establishing relative clauses challenge the familiarity theory, in that the non-ambiguity of the noun is not already mentioned in the discourse before the noun gets mentioned, but instead it is established by the relative clause after the noun gets mentioned. A uniqueness-establishing relative clause is given in (24). Again, a noun like *walad* ‘boy’ is a sortal noun that gets its uniqueness from the restrictive relative clause in (24a); this is also the case with the relational noun *ixt* ‘sister’ in (24b).

- (24) a. dʒa: **im/*al-walad** illi gil-t lik ʕala:-hu.
 come.PFV.3SG.M IM/*AL-boy COMP say.PFV.1SG to.you about-him
 ‘The boy that I told you about came.’
- b. **Im/*al-ixt** illi la:bsah aħmar aflahan.
 IM/*AL-sister COMP wear-IPFV.3SG.F red go.out.PFV.3SG.F
 ‘The sister that was wearing red went out.’

One might wonder if non-restrictive relative clauses trigger the occurrence of *im-*, similar to restrictive relative clauses. The simple answer is no; non-restrictive relative clauses typically do not require uniqueness and therefore do not trigger the occurrence of *im-*. See (25).

- (25) Darasna aʃ/***im-ʃams** illi hi akbar al-kawa:kib.
 study.PFV.1PL AL/*IM-sun COMP 3SG.F big.SUPERLA AL-stars
 ‘We studied the sun, which is the biggest of the stars’

The last environment that triggers the occurrence of *im-* is associative anaphora, where the uniqueness is established by an entity already mentioned in the discourse. The relationship between the noun and its definiteness licencer is not co-referential, and uniqueness is derived by the virtue of a link between the two elements. As given in (26a), *im-gaza:z* ‘the windshield’ is part of the car mentioned earlier in the utterance; Schwarz (2009) refers to this type of relationship as a “part-whole” relationship. Another type of associative anaphora is the “producer-product” associative anaphora, as demonstrated in (26b), where the noun is the producer of its licencer. For Schwarz (2009), the part-whole associative anaphora is different from the producer-product associative anaphora in that the noun in the former is a present discourse participant, while it is not in the producer-product relationship. In Tihami, both cases require *im-*.

- (26) a. ʕali iʃtara sayyarah. **Im/*al-gaza:z** kana maksu:r-in.
 Ali buy.PFV.3SG.M car IM/*AL-windshield COP.3SG.M broken-NUN
 ‘Ali bought a car. The windshield was broken.’ (Associative anaphora:
 part-whole relationship)

- b. Garait kita:b-in. **Im/*al-mu?alif** faransi.
 read.PFV.1SG book-NUN. IM/*AL-author French
 ‘I have read a book. The author is French.’ (Associative anaphora:
 producer-product relationship)

To recapitulate, the definite *im-* occurs with nouns that are inherently non-unique. When the context makes the referents of these nouns unique, then *im-* is required. This contextual (i.e., pragmatic) uniqueness comes in many forms, such as immediate situation contexts, anaphoric contexts, uniqueness-establishing relative clause contexts, and associative anaphora contexts.

5.3 Shifting

We have seen so far that *al-* appears with inherently unique entities and *im-* occurs with inherently non-unique entities; in the latter, the noun gets its uniqueness from the broader context. However, we cannot conclude that inherently unique nouns cannot occur with *im-* at all, or vice versa—i.e., that non-unique nouns cannot occur with *al-*. Thus, a crucial component of Löbner’s view is that uniqueness can be *lexical*. In Tihami, this does not quite work. In fact, it is possible for some inherently unique nouns to appear with *im-* if they are forced by the context—i.e., occurring in some of the definite *im-* environments discussed in §5.2. For instance, consider the noun *malik* ‘king’ which is inherently unique. We predict that it appears with *al-*, and it does as shown in (27a). However, if the sentence is uttered in a particular linguistic context, for instance with a uniqueness-establishing relative clause, as shown in (27b), then *im-* is the preferred definite, referring to a particular king in the time and place situated in the context of the utterance.

- (27) a. **al-malik** as^ydara amir
 AL-king announce.PFV.3SG.M decree
 ‘The king announced a decree.’
 b. jiftah **im-malik** illi kana fi im-gaṣah?
 see.PFV.2SG IM-king COMP COP.3SG.M in IM-hall
 ‘Have you seen the king that was in the hall?’

In fact, this meaning difference is also observed in sub-varieties of Tihami such as Faifi. Examples in (28), are adopted from Alfaifi & Davis (2021: 43), in which the interpretation of the individual noun differs depending on the context; so the referent of *guran* ‘Quran’ in (28a) is the individual noun *Quran*, while in (28b) refers to a particular copy of the Quran, which the speaker specified being on the table.

- (28) a. Garayt **il-gur?a:n**.
 read.PFV.1SG AL-Quran
 ‘I read the Quran.’
 b. Garayt **m-gur?a:n** ḏi: ṣala m-t^ya:wilah.
 read.PFV.1SG IM-Quran that on IM-table
 ‘I read the Quran that is on the table.’ (Alfaifi & Davis 2021: 43)

However, this shifting phenomenon is not absolute as there are some inherently unique nouns that resist shifting, appearing only with *al-*. For instance, a noun like *umm* ‘mother’ refuses to take *im-* even if forced by the context. Consider the pair of examples given in (15) above, repeated in (29) below.

- (31) a. **Garáž** zářila novotou.
garage shined novelty.INSTR
'The garage shined with novelty.'
Context: 'I approached a friend's house.'
- b. **Ta** **garáž** zářila novotou.
DEM garage shined novelty.INSTR
'The garage shined with novelty.'
Context: 'A friend showed me his new garage.' Czech (Šimík 2018: 366)

He formalizes the distinction between inherent and non-inherent uniqueness using situation semantics. Inherent uniqueness, represented by bare NPs, quantifies over all situations that are similar to the topic situation, and across these situations, there is only one entity that satisfies the uniqueness of NP. On the other hand, non-inherent uniqueness, represented by the demonstrative NP, is also captured by quantifying over all situations, but, uniqueness is only restricted to the topic situation, and it is not the case that in all situations that are like the topic situation such that there is only one entity with property *P* that satisfies the uniqueness of NP.

A crucial element of Šimík (2018) analysis is a possibilistic situation semantics (Kratzer 1989; 2021). Situations, in Kratzer's sense, are parts of the possible world, and these parts stand in relationship with each other. They can be used to restrict the speaker's assertion about the utterance, which can come in two ways: the topic situation and the resource situation. While the topic situation restricts the whole utterance, the resource situation restricts a determiner phrase or a quantifier. The resource situation mostly corresponds to the topic situation, but this is not always the case.

The second crucial element of Šimík's analysis is quantification over situations. Šimík utilizes the idea that we can quantify over situations that are "like" (i.e. \approx) some relevant evaluated situation, typically the topic situation.¹² The purpose of "likeness" is to restrict the quantification of inherent/non-inherent uniqueness to contextually relevant situations, excluding abnormal situations. It is particularly important with respect to evaluating inherent uniqueness entities. It eliminates situations where the inherent uniqueness possibly fails. For example, the inherent uniqueness of the *sky*, for instance, is evaluated relative to the actual situation and situations like the actual situation, eliminating possible situations/worlds where there might be more than one sky, a science fiction situation/world, where there might be no sky at all, or a future situation where the sky might not exist.

We will implement the same mechanisms of Šimík's analysis to account for the distinction between the two definites of Tihami Arabic. The meanings of the two definite markers are characterized as follows: while the uniqueness of *al-* is evaluated with respect to all situations that are like the topic situation, the uniqueness of *im-* is evaluated with respect to the topic situation only. That is, the definite *al-* indicates that in all situations like the topic situation there is exactly one entity with the property *P* in those situations, while *im-* indicates that there is exactly one entity in the topic situation and it is not the case that in all situations that are like the topic situation there is exactly one entity with property *P* in those situations. An informal representation of the meanings of

¹² The idea of "likeness" is modeled using Kratzer's (1981; 1989; 2012) modal semantics, in which a modal like *must* relies on two conversational backgrounds: the modal base and ordering source. The modal base picks out a relevant situation among the possible situations, and the ordering source ranks the situations, allowing some similar situations to be more important than the others.

the two definite markers of Tihami is presented in (32) and (33), respectively, modeled after Šimík's analysis for Czech.

- (32) **The inherently unique definite marker *al-*:**
 For any property P , entity x , and topic situation s_0 , such that $P(s_0)(x) = 1$, x is inherently uniquely identifiable in s_0 iff
 $\forall s[s \approx s_0 \rightarrow \exists!y[P(s)(y)]]$
 All situations that are like s_0 are such that there is exactly one entity with property P in those situations.
- (33) **The inherently non-unique definite marker *im-*:**
 For any property P , entity x , and topic situation s_0 , such that $P(s_0)(x) = 1$, x is non-uniquely identifiable in s_0 iff
 $\exists!z[P(s_0)(z)] \wedge \neg \forall s[s \approx s_0 \rightarrow \exists!y[P(s)(y)]]$
 Exactly one entity is P in s_0 and it is not the case that all situations that are like s_0 are such that there is exactly one entity with property P in those situations.

Let us take examples to see how this captures the two definite markers in Tihami. The sky example in (34), repeated from the above, indicates that the definite description *the sky* has one referent in all situations that are like the topic situation s_T . Therefore, the inherent uniqueness of the sky means that in all situations that are like the actual situation, there is exactly one sky in those situations.

- (34) **As-sama:** ma:l-ha ?aʕmidah.
 AL-sky NEG-for-it pillars
 'The sky has no pillars.'

Similarly, generic NPs are inherently unique and are often evaluated with respect to relatively large topic situations, or possibly the whole world (the maximal situation). Consider the example in (35) again. The *dog* is a dog kind, a taxonomic reference, that in all worlds that are similar to the actual world, is such that there is exactly one dog-kind in them.

- (35) **Al-kalb** min anwa:f aθ-θadiyat.
 AL-dog.SG.GEN from types AL-mammal.PL
 'The dog is a type of mammal.'

By contrast, the referent of the non-unique noun like *bayt* 'house' in (36) is restricted to the topic situation. That is, a house with property x in s_T , such that it is not the case that in all situations that are like s_T there is exactly one house with property x .

- (36) **Im-bayt** lu-hu ?aʕmidah.
 IM-house for-it pillars
 'The house has pillars.'

In a nutshell, the uniqueness of the two definite markers is evaluated with respect to different sets of situations; while *im-* restricts the evaluation of its uniqueness to the topic situation, irrespective of the other similar situations, *al-* evaluates its uniqueness with respect to all other situations that are like the topic situation. This indicates that the two definite markers pick out different situations for their evaluation. Supporting evidence comes from places where the choice of the definite affects the choice of situation. Consider (37) and (38) with respect to the given context. Again, a noun like *king* is a functional noun which is inherently unique due to the fact that a king is typically a unique

king for whom he is a king of, and therefore it is typically marked with *al-*. However, it is one of the nouns that can shift if forced by the context, pointing out a non-unique king. In (37) and (38), the choice of the definite marker picks out a referent based on their restricted situation.

Context:

Ali and Saad, who are speakers of Tihami, are attending an environmental summit with various global leaders present, including two kings. One king reigns over Ali and Saad's homeland, while the other is the king of an African nation.

- (37) **Al-malik** alqa: kalmah qawiyyah.
 AL-king give.PFV.3SG.M word strong
 ✓‘The king (for Ali and Saad) gave a strong speech.’
 ×‘The (African) King gave a strong speech.’
- (38) **Im-malik** alqa: kalmah qawiyyah.
 IM-king give.PFV.3SG.M word strong
 ×‘The king (for Ali and Saad) gave a strong speech.’
 ✓‘The (African) King gave a strong speech.’

In (37), the definite *al-* picks out a referent that is unique for the discourse participants in all situations that are like the topic situation (i.e., the environmental summit). The referent is *Ali's and Saad's king*, ruling out the possibility *the African king*. In this case, the resource situation deviates from the topic situation, and chooses situations where the referent is unique. Here, *the African king* cannot afford to be the unique king for the discourse participants in all situations like the topic situation. By contrast, the referent of *im-malik* in (38) aligns with the topic situation and chooses the king who happens to be only unique in the topic situation and not the only king in other situations like the topic situation. The referent in this case is *the African king*, ruling the *Ali's and Saad's king* as a referent. Briefly, the choice of the definite picks out a situation in which uniqueness holds. we will now explore how the two definite markers in Tihami fit within the broader typology of definiteness.

6 Uniqueness Scale

Based on their type of uniqueness, the two definites of Tihami can be mapped into the scale of uniqueness that was first introduced by Löbner (2011: 320) and later revisited by Ortmann (2014: 314), who provides cross-linguistic evidence for the structure of the scale. The scale of uniqueness is shown in (39).

- (39) **Scale of uniqueness**
 deictic SN < anaphoric SN < SN with establishing relative clause < producer-product associative anaphora < part-whole associative anaphora < complex functional concepts < IN/FN < proper names < personal pronouns

(Ortmann 2014: 314)

The scale attempts to organize the occurrences of items that are inherently unique and items that are inherently non-unique. The top of the scale is concerned with items that

are inherently non-unique such as sortal nouns and relational nouns, and because they are inherently non-unique, they derive their uniqueness from the context through various techniques such as deictic use of the definite, anaphoric use of the definite, using a uniqueness establishing relative clause, or through associative anaphora use of the definite. The bottom of the scale is for items that are inherently unique such as complex functional concepts, functional nouns, and individual nouns, including proper names and pronouns. The more we go toward the bottom of the scale the less we need a definite article; this is due to the fact that the definite article on inherent uniqueness items is semantically redundant, as is the case with proper names and pronouns that typically appear with no definite article in many languages (Löbner 2011: 285).

Various languages with one or two definite systems correspond to the scale of uniqueness. However, the cut-off point between inherent uniqueness and non-inherent uniqueness varies from one language to the other. For the sake of demonstrating the point and for the sake of demonstrating where Tihami Arabic stands in the typology of definiteness, let us compare three languages with one and two definite article systems, namely English, Fering, and Tihami Arabic.

Firstly, English is a one-definite-system language, with the definite article *the* occurring with items that are inherently non-unique as in (40a)-(40d), or with items that are inherently unique as in (40f)-(40h) except with proper names and pronouns (see (40h) and (40i), respectively).

- (40) a. **The book** is red. (pointing out to the book)
 b. I bought a book and a magazine. **The book** was expensive.
 c. **The book** that I told you about was expensive.
 d. John read a book. **The author** is French.
 e. John bought a car. **The windshield** was broken.
 f. I met **the most successful tennis player of the year**.
 g. **The king** of Morocco will visit us tomorrow.
 h. *I will meet **the John** tomorrow.
 i. ***The she** will meet us tomorrow.

Another language whose definite system corresponds to the scale of uniqueness is Fering, a North Frisian dialect (Ebert 1971). The language has two definite markers *a* and *di* that correspond to inherent uniqueness and inherent non-uniqueness based on their distribution (Ebert 1971; Ortmann 2014). Particularly, *di* typically occurs with items that are inherently non-unique such as *hingst* ‘horse’ as given in (41). With items that are inherently unique, however, the definite marker *a* is used as the case with *san* ‘sun’ in (42).

- (41) Oki hee an hingst keeft. *A /Di **hingst** haaltet.
 Oki has a horse bought the_{weak} /the_{strong} horse limps
 ‘Oki has bought a horse. The horse limps.’ Fering (Ebert 1971: 160)

- (42) A **san** skiinjt.
 the_{weak} sun shine
 ‘The sun is shining.’ Fering (Ebert 1971: 161)

The cut-off point between the semantic uniqueness side and pragmatic uniqueness side of the scale is associative anaphora and complex functional concepts. Ebert (1971: 118) and Schwarz (2009: 62) report that associative anaphora, be it producer-product associative anaphora or part-whole associative anaphora, is compatible with both definite markers. Although the example in (43) shows that *a* occurs with part-whole associative

anaphora, and example (44) demonstrates that producer-product anaphora occurs with *di*, this does not rule out the possibility of the other way around.

- (43) Wi foon a sark uun a maden faan't taarep. A **törem** stän wat
 we found the church in the middle of the village the_{weak} tower stood a little
 skiaf.
 crooked
 'We found the church in the middle of the village. The tower was a little crooked.'
 Fering (Ebert 1971: 118)
- (44) Peetji hee unn Hamboreg an bilj keeft. Di **mooler** hee ham an
 Peter has in Hamburg a painting bought. The_{strong} painter has him a
 guden pris maaget.
 good price made
 'Peter bought a painting in Hamburg. The painter made him a good deal.' Fering
 (Schwarz 2009: 62)

Similarly, Ebert (1971: 163) reports that both articles are compatible with complex functional concepts such as superlatives as in *a/die huuchst bereg van Feer* 'the highest mountain of Föhr'. This double occurrence of the definite article can only be explained if we consider the scale of uniqueness, where the division between semantic and pragmatic uniqueness becomes unclear as we go toward the middle of the scale.

Proper names typically do not occur with any definite marker, except a few country names that occur with the definite *a*, such as Turkey (Ebert 1971: 71), as demonstrated in (45).

- (45) A **Türkäi**.
 the_{weak} Turkey
 'The (country of) Turkey.' Fering (Ebert 1971: 71)

Finally, Tihami can be mapped to the scale of uniqueness as well. We have seen that *im-*, the pragmatic uniqueness definite marker, occurs with entities that are inherently non-unique such as sortal nouns and relational nouns, as discussed in §5.2. The definite *al-*, the semantic uniqueness definite marker, on the other hand, occurs with entities that are inherently unique such as individual and functional nouns, as well as generics, as discussed in §5.1.

The scale of uniqueness, presented in (39), enables us to compare the distribution of definite articles in different languages, with one article system such as English or two-definite system such as Fering and Tihami Arabic. As noted by Ortmann (2014), the cut-off point between semantic uniqueness and pragmatic uniqueness becomes unclear toward the middle of the scale, as demonstrated in Table 2. Table 2 shows that while the English definite uses *the* in all definite environments, except with proper names and personal pronouns, Fering and Tihami Arabic have different distributions. For Fering, the middle point of the scale, namely with associative anaphora and complex functional concept, is more flexible with both types of definites; the top side of the scale shows a clear preference for pragmatic uniqueness article, while the lower side of the scale prefers the semantic definite article. Tihami fits nicely into the scale of uniqueness, showing that *im-* occurs on nouns that are not inherently unique and they derive their uniqueness from the context; and, *al-* shows up on items that are inherently unique, and deletes on items that are more semantically unique such as on personal pronouns and some proper names; the dashed line in Tihami's column signifies that the definite marker *al-* is sometimes used

with proper nouns, including certain country names such as *al-maksi:k* ‘Mexico’. Similar to Fering, the cut-off point between semantic and pragmatic uniqueness in Tihami occurs with complex functional concepts which possibly occur with both definite markers (see footnote 10).

	ENGLISH	FERING	TIHAMI ARABIC
deictic SN	<i>the</i>	<i>det/di</i>	<i>im-</i>
anaphoric SN			
SN with a uniqueness establishing relative clause		<i>a/at</i>	<i>al-</i>
product-producer associative anaphora			
part-whole associate anaphora			
complex-functional concepts			
IN/FN			
proper names	-	-	-
personal pronouns	-	-	-

Table 2: The distribution of definite articles in English, Fering, and Tihami Arabic, with respect to the uniqueness scale

Overall, Tihami Arabic behaves similarly to other languages in its sensitivity to the scale of uniqueness. Although we don’t have a solid piece of evidence for historical change, the dashed lines in Fering and Tihami Arabic, in Table 2, might suggest a transitional period for the two definites in both languages.

7 Conclusion

This study has examined the distribution of the two definite articles in Tihami Arabic—*im-* and *al-*—and their alignment with the typology of definiteness. By appealing to Löbner’s (1985; 2011) framework, which distinguishes between semantic and pragmatic uniqueness (or inherent and non-inherent uniqueness), We have demonstrated that these two definites exemplify this dichotomy effectively. Specifically, *al-* generally occurs with entities that are inherently unique, while *im-* is used with non-unique entities. Using situation semantics, we argue that the uniqueness of *im-* is evaluated relative to the topic situation, while *al-* is evaluated relative to all situations that are similar to the topic situation, following Šimík (2018). This distinction aligns well with the uniqueness scale proposed by Löbner (2011) and revisited by Ortmann (2014), confirming that Tihami Arabic patterns similarly to other languages in this regard.

In contrast, Schwarz’s (2009; 2013; 2019) uniqueness-familiarity analysis of definiteness does not adequately account for the distribution of Tihami’s articles. Particularly, Schwarz’s analysis does not consider the lexical semantics of the definite entities, which might cast doubt on its generalizability to all languages that exhibit two definite markers. The findings show the importance of considering the lexical semantics of definite entities when analyzing definiteness, paving the way for future research that includes a wider array of languages and linguistic contexts.

Abbreviations

1, 2, 3 = 1st/2nd/3rd person, AL- = The definite marker *al-*, COP = Copula, COMP = Complementizer, COMPARA = Comparative adjective, CONJ = Conjunction, D =

Determiner, DEM = Demonstrative, DEF = Definite, DP = Determiner phrase, DUAL = Dual, EXPL = Expletive, FEM = Feminine, FOC = Focus, FOCF = Focus phrase, FUT = Future tense, GEN = Generic, IM- = The definite marker *im-*, IN = Individual noun, INSTR = Instrumental, IPFV = Imperfective, M = Masculine, NEG = Negation, N = Noun, NP = Noun phrase, NUM = Numeral, NUN = Nunation, O = Object, PFV = Perfective, PL = Plural, PROX = Proximal demonstrative, RN = Relational noun, SG = Singular, SN = Sortal noun, STRONG = Strong definite, SUPERLA = Superlative adjective, TOPP = Topic phrase, WEAK = Weak definite

References

- Abbott, Barbara. 2006. Definiteness and indefiniteness. In *The handbook of pragmatics*, 122–149. Blackwell Publishing. <https://doi.org/10.1002/9780470756959.ch6>
- Aboh, Enoch O. 2004. Topic and focus within D. *Linguistics in the Netherlands* 21(1). 1–12. <https://doi.org/10.1075/avt.21.04abo>
- Ahn, Dorothy. 2017. Definite and demonstrative descriptions: A micro-typology. In Erlewine, Michael Yoshitaka (ed.), *Proceedings of generative linguistics in the old world in asia 11*, vol. 1 (MIT Working Papers in Linguistics). 33–48. MIT. <http://mitwpl.mit.edu/catalog/mwpl84/>.
- Alahmari, Musa. 2015. *The development of m- definite article in Arabic: A historical linguistic perspective*: Indiana University Bloomington Master's thesis. Unpublished.
- Alfaifi, Abdullah & Davis, Stuart. 2021. An examination of the /m-/ definite article in upper Faifi Arabic. *Zeitschrift für arabische Linguistik* 2021(73). 36–52. <https://doi.org/10.13173/zeitarabling.73.0036>
- Alqahtani, Khairiah. 2015. *A sociolinguistic study of the Tihami Qahtani dialect in Asir, Southern Arabia*: University of Essex dissertation. <https://repository.essex.ac.uk/15298/>.
- Arkoh, Ruby & Matthewson, Lisa. 2013. A familiar definite article in Akan. *Lingua* 123. 1–30. <https://doi.org/10.1016/j.lingua.2012.09.012>. <https://www.sciencedirect.com/science/article/pii/S0024384112001969>
- Behnstedt, Peter & Goldbloom, Gwendolin. 2016. *Dialect atlas of north Yemen and adjacent areas*. Leiden, The Netherlands: Brill. <https://doi.org/https://doi.org/10.1163/9789004326422>. <https://brill.com/view/title/32008>
- Carlson, Gregory. 2019. *Genericity* 232–273. Berlin, Boston: De Gruyter Mouton. <https://doi.org/doi:10.1515/9783110589443-008>
- Chierchia, Gennaro. 1998. References to Kinds across Languages. *Natural Language Semantics* 6. 339–405.
- Cisneros, Carlos. 2019. Definiteness in Cuevas Mixtec. In Aguilar-Guevara, Ana & Loyo, Julia Pozas & Maldonado, Violeta Vázquez-Rojas (eds.), *Definiteness across languages*, 39–81. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.3252014>
- Dayal, Vaneeta. 2004. Number marking and indefiniteness in kind terms. *Linguistics and Philosophy* 27. 393–450. <https://doi.org/10.1023/B:LING.0000024420.80324.67>
- Ebert, Karen Heide. 1971. *Referenz, sprechsituation und die bestimmten artikel in einem nordfriesischen dialekt*, vol. 4. Nordfriisk Instituut. <https://api.semanticscholar.org/CorpusID:169906833>.
- Giusti, Giuliana. 2005. At the left periphery of the Romanian noun phrase. *On Space and Time in Language* 23–49.
- Greenman, J. 1978. A sketch of the Arabic dialect of the Central Yamani Tihāmah. *Al-'Arabiyya* 11(1/2). 20–34. <http://www.jstor.org/stable/43192496>.

- Hawkins, J. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. Routledge 1st edn. <https://doi.org/10.4324/9781315687919>
- Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13(2). 137–177. <https://doi.org/10.1515/9783110867602.1>
- Heim, Irene. 2012. 41. *definiteness and indefiniteness* 996–1025. Berlin, Boston: De Gruyter Mouton. <https://doi.org/doi:10.1515/9783110255072.996>. <https://doi.org/10.1515/9783110255072.996>
- Heim, Irene Roswitha. 1982. *The semantics of definite and indefinite noun phrases*: University of Massachusetts Amherst dissertation. <https://scholarworks.umass.edu/dissertations/AAI8229562>. Doctoral Dissertations Available from Proquest. AAI8229562.
- Ingason, Anton Karl. 2016. *Realizing morphemes in the icelandic noun phrase*: University of Pennsylvania dissertation.
- Irani, Ava. 2019. On (in)definite expressions in American Sign Language. In Aguilar-Guevara, Ana & Loyo, Julia Pozas & Maldonado, Violeta Vázquez-Rojas (eds.), *Definiteness across languages*, 113–151. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.3252018>
- Jenks, Peter. 2015. Two kinds of definites in numeral classifier languages. *Semantics and Linguistic Theory* 25. 103. <https://doi.org/10.3765/salt.v25i0.3057>
- Jenks, Peter. 2018. Articulated Definiteness without Articles. *Linguistic Inquiry* 49(3). 501–536. https://doi.org/10.1162/ling_a_00280. https://doi.org/10.1162/ling_a_00280
- Joshua Project. 2023. Tihami Arab in Saudi Arabia. https://joshuaproject.net/people_groups/15484/SA.
- Kamp, Hans. 1984. A theory of truth and semantic representation. In Groenendijk, Jeroen & Janssen, Theo M. V. & Stokhof, Martin (eds.), *Truth, interpretation and information: Selected papers from the third amsterdam colloquium*, 1–42. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110867602.1>
- Kratzer, Angelika. 1981. *The notional category of modality* 38–74. Berlin, Boston: De Gruyter. <https://doi.org/doi:10.1515/9783110842524-004>. <https://doi.org/10.1515/9783110842524-004>
- Kratzer, Angelika. 1989. An investigation of lumps of thought. *Linguistics and Philosophy* 12. 607–653.
- Kratzer, Angelika. 2012. *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199234684.001.0001>. <https://doi.org/10.1093/acprof:oso/9780199234684.001.0001>
- Kratzer, Angelika. 2021. Situations in Natural Language Semantics. In Zalta, Edward N. (ed.), *The Stanford encyclopedia of philosophy*, Metaphysics Research Lab, Stanford University Winter 2021 edn.
- Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4(4). 279–326. <https://doi.org/10.1093/jos/4.4.279>. <https://doi.org/10.1093/jos/4.4.279>
- Löbner, Sebastian. 2011. Concept Types and Determination. *Journal of Semantics* 28(3). 279–333. <https://doi.org/10.1093/jos/ffq022>. <https://doi.org/10.1093/jos/ffq022>
- Ortmann, Albert. 2014. *Definite article asymmetries and concept types: Semantic and pragmatic uniqueness* 293–321. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-01541-5_13. https://doi.org/10.1007/978-3-319-01541-5_13
- Prochazka, Theodore. 1988. Gleanings from southwestern Saudi Arabia. *Zeitschrift für Arabische Linguistik* (19). 44–49. <http://www.jstor.org/stable/43530636>.
- Rabin, Haim. 1951. *Ancient West-Arabian*. Taylor’s Foreign Press.

- Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26. 287–350. <https://doi.org/10.1023/A:1024157132393>
- Russell, Bertrand. 1905. On denoting. *Mind* 14(56). 479–493. <http://www.jstor.org/stable/2248381>.
- Schwarz, Florian. 2009. *Two types of definites in natural language*. University of Massachusetts Amherst. <https://doi.org/https://doi.org/10.7275/1077454>
- Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559. <https://doi.org/10.1111/lnc3.12048>
- Schwarz, Florian. 2019. Weak vs. strong definite articles: meaning and form across languages. *Studies in Diversity Linguistics* (25). <https://doi.org/10.5281/zenodo.3265959>
- Šimík, Radek. 2018. Inherent vs. accidental uniqueness in bare and demonstrative nominals. In *Advances in formal slavic linguistics*, 365–391. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.5155544>
- Strawson, P. F. 1950. On referring. *Mind* 59(235). 320–344. <http://www.jstor.org/stable/2251176>.
- Turner, Michael Lee. 2018. *Definiteness in the Arabic dialects*: The University of Texas at Austin dissertation. <http://hdl.handle.net/2152/68823>.
- Turner, Mike. 2021. Definiteness systems and dialect classification. *Languages* 6(3). <https://doi.org/10.3390/languages6030128>. <https://www.mdpi.com/2226-471X/6/3/128>
- Watson, Janet. 2018. South Arabian and Arabic dialects. In *Arabic Historical Dialectology: Linguistic and Sociolinguistic Approaches*, Oxford University Press. <https://doi.org/10.1093/oso/9780198701378.003.0011>. <https://doi.org/10.1093/oso/9780198701378.003.0011>
- Šereikaitė, Milena. 2019. Strong vs. weak definites: Evidence from Lithuanian adjectives. In Aguilar-Guevara, Ana & Loyo, Julia Pozas & Maldonado, Violeta Vázquez-Rojas (eds.), *Definiteness across languages*, 83–111. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.3252016>