



Asiri, Amer & Gluckman, John. 2025. Revisiting definiteness: Insights from two types of definite articles in Tihami Arabic. *Glossa: a journal of general linguistics* 10(1). pp. 1–40. DOI: <https://doi.org/10.16995/glossa.18227>



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

Amer Asiri, Department of English, College of Education, Majmaah University, Al Majmaah, Saudi Arabia, aj.asiri@mu.edu.sa

John Gluckman, University of Kansas, United States, johnglu@ku.edu

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). Analyses of these languages posit that one definite article encodes solely uniqueness, while the other article includes an anaphoric component. This study presents data from Tihami Arabic which exhibits 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. We formally capture this distinction using Šimík's (2021) situation-based semantics. Our findings reinforce Löbner's/Šimík's approach, and extend our cross-linguistic perspective on languages with two types of definite articles.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



- (3) a. **Im-bayt** lu-hu ʔaʕmidah.
 IM-house for-it pillars
 ‘The house has pillars.’
 b. **As-sama:** ma:l-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, (4) demonstrates that both definite articles appear in anaphoric contexts, (4a) and (4b); in this case, the choice is lexically determined.

- (4) ʕali ga:bala umm-in wa walad-in.
 Ali meet.PFV.3SG.M mother-IN and boy-NUN.
 ‘Ali met a mother and a boy.’
 a. Sallam ʕala **im/*al-walad**.
 shake.hand.PFV.3SG.M on IM/*AL-boy
 ‘He shook the boy’s hand.’
 b. Sallam ʕala ***im/al-mother**.
 shake.hand.PFV.3SG.M on *IM/AL-umm.
 ‘He shook the mother’s hand.’

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 broadly viewed in terms of inherent uniqueness or contextual uniqueness.³ We argue that *im-* attaches to nouns that are not inherently 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. We formalize the lexical classification of Löbner (1985; 2011) in terms of the situation-based framework of Šimík (2021). Briefly, the referent of *im*-NP is evaluated solely relative to the topic situation, while the referent of *al*-NP is evaluated relative to all situations that are similar to the topic situation. Adopting a situation-based semantics covers not just Löbner’s lexical classification, but also extends to cases of “shifts,” where an inherently unique noun is interpreted as non-inherently unique, and vice versa.

publications of the dialect were also consulted, such as Rabin (1951); Greenman (1978); Prochazka (1988); Alahmari (2015); Alqahtani (2015); Alfaifi (2016); Behnstedt & Goldbloom (2016); Watson (2018); Alfaifi & Davis (2021); and Al-Ariqy (2023). We note that there are no corpora of Tihami which we were able to draw from.

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. Ortman 2014).

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 corresponds to uniqueness) versus *strong* (which 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. We are also using the phrase *uniqueness-familiarity distinction* to characterize previous work on two-definite systems. This phrase potentially obscures the fact that in these analyses familiarity entails uniqueness; we will return to this point later.

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. Before concluding in Section 7, we present the analysis for the two definites in Tihami in Section 5.

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 argues that languages with two distinct definite markers correspond to 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 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 of the property of the NP. 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. Uniqueness is exemplified by expressions such as *the king of Jordan* and *the sun*, where the definite description picks out a unique individual that fits the descriptions of the king of Jordan and the sun, respectively. The reference quality, however, does not 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. Weak familiarity involves identifying referents through different mechanisms, such as being directly perceptible to discourse participants, being implied by the context, or being part of shared cultural knowledge (Roberts 2003: 304). For example, a referent can be identified by discourse participants if it is perceptually accessible, contextually established, or globally familiar in the general culture. The other type of familiarity is 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 previous 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 does not 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)

The weak definite, on the one hand, correlates with a uniqueness reading of the definite description, referring 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, is found in contexts in which the definite description refers back to a familiar (unique) entity, already established in the context. As exemplified in (9), the definite description *dem Buch* ‘the book’ refers to the book introduced earlier in the discourse by the indefinite *ein Buch* ‘a book.’ Note, importantly, that the familiar use entails uniqueness of the book as well; that is, uniqueness and familiarity are not in complementary distribution.

- (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 pure 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 reveption 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. An example is given in (11), where the

crisper is part of the fridge, and the definiteness of *im Gemüsefach* ‘in 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** /#in
the fridge was so big that the pumpkin without.problem in.the_{weak} /in
dem Gemüsefach untergebracht werden konnte.
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 *dem Autor* ‘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. Analytically, the weak article encodes uniqueness, while the strong article encodes uniqueness plus an anaphoric element. Overall, this is just one example of many unrelated languages that embody this distinction, such as Akan, American Sign Language, Austro-Bavarian, Dutch, Fering, Icelandic, Korean, Lithuanian, and Mandarin, Mauritian Creole, Thai, Upper Silesian, Upper Sorbian (Ebert 1971a; b; Wespel 2008; Schwarz 2009; 2013; 2019; Arkoh & Matthewson 2013; Ortmann 2014; Simonenko 2014a; Jenks 2015; 2018; Cho 2016; Ingason 2016; Šereikaitė 2016; 2019; Ahn 2017; 2019).

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

Because the uniqueness-familiarity distinction has gained substantial empirical support across various linguistic studies, it is natural to ask whether Tihami’s two definites fit this pattern as well. We show here that this is not a good fit; while there is some overlap, Tihami’s definites largely do not match the descriptive profile of the uniqueness-familiarity dichotomy. This section

presents the most common uses of definite articles in Tihami, largely mirroring the definite uses in Schwarz (2009: Chapter 2). We present standard uses such as uniqueness and anaphoric reference, as well as more specialized applications that distinguish between the definite forms, including covariation, lexical effects, restrictive relative clauses, and generic reference.

3.1 Uniqueness uses

As we saw above, in Standard German, the definites contrast in uniqueness contexts. In Tihami however, 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 *tʰa:wilah* ‘table’. Similarly, (13b) shows that the noun *gamar* ‘moon’ is compatible with both definites and it refers uniquely.⁴

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

3.2 Familiarity uses

Both definites in Tihami are also compatible with (strong) familiarity contexts, where the referent is already established in the discourse. Again, this is a distinguishing factor in Standard German discussed above. However, as shown in (14a), in Tihami both definite articles are possible on the second occurrence of the noun *ʃams* ‘sun’, which has been previously mentioned in the preceding sentence. Similar considerations apply to *intarnit* ‘Internet’ in (14b).

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)–(ii). See Asiri (2024) for an in-depth discussion of the (in)definiteness paradigm in Tihami.

- (i) ʕali ya:kul **tufa:h.**
 Ali eat.IPFV.3.SG.M apple.PL.FEM
 ‘Ali eats apples.’
 (ii) ʕali ya:kul **tufa:hah.**
 Ali eat.IPFV.3.SG.M apple.SG.FEM
 ‘Ali eats an apple.’

We additionally note that, unlike what is reported for Mandarin in Jenks (2018), definiteness marking is not affected by grammatical role; both subjects and objects can appear with either article.

- (14) a. ʕali ʕafa af-ʕams wa al-gamar sawa. **Im-/Aʕ-ʕams** kabi:rah.
 Ali see.PFV.3SG.M AL-sun and AL-moon together. IM/AL-sun big.
 ‘Ali saw the sun and the moon together. The sun is big.’
- b. ʔaham ixtira:ʕain fi ʕayatna al-intarnit wa al-kimbiyu:tar.
 most.important invention.DUAL in life.our AL-Internet and AL-computer.
Im-/Al-intarnit muhimm.
 IM-/AL-Internet important.
 ‘The two most important inventions in our lives are the Internet and the computer.
 The Internet is important.’

3.3 Co-variation uses

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. One such context is when the definite description co-varies with some previous quantification. (We include donkey-anaphoric uses in this category as well.) In Tihami, both *im-* and *al-* appear in co-varying contexts (15)–(16), but they are not interchangeable here. In this respect, Tihami patterns like the Germanic languages reported in Schwarz (2009); Simonenko (2014a; b).

- (15) a. ʕali aʕtʕa kill dʒa:hl-in ku:rah wa impasatʕ akðar min
 Ali give.PFV.3SG.M every child-NUN ball and enjoy.PFV.3SG.M more than
im/*al-dʒa:hl.
 IM/*AL-child
 ‘Ali gave every child a ball and he enjoyed more than the child.’
 (after Schwarz 2009: 4)
- b. Law sʕahib mazraʕah ʕinduh ʕima:r, biyadʕrib **im/*al-ʕima:r**.
 If owner farm has donkey, FUT-beat.3SG.M IM/*AL-donkey
 ‘If a farmer owns a donkey, he beats the donkey.’
 (after Schwarz 2009: 5)
- c. Kull marrah tintahi fi:ha dʒawlah, yaxlutʕ fi:ha
 every time finish.PFV.3SG.FEM in.it round, shuffle.IPFV.3SG.M in.it
im/*al-fa:yiz im-mawra:g wa yawzziʕha min dʒadi:d.
 IM/*AL-winner IM-cards and deal.IPFV.3SG.M.them from anew
 ‘Every time the round is over, the winner shuffles the cards and deals them anew.’
 (after Schwarz 2009: 43)

- (16) a. Fi kill marrah tihtafil as-suṣu:diyyah, al-/*im-malik
 in every time celebrate.IPFV.3SG.FEM AL-Saudi.Arabia, AL-/*IM-king
 yilgi: kalimah.
 give.PFV.3SG.M word
 ‘Every time Saudi Arabia celebrates, the king gives a speech.’
- b. Law ṣayilah raḥan idʒa:zah, al-/*im-umm titḏakar
 If family go.PFV.3SG.FEM vacation, AL-/IM-mother remember.IPFV.3SG.FEM
 wagiḡamsaha.
 sunscreen.her
 ‘If a family goes on a vacation, the mother remembers her sunscreen.’
- c. Yawm ṣayilah tiṣtari lawḥah al-/*im-umm tiṣaligha
 When family buy.IPFV.3SG.FEM painting, AL-/IM-mother hang.IPFV.3SG.FEM-it
 ṣala im-dʒidar.
 on IM-wall.
 ‘When a family buys a painting, the mother hangs it on the wall.’

3.4 Lexical effects

We note that the noun *mother* is a member of a class of nouns that almost always take the definite *al-*, regardless of the type of context it occurs in (17).

- (17) a. Ṣali: sṣa:faḥ *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.’

Other nouns in this class are the days of the week like *al-aḥad* ‘Sunday’ in (18), as well as *as-sama* ‘sky’, *al-mari:x* ‘Mars’, *abb* ‘father’, *ʔilm* ‘science’, *salah* ‘prayer’, along with names of cities, countries, months, and directions. As we discuss below, this is a natural class of nouns: they are “inherently unique”.⁵

⁵ In Faihi Arabic, a subvariety of Tihami, Alfaifi & Davis (2021: 41) note that (non-local) cities (e.g., *al-ga:hirah* ‘Cairo’, *ar-riya:d* ‘Riyadh’), the names of countries (e.g., *al-magrib* ‘Morocco’), days of the week (e.g., *as-sabt* ‘Saturday’ and *al-xami:s* ‘Thursday’), and names of months (e.g., *dul-hijjah* ‘The 12th Hijri Month’) are prohibited with *im-* (see also Alahmari 2015 for a similar observation). Further, Alfaifi (2016: 326) notes that directions always take *al-*, such as *ga:rb:n* ‘west’ or *ja:min* ‘north’.

the uniqueness-familiarity dichotomy. A clear instance of this is that definite *im-* is required on nouns that are modified by a restrictive relative clause (22), parallel to the observation that restrictive relative clauses in Germanic languages tend to require the strong article (Schwarz 2009; Simonenko 2014a; Hanink & Grove 2017).

- (22) a. $\text{ʕat}^{\text{Y}}\text{ni}$ **im-/*al-kita:b** illi yatakallam ʔan ad-dinasora:t.
 give.IMPER.2SG.M-1SG.M IM-/*AL-book COMP talk.IPFV.3SG.M about AL-dinosaurs
 ‘Give me the book that talks about dinosaurs.’
 b. **Im-/*al-madrasah** illi bidʒanbna gafalan.
 IM-/*AL-school COMP next.to.us close.PFV.3SG.FEM
 ‘The school that is next to us closed.’

3.6 Reference to kind uses

Another parallel can be seen with definite generic statements, which tend to correlate with the weak article. In Tihami generic NPs always appear with *al-*, whether the noun itself is singular, plural, countable, or uncountable, as shown in (23a)–(23d) (where we use kind-denoting predicates such as *rare*, *a type of*, and *widespread* to facilitate the relevant reference; cf., Carlson 1977).⁶

- (23) a. **Al/*im-kalb** min anwa:f aθ-θadiyat.
 AL/*IM-dog.SG 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.’

⁶ A reviewer asks whether so-called weak definites like *go to the hospital/store/park* (Carlson et al. 2006) correlate with either definite article in Tihami. In fact, both articles are possible with weak definites, shown in (i) and (ii); the choice appears to be lexically determined. In languages with a “weak/strong” distinction, the weak determiner is always used in these contexts (Schwarz 2013). This further highlights that Tihami’s data does not exactly match what is reported in other languages. We have no comment on the proper analysis of such nouns; cf., Dayal (2004); Schwarz (2014), among others.

- (i) a. ʕali fi: **as-sidʒin** wa saʕad kaman.
 Ali in AL-jail and Saad too
 ‘Ali is in jail and Saad too.’
 b. ʕayyarat ʕali **al-asraʕ** wa saʕad kaman.
 car Ali AL-fast.SUPERLA and Saad too
 ‘Ali’s car is the fastest and Saad’s too.’
 (ii) a. ʕali tasammaʕ li **im-radi:** wa saʕad kaman.
 Ali listen.PFV.3SG.M for IM-radio and Saad too
 ‘Ali listened to the radio and Saad too.’
 b. ʕali ra:ħa **im-biga:lah** wa saʕad kaman.
 Ali go.PFV.3SG.M IM-store and Saad too
 ‘Ali went to the store and Saad too.’

- c. **Ar/*im-ruz** mintafir fi a:sya.
 AL/*IM-rice.PL widespread in Asia
 ‘Rice is widespread in Asia.’
- d. **Al/*im-miyah** na:dirah fi al-marri:x.
 AL/*IM-water.PL rare in AL-Mars
 ‘Water is rare on Mars.’

In sum, while Tihami’s definites do not precisely align with the reported uniqueness-familiarity distinction elsewhere, we also find significant overlaps in distribution. However, on the whole, Tihami appears to exhibit much more freedom in how it uses its two definite articles. As seen in (13), both *im-* and *al-* can establish unique reference, undermining any straightforward mapping between *im-* and anaphoric uses of definiteness. An even greater challenge arises from the sensitivity of both definites to lexical properties of the noun. As demonstrated in (17)–(18), *al-* is completely rejected with certain nouns, while in others, its presence correlates with a particular meaning (19)–(21). This suggests that the choice between *im-* and *al-* is not merely a uniqueness-familiarity distinction but is also governed by deeper lexical and semantic constraints. 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

An alternative to the uniqueness-familiarity 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, and (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). Broadly, he analyses uniqueness as either inherently encoded in the lexical semantics of the definite description or derived from the larger context. This is where he divides definiteness into semantic and pragmatic definites. The core of semantic definiteness is concerned with inherent uniqueness. In contrast, pragmatic uniqueness is concerned with nouns that are not inherently unique and they derive their uniqueness from the context. For example, the noun *sky*, 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 non-unique because its referent can change according to the situation in which the definite description is uttered, that is, 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, as discussed immediately above, is whether the NP is inherently unique or not inherently 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 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.

	not inherently unique	inherently unique
Inherently nonrelational	Sortal nouns (SN):	Individual nouns (IN):
	<i>Dog, stone, book</i>	<i>Sun, weather, Mary</i>
Inherently relational	Relational nouns (RN):	Functional nouns (FN):
	<i>Sister, finger, uncle</i>	<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 not inherently 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 not inherently 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).⁷

While the distinction between the four types of nouns is lexical in nature, Löbner (2011) suggests that these are the *prototypical* types of nouns and predicts that nouns might shift depending on the associated context. For example, the sortal noun *dog* shifts from being a sortal noun to the concept of a dog in (24) that plays a certain role (pet, watchdog, etc.) in a particular range of contexts of utterance. Similarly, an individual noun like *moon* in the context of earth—which is inherently unique—can shift to being a relational noun—non-inherently unique—in (25) in the context of discussing the Planet Mars—which happens to have two moons.

(24) Beware of **the dog**! (written on a board attached to a garden gate) (Löbner 2011: 285)

(25) **The moon** is so small. (mentioned while discussing one of the two moons of the planet Mars)

The key criterion of Löbner's approach is whether the referent of the definite description is determined by inherent or contextual 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

⁷ 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'.

proper names and pronouns in English which do not take a definite article (e.g., *John*, *Bill*, *she*, *they*), for example.

While individual nouns and functional nouns have been recognized as encoding 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 (26). In this context, although *wrestler* and *daughter* are not inherently unique nouns, their uniqueness is derived from the ordinal and superlative modifiers that are considered to be inherently unique.

- (26) 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) (27). In this example, *the woman* becomes unique by virtue of its following restrictive relative clause *he went out with last night*.⁸

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

In general then, Löbner's approach to definiteness claims that definiteness is uniqueness, distinguishing between two types of uniqueness (namely, semantic and pragmatic uniqueness). In the strictest sense, this distinction is lexical: some nouns are "born" unique, while others are contextually determined. However, practically, Löbner leaves room for nouns that can "shift" between semantic and pragmatic uniqueness. To explain this shifting, even inherently unique nouns must be able to be influenced by the context. That is, uniqueness is not necessarily hard-wired into the lexicon, but is in fact *entirely* context-dependent (see also footnote 8). The formal semantics that we ultimately adopt will start from this premise. However, in the next section we will demonstrate how Tihami fits the descriptive classification that Löbner proposes.

⁸ 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 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 and functional nouns. Because the nouns *al-* occurs with are inherently unique, they do not depend on a context for establishing uniqueness. On the other hand, *im-* occurs with nouns that are not inherently unique such as sortal nouns and relational nouns, and consequently they require contexts to draw their uniqueness from, 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 not inherently unique in the right context and vice versa—non-inherently unique nouns shift to being inherently unique. This section underscores the idea that contextual uniqueness is the key factor in Tihami’s definiteness system. In Section 6, 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 a not inherently unique NP is an NP where its referent is different across multiple situations, following Šimík (2021).

5.1 *Al-* as a semantic uniqueness definite

The definite *al-* appears on nouns that are understood 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 contexts. This leads us to the conclusion that *al-* is a semantic uniqueness definite.

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 §5.3.⁹

Initially, let us discuss how *al-* occurs with individual nouns. Löbner (2011: 284) classifies individual nouns into sub-types, namely *role terms*, *institution terms*, *unique objects* or *events*, and *abstract terms*. Interestingly, these sub-types all trigger the appearance of *al-* in Tihami, as demonstrated in (28).

- (28) 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)

⁹ 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-samar* ‘the sky’ while some others, for example, *aʕ/im-fams* ‘the sun’ are more flexible.

- 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: tigarin **al/*im-ħubb bi ad/*im-d^ʕulm.**
 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-rasul* ‘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.¹⁰ Further, a unique object or event is unique in the universal domain as shown in (28c)–(28d); 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 (28d). Abstract terms, as in (28e), likewise, are inherently unique since they refer to one concept, and such a concept can be uniquely referred to in multiple situations, independent of context.¹¹

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 possessum to the possessee, functional nouns require *al-* as a definite marker. Take (29) 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.¹²

(29) **Context:** *Ali and Saad are discussing the visit of their king to their area.*

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

¹⁰ In other work, this would be considered a weak definite (e.g., Carlson et al. 2006; Aguilar-Guevara & Zwarts 2010).

¹¹ 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, simply because these nouns are semantically unique without the article, which leads to semantic redundancy (Löbner 1985: 311).

¹² 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 does not 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. We return to syntactic issues in §5.5.

As we mentioned in the introductory section, there are other nouns that are functional and require *al-*. For instance, *umm* ‘mother’ and *abb* ‘father’ obligatorily appear with the definite *al-*. For any individual, there is a unique mother and father. Note that other relational nouns like, say, *ixt* ‘sister’, are not inherently unique: someone can have more than one sister. Consequently, *ixt* is not required to appear with *al-*.

In summary, the distribution of *al-* can be determined, at least in part, by the lexical semantics of the noun. Nouns that are conceived of as inherently unique all appear with *al-*, and not with *im-*. We return shortly to more complex cases of inherent uniqueness.

5.2 *Im-* as a pragmatic uniqueness definite

In contrast to *al-*, the definite *im-* occurs with nouns that are not inherently 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 *tʿawilah* ‘table’, *bayt* ‘house’, *ʃariʃ* ‘street’, and *kitab* ‘book’ are all sortal nouns that are characterized as being not inherently unique. The other type of nouns that are not inherently unique are relational nouns like *kafarah* ‘tire’, *xal* ‘uncle’, *yadd* ‘hand’, and *hiḏyan* ‘shoes’; these nouns are different from sortal nouns in that they stand in relation to an additional argument. Due to their non-inherent 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 (30). In these examples, the definiteness of the relational noun *sahib* ‘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-*.

- (30) a. **Context:** *Ali has only one friend known to the family members. A family member says the utterance immediately above.*

ʃali aflaha maʔa **im/*al-sa:hib.**

Ali go.PFV.3SG.M with IM/*AL-friend

‘Ali went with his/the friend.’

- b. **Context:** *This is uttered in a room with only one air conditioner.*

ʃayyil **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 (31a), where the uniqueness of *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 (31b).

- (31) a. Walad-in wa bint-in daxalu im-qaḥah. ʔali ʔarafa
 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. ʔift ixtain. **im/*al-ixt** im-sayirah la:bs-ah aḥmar.
 see.PFV.1SG.M sister.DUAL. IM/*AL-sister IM-small wear.IPFV.3SG-FEM red
 ‘I have seen two sisters. The younger sister was wearing red.’

We note one alternative solution: for Simonenko (2014b), the relationship between a restrictive relative and a definite article is formalized as an “anti-uniqueness” presupposition associated with certain determiner heads. On Löbner’s approach, anti-uniqueness would not part of the meaning of the article, because (non-)uniqueness is either encoded lexically or derived from the context.

Another 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 licenser is not co-referential, and uniqueness is derived by the virtue of a link between the two elements. As given in (32a), *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 (32b), where the noun is the producer of its licenser. For Schwarz (2009), the part-whole associative anaphora is different from the producer-product associative anaphora in that the licensing noun in the former is a present discourse participant, while it is not in the producer-product relationship. In Tihami, both cases require *im-*.

- (32) a. ʔali iʔtara sayyarah. **Im/*al-gaza:z** kana maksur-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-muʔalif** faransi.
 read.PFV.1SG book-NUN. IM/*AL-author French
 ‘I have read a book. The author is French.’
 (Associative anaphora: producer-product relationship)

As the uniqueness of the definite *im-* is drawn from the context, it is no surprise that *im-* occurs on elements in co-varying contexts, where the uniqueness of the nominal varies with the selected situation. In each selected situation, the uniqueness of the definite description is evaluated relative to that situation. As we consider the examples in (33), reproduced from (15), we find that the referent of the definite description varies from a particular set of individuals as in (33a), where *the child* is selected from the set of children which Ali gave every one of them a ball. In (33c), there is a possibly different winner in each of the rounds of cards.

- (33) a. ʕali aʕtʕa kill dʒa:hl-in ku:rah wa impasatʕ akðar min
 Ali give.PFV.3SG.M every child-NUN ball and enjoy.PFV.3SG.M more than
im/*al-dʒa:hl.
 IM/*AL-child
 ‘Ali gave every child a ball and he enjoyed more than the child.’
 (after Schwarz 2009: 4)
- b. Law sʕahib mazraʕah ʕinduh ʕima:r, biyadʕrib **im/*al-ʕima:r.**
 If owner farm has donkey, FUT-beat.3SG.M IM/*AL-donkey
 ‘If a farmer owns a donkey (and a goat), he beats the donkey.’
 (after Schwarz 2009: 5)
- c. Kull marrah tintahi fi:ha dʒawlah, yaxlutʕ fi:ha
 every time finish.PFV.3SG.FEM in.it round, shuffle.IPFV.3SG.M in.it
im/*al-fa:yiz im-mawra:g wa yawzziʕha min dʒadi:d.
 IM/*AL-winner IM-cards and deal.IPFV.3SG.M.them from anew
 ‘Every time the round is over, the winner shuffles the cards and deals them anew.’
 (after Schwarz 2009: 43)

Crucially, not all co-varying contexts require *im-*. A definite description can take *al-* in co-varying contexts, just in the case that the referent is in the class of nouns that are inherently unique. We repeat those examples in (34).

- (34) a. Fi kill marrah tihtaʕil as-suʕu:diyyah, **al-/im-malik**
 in every time celebrate.IPFV.3SG.FEM AL-Saudi.Arabia, AL-/IM-king
 yilgi: kalimah.
 give.PFV.3SG.M word
 ‘Every time Saudi Arabia celebrates, the king gives a speech.’
- b. Law ʕayilah raħan idʒa:zah, **al-/im-umm** titðakar
 If family go.PFV.3SG.FEM vacation, AL-/IM-mother remember.IPFV.3SG.FEM
 waġiʕamsaha.
 sunscreen.her
 ‘If a family goes on a vacation, the mother remembers her sunscreen.’

- c. Yawm ʕayilah tiʕtari lawḥah **al-/im-umm** tiʕaligha
 When family buy.IPFV.3SG.FEM painting, AL-/IM-mother hang.IPFV.3SG.FEM-it
 ʕala im-dʒidar.
 on IM-wall
 ‘When a family buys a painting, the mother hangs it on the wall.’

To recapitulate, the definite *im-* occurs with nouns that are not inherently 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, as well as co-varying contexts.

5.3 Shifting

In this section, we will address cases where nouns “shift” between being treated as inherently and non-inherently unique. As we show, this shift correlates with the presence of *al-* and *im-*. We note that, though (non-)inherent uniqueness is framed as a lexical distinction by Löbner, this kind of shifting is actually covered in his approach, which permits some flexibility in the effect of lexical semantics.

In Tihami, 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 earlier. For instance, consider the noun *malik* ‘king’ which is inherently unique. We predict that it appears with *al-*, and it does as shown in (35a). However, if the sentence is uttered in a particular linguistic context, for instance, with a restrictive relative clause, as shown in (35b), then *im-* is the preferred definite, referring to a particular king in the time and place situated in the context of the utterance, who is not unique for the discourse participants (see similar examples in (19)–(21) above).

- (35) a. **Al-/Im-malik** asʕdara amir.
 AL-/IM-king announce.PFV.3SG.M decree
 ‘The king announced a decree.’
 b. ʕiftah ***al-/im-malik** illi kana fi im-ḡaḡah?
 see.PFV.2SG AL-/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 (36), 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 (36a) is the individual noun *Quran*, while in (36b) refers to a particular copy of the Quran, which the speaker specified being on the table.

- (36) a. Garayt **il-/m-gur?a:n.**
 read.PFV.1SG AL-/IM-Quran
 ‘I read the Quran.’
 b. Garayt ***il-/m-gur?a:n** ði: ʕala m-tʰa:wilah.
 read.PFV.1SG AL-/IM-Quran that on IM-table
 ‘I read the Quran that is on the table.’ (Alfaifi & Davis 2021: 43)

The effect of relative clauses on the choice of definite article has been well-established in the literature outside of Tihami (cf., Schwarz 2009). There are a few options to analyze this behavior. The first is to follow precedent and argue for a syntactic solution: only *im-* provides the syntactic structure that a relative clause needs (Wiltschko 2012; Simonenko 2014a; Hanink & Grove 2017), this would require us to assume that *im-* and *al-* occupy distinct syntactic positions or embed different structures. We discuss structural considerations in Section 5.5.

However, Löbner’s approach offers an alternative explanation that does not necessarily rely on structure. Restrictive relative clauses, by definition, narrow down the set of individuals given by the head noun. For inherently unique individuals, that set is already a singleton set; it cannot be further restricted (cf., Wiltschko 2012). A restrictive relative clause forces a non-inherently unique interpretation of the head noun (to the extent that one is possible), and therefore *im-* is chosen as the definite. Note that non-restrictive relative clauses do not affect the choice of definite article in Tihami (37).¹³

- (37) 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.’

It is also possible for non-unique nouns such as sortal nouns and relational nouns to shift to being unique. This is the case with generics. Although Löbner (2011: 279–80) abstracts away from discussing generics, inherent uniqueness has a clear extension to genericity as noted by Šimík (2021: 381).¹⁴ Generics can refer to kinds or be used in characterizing statements. It is virtually possible for any type of nouns to be generic. A singular definite *the dog* can refer to a taxonomic individual, i.e., *The dog is a mammal*. While the bare plural *dogs* can be used to make a characterization about no particular dog: *Dogs bark*.

¹³ We note one alternative solution: for Simonenko (2014b), the relationship between a restrictive relative and a definite article is formalized as an “anti-uniqueness” presupposition associated with certain determiner heads. On Löbner’s approach, anti-uniqueness would not be part of the meaning of the determiner, because (non-)uniqueness is either encoded lexically or derived from the context.

¹⁴ Definite generics have been addressed by many researchers (Carlson 1977; Krifka et al. 1995; Chierchia 1998; Longobardi 2001; Dayal 2004, among many others). Languages with articles use definite NPs to refer to kinds either obligatorily such as in Spanish (Borik & Espinal 2015) or in variation with bare NPs such as Brazilian Portuguese (Schmitt & Munn 1999).

In Tihami, generic definite NPs always appear with *al-*. Intuitively, this is because a generic NP always refers to the same generic individual. While an individual referent for *the dog* might vary across contexts, the generic or taxonomic reference (as in, e.g., Chierchia 1998 and Dayal 2004) does not. This is why *al-* is required in Tihami. Consider (23a), repeated in (38) below for convenience. In this example, *kalb* ‘dog’ is a sortal noun that shifts to generic by virtue of the kind denoting predicate *a type of*.

- (38) **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.’

We believe that this “genericity” shift from non-inherent to inherent uniqueness explains that ambiguities noted earlier and repeated in (39)–(40).

- (39) a. ʕali ʕafa: **al-ardʕ**.
 Ali see.PFV.3SG.M AL-earth
 ‘Ali has seen the earth.’
 b. ʕali ʕafa: **im-ardʕ**.
 Ali see.PFV.3SG.M IM-earth
 ‘Ali has seen the (piece of) land.’
- (40) a. **An-na:r** ha:rah.
 AL-fire hot
 ‘The hell fire is hot.’
 b. **Im-na:r** ha:rah.
 IM-fire hot
 ‘The fire is hot.’

In this case, we suggest that *al-ardʕ* ‘the earth’ and *an-na:r* ‘hell fire’ are functionally kinds; they refer to generic uses of “earth” and “fire”. With *im-*, they refer to an individual piece of land or fire, respectively.

In short, nouns that are labeled as inherently unique can shift, to a large extent, to be non-unique nouns. On a similar note, nouns that are labeled as not inherently unique can shift to be inherently unique, and they provide generic interpretations.

As a final note, we believe that shifting—and inherent uniqueness in general—is at least in part the result of the historical development of the two definite markers in Tihami. We think that Arabic dialects begin with a single definite marker, *al-*, and that Tihami later distinguishes between what is inherently unique and what is not. As nouns began to use *im-*, several nouns had their definite marker lexicalized, resisting the shift to use the definite *im-*. Inherently unique

nouns with prestigious indications and those of high value to Tihami society tend to be more lexicalized and the last nouns to shift, if they do at all.

6 Formalizing non-inherently unique and inherently unique definites

The analysis we adopt for the two definite markers is an extension of the accounts in Löbner (1985; 2011) and Šimík (2021), both of which are grounded in a theory of uniqueness. Despite their fundamental difference—Löbner’s account being lexically oriented and Šimík’s being situation-based—they converge in ways supported by the Tihami Arabic data. Löbner’s framework categorizes definites based on inherent uniqueness or non-inherent uniqueness, while allowing for shifts between these categories. Šimík’s model complements this by explaining how the uniqueness of definite descriptions is evaluated in context, especially for nouns that shift and are capable of taking both definite markers, depending on the type of context. While nouns that resist shifting have been addressed earlier, the focus now turns to those nouns that exhibit flexibility by undergoing shifting and being compatible with both definite forms. This phenomenon will be analyzed through the framework provided by Šimík’s account.

The analysis in Šimík (2021) aims to explain inherent uniqueness and non-inherent (called “accidental”) uniqueness in Czech, which are realized by bare versus demonstrative NPs.¹⁵ See (41).

- (41) a. **Context:** *I approached a friend’s house.*
 Garáž zářila novotou.
 garage shined novelty.INSTR
 ‘The garage shined with novelty.’
 b. **Context:** *A friend showed me his new garage.*
 Ta garáž zářila novotou.
 DEM garage shined novelty.INSTR
 ‘The garage shined with novelty.’

Czech (Šimík 2021: 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 captured by universally quantifying over situations that are like the topic situation, but in

¹⁵ Šimík (2021) uses the term *accidental uniqueness* which is mapped to our non-inherent uniqueness. For the sake of consistency with Löbner’s approach, we will use the term *non-inherent uniqueness* rather than *accidental uniqueness*.

which the number of NPs satisfying the NP description varies. Here, the uniqueness of the NP is restricted to the topic situation and it is not the case that in all situations that are like the topic situation, there is one entity with property P that satisfies the uniqueness of NP; there may be none or more than one.

A crucial element of Šimík's (2021) 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.¹⁶ The purpose of "likeness" is to restrict the quantification of inherent/non-inherent uniqueness to contextually relevant situations, excluding abnormal situations. It eliminates situations where inherent uniqueness possibly fails because of "abnormal" conditions. For example, the inherent uniqueness of the *sky* 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/past 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, in relation to other relevant situations. 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 the two definite markers of Tihami is presented in (42) and (43), respectively, modeled after Šimík's analysis for Czech.¹⁷

¹⁶ 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 is a set of worlds compatible with what we know, what we believe, what we hope, what is normal, etc. The ordering source ranks the worlds in the modal base so that, in the case of likeness, worlds that are more similar to the topic world are ranked higher than others.

¹⁷ We feel that a slight clarification is in order. In the definitions from Šimík (2021), the noun is evaluated with respect to the "topic situation." This may give the impression that this situation is the topic situation for the sentence. Of course, this is not always true. In some cases, the (Austinian) topic situation may very well be distinct from the situation used

(42) **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.

(43) **The not inherently 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-inherently 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 (44), 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.

(44) **Context:** *Ali is looking directly at the sky.*

As-sama: ma:l-ha ?aʕmidah.

AL-sky NEG-for-it pillars

‘The sky has no pillars.’

$$\forall s[s \approx s_T \rightarrow \exists!x[SKY(x)(s_T)]]$$

By contrast, the referent of the non-unique noun like *bayt* ‘house’, in (45), repeated from the above, is restricted to the topic situation, relative to the other similar situations. That is, there is a unique house in s_T , but it is not the case that in all situations that are like s_T there is exactly one house. It is possible that the speaker does not have a house, or that there are multiple houses.

(45) **Context:** *Two family members are talking about their only house.*

Im-bayt lu-hu ?aʕmidah.

IM-house for-it pillars

‘The house has pillars.’

$$\exists!x[HOUSE(x)(s_T) \wedge \neg \forall s[s \approx s_T \rightarrow \exists!y[HOUSE(y)(s)]]]$$

to evaluate the definite expression. In such cases, it would be more appropriate to use the term “resource situation” for evaluating the noun.

We note the importance of the topic situation (or more accurately, the resource situation) in evaluating the referent of the NP. Consider (46) and (47) 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 in his kingdom, and therefore it is marked with *al-*. However, it is one of the nouns that can shift in a particular situation. In (46) and (47), the choice of the definite marker picks out different referents.

Context:

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

- (46) **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.’

$$\forall s[s \approx s_1 \rightarrow \exists!x[\text{KING}(s)(x)]]$$

(where s_1 is a situation based on the usual shared experience of Ali and Saad; in that situation, there is normally a single king, namely their king)

- (47) **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.’

$$\exists!z[\text{KING}(s'_1)(z)] \wedge \neg \forall s[s \approx s'_1 \rightarrow \exists!x[\text{KING}(s)(x)]]$$

(where s'_1 is the environmental summit meeting)

In this context, the choice of *im-* versus *al-* depends on which situation the speaker is using as their resource situation to evaluate the noun. In (46), the speaker chooses a situation like “our homeland”. For the speakers of Tihami, the situations that are like this situation will all have exactly one king, so the speaker must be referring to their homeland king—it is the only king that is guaranteed to appear in all relevant situations for native Tihami speakers in the context. In (47), however, by uttering *im-*, the speaker is using a different resource situation, specifically, the environmental summit. As we look at situations that are similar to the environmental summit, it is possible to find situations without any kings or with multiple kings. The speaker therefore uses *im-* to indicate the African king, since this individual is not

constant across like situations. As noted by Šimík (2021), this approach requires some extra explanation for why a definite description is permitted at all in the environmental summit context above because there are two kings: the Tihami speakers' king and the African king. Uniqueness does not hold. Ultimately, Šimík (2021: 380) suggests that inherently and non-inherently unique individuals are not “competitors” when calculating uniqueness in a situation. The inherently unique king—“our king”—is not considered when calculating uniqueness in a particular situation. Note that if we slightly modify the situation so that there is a European king at the conference in addition to the African king and the Saudi king, *im-* is not possible at all, as it leads to a presupposition failure.¹⁸

Similar considerations apply for (48) and (49). With *al-*, Ali, the speaker, invites Saad, the listener, to accommodate a resource situation like “our school”. Across all situations that are like this situation, there is only one referent for the principal: the principal of Ali and Saad's school. With *im-*, however, the resource situation is the education forum. As we (i.e., the interlocutors) look across similar situations, it is not the case that there is only one school principal in them.¹⁹

Context:

Ali and Saad attended an education forum with many teachers and students from many schools. Among the attendees are two principals: Ali and Saad's school principal, and another principal from a rival school. Ali later said:

¹⁸ Šimík suggests that we might appeal to the notion of “anti-uniqueness”. The precise mechanism of the way the domain is restricted for an anti-uniqueness inference is left for future research (for relevant anti-uniqueness discussion see Hawkins 1978; Heim 1991; Sauerland 2008). Alternatively—as suggested by a reviewer—we might appeal to one of the other competition-based accounts (e.g., Jenks 2018 or Ahn 2019) to solve this issue. For example, we might adopt Jenks' intuition behind *Index!* which is a consequence of *Maximize Presupposition* (Heim 1990). Since the presupposition associated with the non-inherently unique definite is stronger, given the chance, we use that one. However, this still does not solve the issue of the two kings in the environmental summit context.

¹⁹ A reviewer notes that the example of *im/*al-mikayyif* ‘the air conditioner’ in (30b) above is somewhat surprising given that in Šimík's analysis of definiteness in Czech, he observes that in the following situation the bare NP, which correlates with inherent uniqueness, is appropriate.

- (i) *Situation:* Two student assistants A and B are at their shared work desk which they share with other student assistants and where there's a computer and a couple of other things, including a book (it doesn't really matter to whom the book belongs). A is looking for a pencil, B says:

Nějaká tužka je vedle {počítače / #toho počítače}
some pencil is next.to computer DEM computer

‘There's a pencil next to the computer.’

Czech (Šimík 2021: 377)

We actually think that Šimík's judgment is more surprising than that of Tihami, since the similarity set for the topic situation very plausibly includes situations in which there is no computer at all on the desk. This difference appears to speak to a larger issue about how we choose resource situations and construct similarity sets—a topic we cannot fully address here.

- (48) **Al-midi:r** **xaradʒa** **bisurfah.**
 AL-principal go.out.PFV.3SG.M fast
 ✓‘The principal (of Ali and Saad’s school) went out fast.’
 ×‘The principal (from the rival school) went out fast.’
 $\forall s[s \approx s_1 \rightarrow \exists!x[\text{PRINCIPAL}(s)(x)]]$
 (where s_1 is a situation based on the usual shared experience of Ali and Saad, where their principle is uniquely relevant)
- (49) **Im-midi:r** **xaradʒa** **bisurfah.**
 IM-principal go.out.PFV.3SG.M fast
 ×‘The principal (Ali and Saad’s school) went out fast.’
 ✓‘The principal (from the rival school) went out fast.’
 $\exists!z[\text{PRINCIPAL}(s'_1)(z)] \wedge \neg \forall s[s \approx s' \rightarrow \exists!x[\text{PRINCIPAL}(s)(x)]]$
 (where s'_1 is the education forum meeting to the exclusion of the principal of Ali and Saad’s school)

This idea extends to definite generics as well, as noted in Šimík (2021: 381–382). We assume, following Carlson (1977); Krifka et al. (1995); Krifka (2003); Dayal (2004), that definite generics make reference to kinds. While *the dog*, for example, can pick out a contextually unique individual $\lambda s.\iota x[\text{DOG}(x)(s)]$, it can also refer to a taxonomic kind: $\lambda s.\iota x[\text{DOG}_k(x)(s)]$. In this case, it refers not to a particular individual, but rather to the species *Canis*. The use of *al-* with kinds follows, as across like situations, the taxonomic definition of DOG—dog-kind—is constant and unique.²⁰ A statement like the one in (50a) is often evaluated relative to a large topic situation or possibly the whole world (which is the maximal situation). Here, the statement satisfies the definition in (50b), in which all worlds that are similar to the actual world (that have similar parameters) are such that they have only one dog-kind in them.

- (50) a. **Al-kalb** **min anwa:f aθ-θadiyat.**
 AL-dog.SG.GEN from types AL-mammal.PL
 ‘The dog is a type of mammal.’
 b. $\forall s[s \approx s_0 \rightarrow \exists!x[\text{DOG}_k(x)(s)]]$

We leave open the issue of weak definites, as in expressions like *go to the hospital/store/park*, which have been argued to be kind-denoting in their semantic representation (Aguilar-Guevara & Zwarts

²⁰ Obviously we can be more explicit about the mapping between properties and kinds, adopting versions of the $\cap$ and $\cup$ operators, as in e.g., Chierchia (1998), among others.

- (i) a. $\cap P = \lambda s.\iota x[P(x)(s)]$, defined iff x is in KINDS
 b. $\cup k = \lambda s\lambda x.x < k(s)$, for any kind k

2010). In Tihami, as discussed in footnote 6, both definites are possible with weak definites, and it is not entirely clear what determines the use of *al-* versus *im-* in these constructions. (We note that the issue is not resolved by appealing to a uniqueness-familiarity distinction for Tihami, since weak definites consistently use the (morphologically) weak form of the definite in those languages that have the weak/strong distinction.)

Turning to anaphoric contexts, Šimík (2021: 382) suggests that anaphoric reference can be achieved not by tracking individual referents within the discourse, but rather by tracking situations. In this case, the noun is evaluated relative to a salient situation (a resource situation). Recall from earlier that both *im-* and *-al* are possible in anaphoric contexts, depending on the noun.

- (51) Walad-in wa bint-in daxalu im-qaṣah. ʕali ʕarafa **im/*al-bint**.
 boy-NUN and girl-NUN enter.PFV.3PL IM-hall. Ali recognize.PFV.3SG.M IM/*AL-girl
 ‘A boy and a girl entered the hall. Ali recognized the girl.’

In (51), the situation that is relevant for interpreting the second occurrence of *bint* ‘girl’ is the minimal situation of Ali meeting the two children. Across like situations, there need not be a unique individual which satisfies the description of “the girl.” Thus, *im-* is required in this anaphoric context. However, in (52), anaphoric reference to the day that we met does not trigger *im-*, because, across like-situations, there is a unique referent for the day Sunday.

- (52) Tiga:balna al-ahad wa al-iḥnain. ***Im/Al-aḥad** kanna wana:sah.
 meet.PFV.1PL AL-Sunday CONJ AL-Monday. *IM/AL-Sunday COP.3SG.M fun
 ‘We met on Sunday and Monday. Sunday was fun.’

Nothing different needs to be said about associative anaphora. When evaluating the definite NP, the resource situation is given by the preceding sentence.

- (53) a. ʕali ijtara sayyarah. **Im/*al-gaza:z** kana maksur-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-muʔalif** faransi.
 read.PFV.1SG book-NUN. IM/*AL-author French
 ‘I have read a book. The author is French.’
 (Associative anaphora: producer-product relationship)

- (54) ʕali ga:bala ʕayilah. ***Im/Al-umm** ka:nan dʒami:lah.
 Ali meet.PFV.3SG.M family. *IM/AL-mother COP.3SG.FEM beautiful.
 ‘Ali met a family. The mother was beautiful.’

The situation for evaluating the windshield in (53a) is the preceding Ali-buying-a-car situation. Likewise, the situation for evaluating the author in (53b), is the preceding me-reading-a-book situation. Importantly, in both of these situations, the window and the car are not inherently unique. In contrast, in (54), in the situation of Ali-meeting-a-family, there is an inherently unique referent for mother for the family mentioned in the utterance.

The same idea extends to co-varying contexts. In a context like (55). As we noted above, in examples paralleling those in the literature, *im-* is preferred. This is because the uniqueness of the definite description is satisfied in each chosen situation, and as we look into similar situations, the winner is different—or there may not be a winner.

- (55) Kull marrah tintahi fi:ha dʒawlah, yaxlutʰ fi:ha
 every time finish.PFV.3SG.FEM in.it round, shuffle.IPFV.3SG.M in.it
im/*al-fa:yiz im-mawra:g wa yawzziʃha min dʒadi:d.
 IM/*AL-winner IM-cards and deal.IPFV.3SG.M.them from anew
 ‘Every time the round is over, the winner shuffles the cards and deals them anew.’
 (after Schwarz 2009: 43)

Here, we assume that what is being quantified over are situations, which subsequently act as resource situations for the noun (cf., Elbourne 2005). In (55), each round is a situation in which there is a winner, and a similarity set is calculated for each winner. In those sets, it may well be that there is no winner (e.g., perhaps the game was canceled).

Importantly, we can also construct scenarios where there is a unique individual in all of the similar situations of all of the quantified situations, as in (56).

- (56) Yawm ʃayilah tiʃtari lawhah **al-umm** tiʃaligha ʃala
 When family buy.IPFV.3SG.FEM painting, AL-mother hang.IPFV.3SG.FEM-it on
 im-dʒidar.
 IM-wall
 ‘When a family buys a painting, the mother hangs it on the wall.’

Here, even though *umm* ‘mother’ is co-varying with families (or family-buying-a-picture situations), we still see *al-* as the chosen article. This is because there will always be a unique mother for any family in all relevant situations.

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.

6.1 Syntactic factors

While we are not proposing a syntactic analysis, it is possible that the two definites are structurally distinct. This is suggested in their interaction with demonstratives. When there is a prenominal demonstrative, only *al-* is permitted. When there is a post-nominal demonstrative, only *im-* is possible.²¹

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

Such data suggest that *im-* and the demonstrative compete for a syntactic position, and that that position is higher than *al-*. (Or equally, the definite articles are realizations of the same syntactic head, but that head is situated in distinct syntactic structures.)

The idea that the definite articles correlate with different structures is theoretically significant, and it has strong support in the literature (Elbourne 2008; Schwarz 2009; Simonenko 2014a; Hanink 2018). It is in fact of central importance in Schwarz’s formalism (and later Simonenko 2014a; Jenks 2018 and Hanink 2018; see also Wiltschko 2012). In this line of work, the meaning components of the strong definite determiner are arrayed in the syntax, capturing Schwarz’s proposal that the strong determiner is built by introducing an index “on top” of the weak article (cf., Elbourne 2008).

- (58) Weak article
 a. $\llbracket D_{\text{weak}} \rrbracket = \lambda s_r \lambda P. \iota x. P(x)(s_r)$
 b. $[_{DP} D [_{NP} N]]$

²¹ The same is true for the distal demonstrative; (ia) and (ib). While the strong definite in German can function demonstratively, similar to the English distal demonstrative *that* (Schwarz 2009: 34–37), one might wonder whether the two definite articles in Tihami compete in demonstrative contexts. However, they cannot be used as distal demonstratives, and the presence of a demonstrative does not influence the selection between the two definites on the associated noun. Instead, the choice is solely governed by the noun’s syntactic positions relative to the demonstrative as in (i).

- (i) a. *ʕali dʒa: fi ta:k *im/as-sayyarah ma:hu ta:k.*
 Ali come.PFV.3SG.M in that *IM/AL-car NEG that
 ‘Ali came in that car not that one.’
 b. *ʕali dʒa: fi im/*al-sayyarah ta:k ma:hu ta:k.*
 Ali come.PFV.3SG.M in IM/*AL-car that NEG that
 ‘Ali came in that car not that one.’

(modeled after Schwarz 2009: 34)

(59) Strong article

- a. $\llbracket D_{\text{strong}} \rrbracket = \lambda s_r \lambda P \lambda y. \iota x. P(x)(s_r) \ \& \ x = y$
- b. $[_{DP} \ 1 \ [_D \ [_{NP} \ N \] \]]$

(Schwarz 2009: 264f)

What is crucial in this idea is that the strong article entails the meaning of the weak article; the strong article is the weak article plus an anaphoric component. This meaning entailment is captured in the structural representation. Ideally, we would try to extend the same reasoning to Tihami, but we do not see how it is easily accommodated in the present framework. The issue is that there is no entailment relationship between the meanings of inherent and non-inherent uniqueness given above. It is not possible (within reason) to build one from the other.²²

Moreover, it is not clear what we would gain from “syntactizing” the proposal here. Our main claims are (i) that the weak/strong dichotomy simply does not work for Tihami, and (ii) that Tihami’s system is best captured by appealing to (non-)inherent uniqueness. These claims are independent of the formalisms associated with either approach.

All that said, we will emphasize that we are very much amenable to a syntactic analysis for the Tihami facts discussed here. We have taken the analysis from Šimík (2021) “off the shelf”, but it is possible that reconfiguring these ideas might motivate a more straightforward syntax-semantics mapping. But ultimately this is orthogonal to the main claims of the paper. We leave the apparent structural differences in (57) for future work.

7 Conclusion

Cross-linguistic variation in expressing definiteness reveals a complex interplay between form and meaning. Languages utilize a range of strategies to mark definiteness, spanning from bare nominals to the use of definite articles or demonstratives/pronouns. When a language uses more than one strategy for expressing definiteness, these different strategies typically correspond to distinct semantic implications (Schwarz 2009; Jenks 2015). A key distinction in languages that deploy more than one strategy is the uniqueness-familiarity distinction. Schwarz’s (2009) analysis is rooted in the distinction between weak definite article (encoding uniqueness) and strong definite article (encoding uniqueness plus anaphoricity). However, not all languages align perfectly with this binary framework; one example of such a language is Reunion Creole (Réyoné), a French-related creole language, which exhibits a three-way definiteness split, distinguishing sortal, individual, and functional concepts, as outlined in Löbner’s theory, using bare nouns for individual concepts, demonstrative *-la*-marked noun phrases for pragmatic definiteness with sortal

²² That said, we want to point out one potential way forward on this front. Since the non-inherent definite contains the negation of the universal quantification found in the inherent definite, we might capture the difference in terms of a scalar implicature: the non-inherent definite looks like “some (but not all).” We leave this idea for future work.

concepts in deictic and anaphoric contexts, and *lo*-marked noun phrases dedicated to functional concepts (Albers 2024). Tihami Arabic is another example of a language that does not fall under this distinction, as has been discussed in this study.²³ It remains an open question about what accounts for the crosslinguistic variation.

In this study, we have examined the distribution of the two definite articles of Tihami: *im*- and *al*-. 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. Data reveal that *al*- attaches to inherently unique nouns, while *im*- marks nouns whose uniqueness is discourse-dependent, while allowing nouns to shift between inherent and non-inherent uniqueness. Analytically, we follow Šimík (2021), arguing 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.

²³ An additional possible language that does not fall under the uniqueness-familiarity distinction of Schwarz (2009) is Haitian Creole which employs its overt marker *la* across both situational and anaphoric contexts, suggesting a distinct semantic contrast possibly tied to situation pronouns rather than a weak-strong distinction. (See Schwarz 2019.)

Abbreviations

1, 2, 3 = 1st/2nd/3rd person, AL- = The definite marker *al-*, COP = Copula, COMP = Complementizer, CONJ = Conjunction, D = Determiner, DEM = Demonstrative, DP = Determiner phrase, DUAL = Dual, EXPL = Expletive, FEM = Feminine, FN = Functional noun, FUT = Future tense, GEN = Generic, IM- = The definite marker *im-*, IMPER = Imperative, IN = Individual noun, INSTR = Instrumental, IPFV = Imperfective, M = Masculine, NEG = Negation, N = Noun, NP = Noun phrase, NUN = Nunation, O = Object, PFV = Perfective, PL = Plural, RN = Relational noun, SG = Singular, SN = Sortal noun, STRONG = Strong definite, SUPERLA = Superlative adjective, WEAK = Weak definite

Acknowledgements

The authors extend the appreciation to the Deanship of Postgraduate Studies and Scientific Research at Majmaah University for funding this research work through the project number: R-2025-1892.

Competing interests

The authors have no competing interests to declare.

References

- Abbott, Barbara. 2006. Definiteness and indefiniteness. In *The handbook of pragmatics*, 122–149. Blackwell Publishing. DOI: <https://doi.org/10.1002/9780470756959.ch6>
- Aguilar-Guevara, Ana & Zwarts, Joost. 2010. Weak definites and reference to kinds. In Li, Nan & Lutz, David (eds.), *SALT 20: Proceedings of the 20th semantics and linguistic theory conference*. 179–196. Ithaca, NY: CLC Publications. DOI: <https://doi.org/10.3765/salt.v20i0.2583>
- 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/>.
- Ahn, Dorothy. 2019. *THAT thesis: A competition-based mechanism for anaphoric expressions*: Harvard University dissertation.
- Al-Ariqy, Mohammed J. 2023. Definite determiner alternation in Yemeni Tihami Arabic. *Coyote Papers* 25. 1–10. <http://hdl.handle.net/10150/668446>. DOI: <https://doi.org/10.3765/amp.v10i0.5564>
- Alahmari, Musa. 2015. *The development of m- definite article in Arabic: A historical linguistic perspective*. Indiana University Bloomington Master's thesis. Unpublished.
- Albers, Ulrike. 2024. A three-way distinction in definiteness: sortal, individual, and functional concepts in Réyoné. *Linguistics* 62(5). 1185–1223. DOI: <https://doi.org/10.1515/ling-2021-0218>

- Alfaifi, Abdullah & Davis, Stuart. 2021. An examination of the /m-/ definite article in upper Faifi Arabic. *Zeitschrift für arabische Linguistik* 2021(73). 36–52. DOI: <https://doi.org/10.13173/zeitarabling.73.0036>
- Alfaifi, Ali. 2016. *Aspects of the morpho-syntax of the Faifa dialect of Saudi Arabia: A descriptive study*. Ulster University dissertation.
- 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. DOI: <https://doi.org/10.1016/j.lingua.2012.09.012>
- Asiri, Amer. 2024. *Definite and indefinite affixes in Tihami Arabic*. University of Kansas dissertation. <https://www.proquest.com/dissertations-theses/definite-indefinite-affixes-tihami-arabic/docview/3069327258/se-2>.
- Behnstedt, Peter & Goldbloom, Gwendolin. 2016. *Dialect atlas of north Yemen and adjacent areas*. Leiden, The Netherlands: Brill. DOI: <https://doi.org/10.1163/9789004326422>
- Borik, Olga & Espinal, M. Teresa. 2015. Reference to kinds and to other generic expressions in Spanish: Definiteness and number. *The Linguistic Review* 32(2). 167–225. DOI: <https://doi.org/10.1515/tlr-2014-0023>
- Carlson, Gregory. 1977. *Reference to kinds in English*. University of Massachusetts at Amherst dissertation.
- Carlson, Gregory & Sussman, Rachel & Klein, N. & Tanenhaus, M. 2006. Weak definite noun phrases. In Davis, Chris & Deal, Amy Rose & Zabbal, Youri (eds.), *NELS 36: Proceedings of the 36th annual meeting of the North East Linguistic Society*, 179–196. Amherst, MA: GLSA.
- Chierchia, Gennaro. 1998. References to kinds across languages. *Natural Language Semantics* 6. 339–405. DOI: <https://doi.org/10.1023/A:1008324218506>
- Cho, Jacee. 2016. The acquisition of different types of definite noun phrases in L2-English. *International Journal of Bilingualism* 21(3). 367–382. DOI: <https://doi.org/10.1177/1367006916629577>
- 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. DOI: <https://doi.org/10.5281/zenodo.3252014>
- Dayal, Vaneeta. 2004. Number marking and indefiniteness in kind terms. *Linguistics and Philosophy* 27. 393–450. DOI: <https://doi.org/10.1023/B:LING.0000024420.80324.67>
- Ebert, Karen. 1971a. *Referenz, Sprechsituation und die bestimmten Artikel in einem Nordfriesischen Dialekt (Fering)*. Kiel: Christian-Albrechts-Universität zu Kiel (Doctoral dissertation).
- Ebert, Karen. 1971b. Zwei Formen des bestimmten Artikels. In Wunderlich, Dieter (ed.), *Probleme und Fortschritte der Transformationsgrammatik*, 159–174. München: Hueber.
- Elbourne, Paul. 2005. *Situations and individuals*. Cambridge MA: MIT Press.
- Elbourne, Paul. 2008. Demonstratives as individual concepts. *Linguistics and Philosophy* 31. 409–466. DOI: <https://doi.org/10.1007/s10988-008-9043-0>

- Greenman, Joseph. 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>. DOI: <https://doi.org/10.2307/3335446>
- Hanink, Emily. 2018. *Structural sources of anaphora and sameness*. University of Chicago Doctoral dissertation.
- Hanink, Emily & Grove, Julian. 2017. German relative clauses and the severed index hypothesis. In Kaplan, Aaron & McCarvel, Miranda K. & Kaplan, Abby & Rubin, Edward J. (eds.), *Proceedings of the West Coast conference on formal linguistics* 34, 241–248. Somerville, MA: Cascadilla.
- Hawkins, John. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. Routledge 1st edn. DOI: <https://doi.org/10.4324/9781315687919>
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. University of Massachusetts Amherst dissertation. <https://scholarworks.umass.edu/dissertations/AAI8229562>.
- Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13(2). 137–177. DOI: <https://doi.org/10.1007/BF00630732>
- Heim, Irene. 1991. Artikel und Definitheit. In von Stechow, Arnim & Wunderlich, Dieter (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, 487–535. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110255072.996>
- Heim, Irene. 2012. *Definiteness and indefiniteness* 996–1025. Berlin, Boston: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110255072.996>
- 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. DOI: <https://doi.org/10.5281/zenodo.3252018>
- Jenks, Peter. 2015. Two kinds of definites in numeral classifier languages. *Semantics and Linguistic Theory* 25. 103. DOI: <https://doi.org/10.3765/salt.v25i0.3057>
- Jenks, Peter. 2018. Articulated Definiteness without Articles. *Linguistic Inquiry* 49(3). 501–536. DOI: 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. DOI: <https://doi.org/10.1515/9783110867602.1>
- Kratzer, Angelika. 1981. *The notional category of modality*, 38–74. Berlin, Boston: De Gruyter. DOI: <https://doi.org/10.1515/9783110842524-004>
- Kratzer, Angelika. 1989. An investigation of lumps of thought. *Linguistics and Philosophy* 12. 607–653. DOI: <https://doi.org/10.1007/BF00627775>

- Kratzer, Angelika. 2012. *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press. DOI: <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.
- Krifka, Manfred. 2003. Bare NPs: Kind-referring, indefinites, both, or neither? In *Semantics and Linguistic Theory (SALT) XIII*, 180–203. Cornell: CLC Publications. DOI: <https://doi.org/10.3765/salt.v13i0.2880>
- Krifka, Manfred & Pelletier, Francis & Carlson, Gregory & ter Meulen, Alice & Link, Godehard & Chierchia, Gennaro. 1995. Genericity: an introduction. In Carlson, Gregory & Pelletier, Francis (eds.), *The generic book*, 1–125. Chicago: University of Chicago Press.
- Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4(4). 279–326. DOI: <https://doi.org/10.1093/jos/4.4.279>
- Löbner, Sebastian. 2011. Concept Types and Determination. *Journal of Semantics* 28(3). 279–333. DOI: <https://doi.org/10.1093/jos/ffq022>
- Longobardi, Giuseppe. 2001. How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Natural Language Semantics* 9. 335–369. DOI: <https://doi.org/10.1023/A:1014861111123>
- Ortmann, Albert. 2014. Definite article asymmetries and concept types: Semantic and pragmatic uniqueness. In Gamerschlag, Thomas & Gerland, Doris & Osswald, Rainer & Petersen, Wiebke (eds.), *Frames and concept types: Applications in language and philosophy*, vol. 94 (Studies in Linguistics and Philosophy), 293–321. Cham: Springer International Publishing. DOI: 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. DOI: <https://doi.org/10.1023/A:1024157132393>
- Russell, Bertrand. 1905. On denoting. *Mind* 14(56). 479–493. <http://www.jstor.org/stable/2248381>. DOI: <https://doi.org/10.1093/mind/XIV.4.479>
- Sauerland, Uli. 2008. Implicated presuppositions. In Steube, Anita (ed.), *The discourse potential of underspecified structures*, 581–600. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110209303.4.581>
- Schmitt, Cristina & Munn, Alan. 1999. Against the nominal mapping parameter: Bare nouns in Brazilian Portuguese. In Tamanji, Pius N. & Hirotani, Masako & Hall, Nancy (eds.), *NELS 29: Proceedings of the 29th annual meeting of the North East Linguistic Society*, 339–353. Amherst, MA: GLSA Publications.
- Schwarz, Florian. 2009. *Two types of definites in natural language*. University of Massachusetts Amherst dissertation. DOI: <https://doi.org/10.7275/1077454>

- Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559. DOI: <https://doi.org/10.1111/lnc3.12048>
- Schwarz, Florian. 2014. How weak and how definite are Weak Definites. In Aguilar-Guevara, Ana & Bruyn, Bert Le & Zwarts, Joost (eds.), *Weak referentiality*, 213–235. DOI: <https://doi.org/10.1075/la.219.09sch>
- Schwarz, Florian. 2019. Weak vs. strong definite articles: Meaning and form across languages. *Studies in Diversity Linguistics* 25. DOI: <https://doi.org/10.5281/zenodo.3265959>
- Šereikaitė, Milena. 2016. *Definiteness in Lithuanian*. Manuscript, University of Pennsylvania.
- 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. DOI: <https://doi.org/10.5281/zenodo.3252016>
- Šimík, Radek. 2021. Inherent vs. accidental uniqueness in bare and demonstrative nominals. In *Advances in formal Slavic linguistics 2018*, 365–391. Berlin: Language Science Press. DOI: <https://doi.org/10.5281/zenodo.5155544>
- Simonenko, Alexandra. 2014a. *Grammatical ingredients of definiteness*. McGill University dissertation.
- Simonenko, Alexandra. 2014b. Structural triggers of the loss of scopelessness. In *Proceedings of the 32nd West Coast conference on formal linguistics*, 191–200.
- Strawson, Peter Frederick. 1950. On referring. *Mind* 59(235). 320–344. <http://www.jstor.org/stable/2251176>. DOI: <https://doi.org/10.1093/mind/LIX.235.320>
- 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). DOI: <https://doi.org/10.3390/languages6030128>
- Watson, Janet. 2018. South Arabian and Arabic dialects. In Holes, Clive (ed.), *Arabic historical dialectology: Linguistic and sociolinguistic approaches* (Oxford Studies in Diachronic and Historical Linguistics), Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oso/9780198767391.003.0004>
- Wespel, Johannes. 2008. *Descriptions and their domains: The patterns of definiteness marking in French-related creoles*. Stuttgart: University of Stuttgart Doctoral dissertation.
- Wiltschko, Martina. 2012. What does it take to host a (restrictive) relative clause? *Working Papers of the Linguistics Circle of the University of Victoria* 21(2). 100–145.

