

Definiteness in Levantine Arabic Heritage Speakers of English

How Heritage Language Affects Cognition

Letizia Cerqueglini
Tel Aviv University, Israel

Abstract Definiteness depends on crosslinguistic semantic variables, including count/mass distinction, which correlates with cognitive individuation of discrete entities and attention to shape rather than substance. Count/mass distinction is represented by definiteness markers in English but not in Levantine Arabic (LA). Replicating experiments by Liu and Gleason, Middleton, and Lucy and Gaskins, this study tested grammar of definiteness, cognitive individuation, and attention to shape vs. substance in LA heritage speakers of English (LAHSEs, aged 18-25). The results show that LA definiteness parameters affect LAHSEs' cognition but not their grammar of definiteness.

Keywords Definiteness. Count/mass nouns. Heritage language. Language-to-cognition correlation. English. Levantine Arabic.

Summary 1 Background. – 1.1 Definiteness. – 1.2 The Definite Article. – 1.3 English Definiteness and the Definite Article. – 1.4 Arabic Definiteness and the Definite Article. – 1.5 Definiteness of Entities Varies across Languages. – 1.6 System Interference. – 1.7 Cognitive Correlates of Definiteness in English and Levantine Arabic. – 2 Aim of the Study and Research Question. – 3 Preliminary Hypothesis. – 4 Heritage Speakers. – 5 Methodology. – 5.1 Preliminary Grammar Tests. – 5.2 Semantic Similarity Test. – 5.3 Spot-the-Odd-One-Out. – 5.4 Match-by-Similarity. – 6 Results. – 6.1 Preliminary Grammar Test. – 6.2 Semantic Similarity Test. – 6.3 Spot-the-Odd-One-Out. – 6.4 Match-by-Similarity. – 7 Discussion. – 8 Conclusions.

1 Background

1.1 Definiteness

Definiteness, found universally, varies greatly across languages in terms of its semantic foundations and manifestations (Lyons 1999). It has been explored in connection to specificity (von Heusinger 2002), salience (von Heusinger 1997a), uniqueness and inclusiveness (Heim 2011; von Heusinger 1997b; Roberts 2003), familiarity and identifiability (Chen 2004; Heim 1983), ellipsis/reference tracking (Nariyama 2003), and anaphoricity/information triggering (Reinhart 1983). Indeed, definiteness consists of a complex interplay of several strictly correlated semantic and pragmatic features (Zucchi 1995) that operate in different languages to varying degrees (Cho, Slabakova 2014). Among these features is countability, i.e., the distinction between mass and count nouns. Chesterman (2005) theorized definiteness/indefiniteness as linguistically “encoded” and psycholinguistically “decoded” based on a cluster of physical properties such as quantity/inclusiveness, genericity/extensivity, and countability/concreteness, all scalar properties that may vary crosslinguistically. Interestingly, quantity, generality, and concreteness can all be considered semantically proximal within the cognitive bedrock of countability (Strik Liever, Bolognesi, Winter 2021).

1.2 The Definite Article

Languages with articles vary widely in their usage (Hawkins 1978). Some uses of articles have been explored for deixis (Himmelman 1997), anaphora (Bosch 1983), and uniqueness (Hawkins 1978), among other closely interconnected parameters and functions. Among many crosslinguistically detected strategies (Czardybon 2017), the presence or absence of the definite article - in the languages that use it - is a marker associated with the semantic-pragmatic parameter of definiteness (Krámský 1972; von Heusinger 1997b). A striking example of crosslinguistic variation in the use of the definite article is revealed by comparing Arabic (in its numerous varieties) and English, as several studies have demonstrated (Harb 2014; Husni, Newman 2015).

1.3 English Definiteness and the Definite Article

The English article system includes the indefinite article *a(n)*, the definite article *the*, and the zero (null) article. Many have attempted to identify explanations for definite/indefinite noun phrases and the semantic features beyond this distinction (Haspelmath 1999). Abbott (2004) discusses the following semantic properties as related to

English definiteness: uniqueness (Russell 1905), inclusiveness (Hawkins 1978), familiarity (Bolinger 1977), strength (Milsark 1977), and specificity (Haspelmath 1997; Partee 1972). Some of these properties correlate with countability.

Languages either allow or require nouns to appear with an overt indefinite or definite article or allow bare nouns to appear without an article. Danon (2001) observes that the use of the definite article with generics varies crosslinguistically in a way that has no possible effect on interpretation. English allows singular proper nouns and abstract, plural, and mass nouns in argument position with no determiner. In English, singular common, concrete, countable nouns require a determiner (definite article, classifier, number, measure). Mass nouns require the use of measure phrases that contain a classifier to be countable, while count nouns do not (Chierchia 1998). English definite and indefinite singular count nouns, bare plural count nouns, and bare mass nouns can convey genericity, while definite plurals are not allowed to express generic meaning except for names of nationalities. Thus, the semantic-pragmatic feature of entities' countability and its grammatical manifestation in the use of the definite article is active in English. In English, linguistic countability is marked by differential use of definite, indefinite, and no-article options, e.g., 'food is necessary', 'dogs bark', but 'the house is furnished', 'the/a dog barks', to put it simply. English mass nouns in generic sentences do not take the definite article, e.g., 'water is healthy'.

1.4 Arabic Definiteness and the Definite Article

Studies on Arabic definiteness have mainly focused on Classical and Standard varieties, with a few exceptions concerning dialectal data (Dickins 2013; Testen 1998), especially on the grammar of nuna-tion (*tanwīn*) and the definite article 'al- (and its variants, e.g., *il-* in some dialects), considering them definiteness/indefiniteness markers (Holes 1995; Badawi, Carter, Gully 2004), state markers (Lyons 1999; Retsö 2010), or information triggers (Jarrah 2016).

According to the Arab grammatical tradition (Sakaedani 2019; Sartori 2019) and modern scholars (Al-Rawi 2005; Hawas 1986; Jaber 2014), definiteness is not expressed only by 'al 'at-ta'rīf, nor does 'al express only definiteness, e.g., 'ams 'yesterday/'al-'ams 'a day in the past' (Kashgari 2015). Definiteness is also acquired through annexation in constructing state nominals (Shlonsky 2004). 'Al- can be: 1. nominal ('ism mawsūl); 2. definite ('al- 'al-ta'rīf, including 'al 'al-dīhniyya for familiarity, 'al 'al-huduriyya for contextuality, 'al 'al-dīkriyya for anaphoricity, and 'al 'al-jinsiyya for 'non-referential' definiteness) (Abu-Melhim 2013); or 3. augmented 'al-zāyda, attached to demonstrative nouns, time adverbials such as 'now', days of the week,

and singular proper names, which are otherwise generally nunated. The situation is different among dialects in the classic language, as nunation is absent or residual, local strategies other than *'al-/il-* can be prefixed to nouns (e.g., *hal-*; Jarrah 2016), and the obligatory definiteness agreement (Danon 2008) is often violated (as in the *yom ha-šišiy* syndrome; Borg 2000; Pat-El 2009). However, the article system does not correlate with considerations of count/mass oppositions.

Indeed, from a typological point of view, Arabic is a determiner language, i.e., it requires noun phrases in argument position to be preceded by a determiner. In classical, standard forms and most dialects, Arabic only allows definite (singular, plural, and mass) noun phrases to express generic meaning without any difference between well-established and less well-established types and noun-level and sentence-level genericity (Krifka et al. 1995). In Levantine Arabic (LA), the sentence

1. il-ḥalīb jayyid *l-ak*
 DEF-milk good PREP-you
 lit. 'The milk is good for you.'

has a generic meaning. Indeed, Arabic varieties use definite articles with non-count generics, unlike English. Furthermore, across Arabic varieties, the default form of many basic nouns is a grammatically singular mass noun, from which countable forms are obtained through a change in the ending or stem (Bettega, D'Anna 2023). This series of LA examples shows different forms for the concept *bēḏ* 'egg-ness':

2. 'addēš əl-bēḏ?
 how much DET-egg(ness)?
 'How much do eggs cost?'

Bēḏ is a singular grammatically masculine mass noun.

3. kəmm bēḏa?
 how much egg?
 'How many eggs?'

Bēḏa is the countable singulative form, obligatory after the adverbial quantifier *kəmm* and obtained by adding to the basic form *bēḏ* the feminine singular ending *-a*.

4. xams bēḏāt
 five eggs
 'Five eggs.'

Bēḏāt is the countable plural form obtained by adding the feminine plural ending *-āt* to the basic form *bēḏ*.

1.5 Definiteness of Entities Varies across Languages

According to Chesterman (2005), English countability correlates with contiguous properties such as extensiveness, inclusivity, and genericity, all of which surface in grammatical and syntactic definiteness, so that words like ‘lightning’, ‘mankind’, ‘evidence’, and ‘furniture’ are also grammatically processed as mass nouns (Iwasaki, Vinson, Vigliocco 2010). In all languages, some nouns can be counted by numbers, while others need classifiers. Count nouns are perceived as possessing properties that allow them to be counted. Referents of mass nouns are considered not easily countable. Importantly, count/mass properties may vary crosslinguistically in reference to the same entity. In English, apples, biscuits, and sandwiches are usually considered countable, but wine, soup, water, pasta, and corn are not (they need a classifier to be counted, e.g., ‘three bottles of wine’). In LA, most edible entities and foods are mass nouns and require classifiers (*ḥabbeh* for many fruits and grains, *kurrah* for ball-shaped food units such as meatballs, *qit* for candies, *mlaffeh* or *‘ilbeh* for most traditional pastries). English and LA definiteness systems diverge according to different countability parameters that determine different usages of the definite article. The following examples show that the same words (fire, water, bread, dogs, uranium, apples) have different definiteness statuses in English and LA but the same degree of genericity, i.e., they are indefinite in English but definite in LA (Fassi Fehri, Vinet 2008):

5a. When fire starts to burn, it spreads

5b.	lam	btabda	n-nār	tift’il,
	conj	start.PRES.III.F.S.	DEF-fire	burn.SUB.III.F.S.
	inna-ha	btunfur		
	CONJ-SUFF.III.F.S.	spread.PRES.III.F.S.		

6a. Water is good for health

6b.	il-mā	mufid	la-ṣ-ṣiḥḥa
	DET-water	good	PREP-DET-health

7a. Bread sells well every day

7b.	il-xubz	byitbī	kaṭīr	kull	yōm
	DET-bread	sell.PRES.III.M.S.	well	every	day

8a. Dogs bark

8b.	il-klāb	byinbaḥu
	DET-dog.PL	bark.PRES.III.PL

- 9a. Uranium is a heavy element
9b. il-yurānyum ʿunṣur tqīl
 DET.uranium element heavy
- 10a. Apples are too expensive
10b. it-tiffāḥ ktīr ḡāliy
 DET-apple much expensive

1.6 System Interference

English and LA definiteness systems interfere with Arabic native speakers' acquisition of English as a second language (Harb 2014; Husni, Newman 2015). There is evidence of the effects of Arabic countability properties in their errors in the target language.¹ Arabic learners of English overuse the definite article in idioms, with abstract and uncountable nouns, and in generic plural noun phrases, e.g., *The value of the time; *He sells the apples at the crossroad; *The milk is nutritious to the body; *I went to the bed; *You cook the rice; *The horses are useful animals. The fact that Arabic-speaking English learners find it difficult to decide whether referents are countable (Butler 2002; Master 1987) supports the hypothesis that the count/mass opposition is language-specific and non-conceptual, i.e., to some extent, arbitrary. Therefore, it is worth observing if and how the cognitive structures involved co-vary crosslinguistically with the linguistic structures. In addition to crosslinguistic comparisons, relevant case studies are those in which different linguistic systems coexist in the same subjects, such as multilingualism, heritage languages, linguistic impairments, and language acquisition.

1.7 Cognitive Correlates of Definiteness in English and Levantine Arabic

The correspondence between the grammatical expression of countability through definiteness and conceptual properties, for example, the individuation of discrete bounded entities vs. non-individuation, has been proposed by Du Bois (1980), Gundel, Hedberg and Zacharski (1993); Koga (1992), Wierzbicka (1988), and Wisniewski, Lamb and Middleton (2003). Furthermore, cognitive tests conducted on native speakers of English have proved the relationship between linguistic countability and the cognitive individuation of discrete bounded entities (Middleton et al. 2004). Lucy and Gaskins (2001) have

¹ Aboras 2020; Alenizi 2013; 2017; Al-Malki, Norazmani, Noor 2014; Naim-Bader 1988.

demonstrated that the presence of count/mass distinction correlates crosslinguistically with attention to the shape rather than the substance of entities. Indeed, they found that English speakers associate the unit of individuation with count nouns and, as a result, classify entities based on their shapes, which are the best indicators of individuated entities. In contrast, speakers of Yucatec Maya (an indigenous language spoken in southeastern Mexico) usually pay attention to the material composition of entities rather than their shapes. Thus, while native speakers of English categorize objects by shape rather than material, native speakers of Mayan languages – where the count/mass distinction is not embedded in grammar but realized through classifiers – tested positive for the opposite cognitive tendency.

English spoken by native speakers thus displays a strong correlation between the semantic-pragmatic property of countability and its grammatical manifestation through definiteness and the tendency toward individuation of discrete bounded entities and attention to shape rather than substance in cognition. Assuming a correlation between language and cognition exists, based on the linguistic data presented so far, one could hypothesize that LA will cognitively behave similarly to Yucatec Maya in the experiment of Lucy and Gaskins (2001) and produce results opposite to those produced by English in the experiment designed by Middleton et al. (2004).

Indeed, the two experiments had not previously been conducted on LA. Fifteen LA native monolingual adult speakers were tested in both experiments to provide the necessary background information for the present research. The hypothesis articulated above was confirmed. LA speakers do not manifest any particular cognitive bias toward the individuation of discrete bounded entities, and they show clear attention to substance rather than shape, consistently with LA's lack of correlation between countability, definiteness, and the use of the definite article at the linguistic level.

A more detailed exposition of the LA data is unnecessary here, as it is not the subject of the present research and will be treated separately in a future study. For the present study, the results provided by the replications of the studies of Lucy and Gaskins (2001) and Middleton et al. (2004) for English and LA speakers represent control data.

2 Aim of the Study and Research Question

Due to the extreme differences between English and LA linguistic and cognitive data mentioned above, linguistic and cognitive responses of LA heritage speakers of English (LAHSEs) are particularly interesting here. Linguistically, LAHSEs' definiteness system and use of the definite article should reflect complete mastery of the English rules of mass/count distinction without influences from LA, which the LAHSE

informants selected for this study ceased to acquire at an early age (three to four) so that their experience with the language remained limited to partial comprehension and the use of some brief/routine speech productions (greetings, thanking formulas, expressions of affection or disappointment, some nouns). Assuming there is a strong correlation between linguistic definiteness and countability, on the one hand, and cognitive bias toward the individuation of discrete entities and preference for the shape over the substance of objects on the other, LAHSEs' cognitive responses should reflect the data elicited from English speakers.

If cognition does not necessarily reflect only the constraints of the dominant language at a certain moment in one's lifetime and can store different kinds of information, experiences, and endowments, some cognitive similarities may exist between LA speakers and LAHSEs.

To test the presence of cognitive similarities between LA and LAHSEs, 15 adult LAHSEs aged 18 to 25 were tested in the experiments designed by Lucy and Gaskins (2001) and Middleton et al. (2004). The identification of cognitive similarities between LA and LAHSEs would demonstrate that despite the late development of definiteness in children (Liu, Gleason 2002), cognitive features that correlate with it are established by age three or four, the age at which LA heritage language acquisition ceased among the informants of this study.

3 Preliminary Hypothesis

I expected LAHSEs to display complete mastery of the English definiteness markers in tasks that entailed different countability values in the linguistic experiments and produce the same results as the English (EN) native speakers' control group. On the other hand, I hypothesized that LAHSEs' cognitive responses would be similar to those of the EN control group, yet I could not exclude the possibility of some similarities with the LA control group. In particular, following preliminary spontaneous observations, I expected LAHSEs to classify known and novel objects by material, not shape. I anticipated that heritage languages would leave traces in the cognition of heritage speakers, as was observed for semantic structures by Polinsky (2011) and Scontras et al. (2017). It seems, therefore, unnecessary to actively speak a language for it to influence one's cognitive structures. This idea is in line with the notion of complex cognition, in which different structures not necessarily related to communicative tasks and proficiency can coexist (Slobin 2014).

4 Heritage Speakers

The term ‘heritage speakers’ refers to descendants of immigrants who inhabit a bilingual environment from an early age. Their dominant language is the host country’s language, but some aspects of the family language (the heritage language) may still affect their linguistic abilities from the periphery of their linguistic consciousnesses. Heritage languages are often accompanied by stories and individual paths of migration, displacement, and minority communities and often remain marginal in the societies in which heritage speakers live and in their communities and daily lives (Pavlou, Grohmann 2021). Importantly, heritage languages can be acquired at different proficiency levels and for various reasons (Pavlou, Fotiou 2022). Indeed, heritage speakers vary widely in the degree of their receptive and productive command of the heritage language. This study considers a particular type of heritage speakers defined by Polinsky as “over-hearers” (2018). These heritage speakers have limited situational competence in their heritage language that is restricted to a more or less extensive understanding. The LAHSEs selected for this research were born to LA-speaking parents in England. They are all over-hearers whose families deliberately deprived them of exposure to Arabic linguistic stimulus very early due to the urgency of integrating into English-speaking society. LA was heard only at gatherings of family and friends, early on becoming less important than English, which was considered necessary for education and perceived as an instrument of social integration and advancement. None of the LAHSEs tested here were proficient in LA, and all possessed only oral comprehension abilities and basic communicative competence (beginner level). I chose LAHSEs to investigate whether and how passive proficiency in a given language can still affect cognition.

5 Methodology

All groups (LAHSEs, EN speakers, LA speakers) were tested using the linguistic and cognitive tests described here. Fifteen LAHSEs aged 18 to 25, born and raised in England, participated in linguistic and cognitive experiments. Fifteen monolingual native LA speakers and fifteen EN native speakers aged 18 to 25 represented the control groups and participated in the same experiments. As for the linguistic experiment, EN speakers and LAHSEs were requested to reply in English, and monolingual LA speakers in LA.

Due to the effort of recruiting a sufficient number of heritage Arabic speakers of English, especially of the ‘over-hearer’ type, for the experiments described here, I selected the informants who participated in this study, to whom I am deeply indebted, from sedentary urban

Arabic-speaking Palestinian and Syrian communities. Therefore, I use the label Levantine Arabic (ISO 639-3), also used in traditional dialectological classifications to define the languages spoken in the predominantly urban and coastal area that includes Syria, Lebanon, Palestine, and Jordan (Al-Jallad 2012; Al-Masri 2015; McLoughlin 2009).

5.1 Preliminary Grammar Tests

Grammar tests consisted of a fill-in-the-blank task, an error correction task, countability judgments of nouns in isolation, and countability judgments of nouns in context. Each test included 20 entries, all elaborated *ad hoc* based on Liu and Gleason's model (2002). These entailed countability-based oppositions correlated with abstractness, genericity (Behrens 2000; Dahl 1975), extension, and inclusiveness (Carlson 1977; Fiengo 1987) under different conditions of numbers, tenses, and argument structure. Sentences were never longer than one line. The EN control group and LAHSE informants underwent the test in English, while the LA control group underwent the same test in LA. LA informants were tested as a control group to allow me to judge LA influences in LAHSEs' performances. Informants had thirty minutes to complete this task.

5.2 Semantic Similarity Test

A semantic similarity test was employed to determine whether the count/mass distinction had consequences for semantic representation in that words that share count or mass status are more semantically similar than words that do not (Iwasaki, Vinson, Vigliocco 2010). The test was based on the 'error induction design', counting and analyzing the mass/count cross-category 'substitution errors'. Semantically related lexical substitution errors (e.g., 'beer' substituted by 'wine') are quite common due to the co-activation of semantically related lexical candidates during a conceptually driven retrieval process (Garrett 1992; Levelt, Roelofs, Meyer 1999). Grammatical classes also appear to play a role in errors of this type (Fromkin 1973). If English speakers' semantic representations are affected by count/mass distinction, their substitution errors are expected to include more cases in which target and error words share count/mass status than those made by LA speakers. I tested words from the domain of food because this domain offers significant within-category variability regarding the count/mass status of picturable items. In this field, there are both solid and non-solid entities, and within substantial items, there are differing degrees of ease of individuation (e.g., 'apple' may be more easily individuated than 'celery') that can overlap

with count/mass status. Finally, Bloom (1994) found that English-speaking children tend to make errors involving count/mass status in the semantic field of foods ('eating a bacon', 'I drop a celery'), suggesting that in this domain, the link between conceptual and grammatical properties may be more arbitrary than in others. If English count/mass status has come to affect English speakers' semantic knowledge as a consequence of the relationship between conceptual properties (object vs. substance or individual vs. portion) and grammatical properties (count vs. mass), we would be expected to observe language-specific effects on semantic substitution errors such that errors in which names of stimulus and target share the count/mass status would be more common for English speakers than for Arabic speakers, for whom this distinction is not grammatically marked.

Speakers were first asked to name, in their native language, items depicted in 30 high-resolution color pictures shown on a computer screen using either a count phrase (e.g., English 'a __'; LA ' __ *wāḥid/a* 'one') or a mass phrase (e.g., English 'some __'; LA ' __') to check their agreement on the property attributed to each entity. Next, I grouped 21 pictures of the food items on whose mass/count status speakers of both groups agreed the most into blocks of seven pictures. Each picture appeared 14 times during the experiment, which included 42 blocks and lasted 30 to 40 minutes. Speakers were asked to name food pictures aloud in their native languages using single words (or a name such as 'green bean') as they appeared on the computer screen.

The experiment began with a name agreement phase in which participants were asked to name each experimental picture without time pressure. I noted any variation from the intended names and provided prompts if the participants could not produce a label for a particular picture. Next, participants performed a set of 42 practice trials on 42 blocks, each containing seven pictures in a row. Each target picture could appear once in each block and could not appear as the last item of one block and the first item of the next. In the practice trials, the blocks were presented in one of four possible locations on the screen, and the participant was instructed to name each aloud. After each block was presented, I altered the presentation rate to accommodate each speaker's speech rate, adjusting it by 100 milliseconds more or less, if necessary, to make the task challenging yet manageable for each speaker. The presentation rate was speeded up if a participant successfully named all pictures without errors and slowed down if a participant could not keep up with the presentation.

I analyzed the lexical errors (i.e., cases in which the word produced for a target was a different word than the one I expected). EN speakers', LA speakers', and LAHSEs' errors were analyzed according to the proportion of errors that preserved the count/mass status of the target label.

5.3 Spot-the-Odd-One-Out

In this experiment, speakers were asked to make semantic judgments on 12 triads of words (translation equivalent in the two languages). Their task was to spot the odd one out and cross out the word less similar to the other two in terms of meaning (Garrard et al. 2004). I hypothesized that if count/mass status affects English speakers' semantic representations, EN speakers and LAHSEs should show a greater tendency than LA speakers to select words that share count/mass. Twelve words were selected and combined in all possible triadic combinations, and the order of the three words in each triad was randomized. Participants completed the task using paper and pencil. The time assigned to complete the task was 90 seconds.

5.4 Match-by-Similarity

A non-linguistic experiment by Lucy and Gaskins (2001) was replicated. It consisted of asking the informants to observe an original object and decide which of two alternative objects was more similar to it. One had the same shape as the original object, while the other had the same material composition. Each informant underwent six tests, four with known objects and two with novel objects. The time allocated was 18 seconds.

According to the hypothesis that linguistic properties of countability affect cognition, EN speakers were expected to prefer the shape alternative and LA speakers the material alternative. The choices of the LAHSEs were the objects of the experimental question.

6 Results

6.1 Preliminary Grammar Test

The results of the EN speakers' and LAHSEs' grammar tests confirm that count/mass noun judgments strongly correlate with competence in definiteness rules in English. The grammatical tests are not discussed here since the linguistic count/mass nominal parameters of LAHSEs are similar to the average outcomes produced by EN speakers (Liu, Gleason 2002) in all respects. Indeed, LAHSEs demonstrated native competence in English.

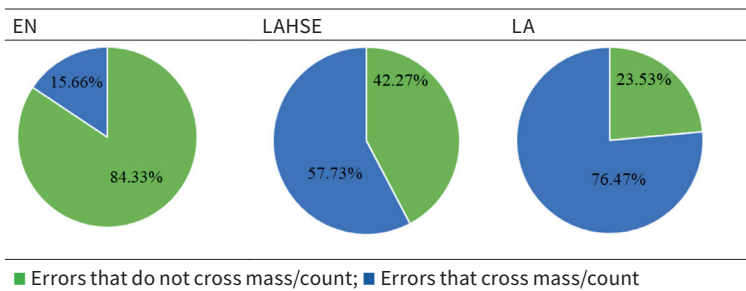
6.2 Semantic Similarity Test

In the semantic similarity test, all groups produced more naming errors under increasing time stress. Each group had an average of 4,400 valid responses. All entities (mainly liquid, edible, processed, or raw) were presented in bowls so that shape would not interfere with categorization. Errors made by the EN control group (83 errors; 1.88%) did not involve the count/mass distinction (70 of 83 errors; 84.33%), so nouns were mistakenly attributed within the same category (mass: ‘water’ for ‘juice’, ‘rice’ for ‘corn’, ‘flour’ for ‘sugar’, ‘oil’ for ‘honey’; count: ‘biscuits’ for ‘candies’, ‘chocolates’ for ‘meatballs’, ‘pastries’ for ‘meat rolls’). Only 13 of 83 errors (15.66%) violated the mass/count boundary in the EN group. Interestingly, a small percentage of errors produced by EN speakers also involved shape-related boundaries (e.g., ‘oranges’ for ‘meatballs’; 5 of 83; 6.02%).

Among the LA speakers, 2.31% of the responses were incorrect. In line with the hypothesis, the errors produced by LA informants often crossed the count/mass distinction (78 of 102 errors; 76.47%). Errors included, for example, ‘rice’ (*ruz*/mass) for ‘sugared almonds’ (*ḥalawiyāt*/count), ‘meat-and-rice balls’ (*kafta*/mass) for ‘biscuits’ (*baskwit*/count), and ‘candies’ (*ḥulwa*/mass) for ‘pastries’ (*mu’janeh*/count).

Among the LAHSEs, 2.2% of the responses were incorrect. Notably, 56 of the 97 errors (57.73%) produced by LAHSE informants, mainly those related to nouns representing processed food types, crossed the count/mass boundary, in line with the LA results. Figure 1 summarizes the results of the semantic similarity test.

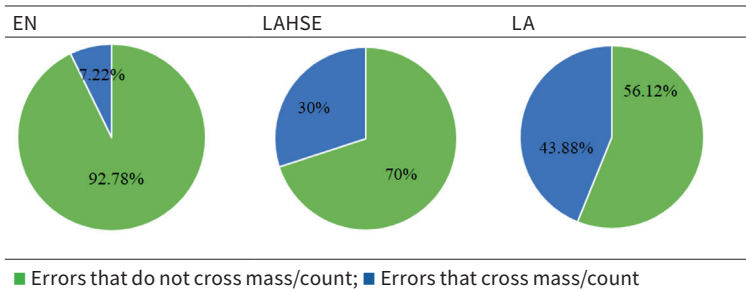
Figure 1 Results of the semantic similarity test



6.3 Spot-the-Odd-One-Out

The test employed 180 triads. Of these, EN speakers violated the count/mass boundary in 13 cases (7.22%), LA speakers in 79 cases (43.88%), and LAHSEs in 54 cases (30%). Figure 2 reports the results of the spot-the-odd-one-out test.

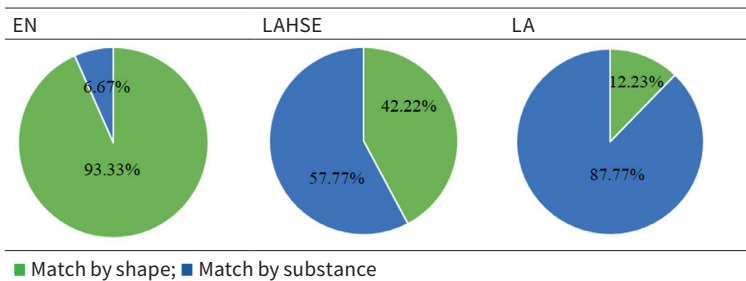
Figure 2 Results of the spot-the-odd-one-out test



6.4 Match-by-Similarity

The cognitive test confirmed the data yielded by previous experiments conducted by Lucy and Gaskins (2001) on EN speakers. Indeed, this group opted mainly for matching objects with the same shape (84 of 90 responses; 93.33%). LA speakers were more oriented toward matching objects of the same substance (79 of 90 responses; 87.77%). LAHSE informants produced an intermediate result: 52 of 90 responses matched objects by material (57.77%) and 38 by shape (42.22%). Figure 3 reports the results of the match-by-similarity test.

Figure 3 Results of the match-by-similarity test



7 Discussion

Except in the case of the preliminary grammatical test, in which the expectations of similarity between the results of the EN and LAHSE speakers were satisfied, the other experiments showed a marked closeness of the LAHSE results with the LA ones. Regarding the match-by-similarity cognitive test, a remarkable discrepancy emerged between linguistic and cognitive representations. This discrepancy could be due to various cultural and environmental factors, yet it is not very surprising in light of the most recent studies on the complexity of the relationships between language and cognition in different domains of experience. The alignment between linguistic and cognitive representations is a recently dispelled myth (Bohnenmeyer et al. 2022). The relation between linguistic and cognitive structures is a complex phenomenon that depends on many factors, such as the domain (more or less dependent on sensory experience) and the language in question and its transmission, which is connected to the conditions of the material and intellectual culture in which speakers are immersed (Cerqueglini 2022).

What is perhaps most striking is the discrepancy between the LAHSE results of the grammatical test, which converge with the EN results, and the LAHSE semantic representations of mass/count, which lie between the EN and LA results, as shown by the semantic similarity test and the spot-the-odd-one-out test. The results of the semantic similarity and spot-the-odd-one-out tests may depend on cultural factors that interfere with linguistic choices in LAHSEs. Indeed, LAHSE informants live in an English linguistic environment, yet food is part of the daily domestic routine and is prepared, measured, served, and discussed according to inherent LA cultural criteria, influenced by mass concepts and related quantifiers and classifiers.

Nonetheless, the failure to notice the count/mass distinction did not occur only in the case of food. In the spot-the-odd-one-out triads, for example, given the trio showing water/rice/biscuits, 12 EN speakers spotted the water (the liquid), while 14 LAHSE informants spotted the biscuits (the only count noun). In the English triplet ‘parquet’ (mass)/‘tile’ (count)/‘brick’ (count), 12 EN speakers pointed to ‘parquet’, the only mass noun, while LAHSEs were much less count/mass-oriented (four pointed to ‘parquet’, six to ‘tile’, and five to ‘brick’). In the corresponding LA triplet *barkē* (mass, ‘parquet’)/*balāṭa* (count, ‘tile’)/*qarmīd* (mass, ‘brick’), LA speakers did not show specific effects of count/mass distinctions. Similar results were obtained for the triplet ‘soap/*ṣabūn*’ (mass)/‘shampoo/*šambū*’ (mass)/‘sponge/*sfinjeh*’ (count). EN speakers generally pointed to ‘sponge’, while LAHSEs and LA speakers made different choices, not oriented by count/mass bias.

8 Conclusions

This article presents a study of the alignment between the grammatical structures of definiteness and semantic and cognitive representations of countability and individuation. Semantic and cognitive similarities between LAHSEs and LA speakers are striking, considering that LAHSEs speak only English fluently, and their competence in LA is only passive and restricted to a scanty vocabulary and set of communicative tasks. LA definiteness grammar, which does not mark count/mass distinctions, does not affect LAHSEs' grammatical proficiency in English definiteness grammar. Probably because the article system is among the most frequently recurring elements in English, its rules are deeply embedded in the grammatical proficiency of its speakers. Being fully competent in English definiteness grammar as native speakers of English, LAHSE informants were expected to have strong biases toward the semantic count/mass opposition and cognitive individuation by shape. Nonetheless, the lability of the semantic count/mass opposition among LAHSEs echoes the LA semantic profile, and LAHSEs' cognitive bias toward matching objects by shape is in line with LA speakers' cognitive decisions.

Regarding the relationship between language and cognition, the case of LAHSE speakers demonstrates that cognition and language are not expressions of the same underlying structures; vast experiential categories are stored in cognition yet are often silent or recessive in language. Thus, a 'thinking for speaking' activity does exist (Slobin 1992) but represents a small part of the whole cognitive potential of an individual. In other words, language is not the only factor that affects cognition. LAHSE informants' experience demonstrates that mental habits and attitudes that lead to specific judgments, evaluations, and decisions are also transmitted via cultural practices. Preparing food in specific quantities, portions, and shapes and serving and consuming it in specific containers and with certain utensils can affect cognition as much as the language in which we think.

Furthermore, in terms of linguistic acquisition studies, this study raises an interesting question. Although definiteness is acquired at a later age than other grammatical competencies, it seems that semantic and cognitive parameters related to it are ready to use at a very early age (three to four), when LAHSEs' acquisition of LA structures begins its decline in favor of English. This suggestive scenario remains open for future studies to explore.

Bibliography

- Abbott, B. (2004). "Definiteness and Indefiniteness". Horn, L.; Gregory, W. (eds), *The Handbook of Pragmatics*. Malden: Blackwell, 122-49.
- Aboras, A. (2020). "The Acquisition of Definiteness and Specificity in English: A Case Study with Saudi-Arabic Learners of English Using an Online Task". *English Language and Literature Studies*, 10(1), 13-25.
- Abu-Melhim, A. (2013). "The Presence or Absence of Definiteness in the Native Language and Its Impact on ESL Learning". *English Language Teaching*, 7(1), 50-65.
- Alenizi, A. (2013). *The Acquisition of the Nongeneric Uses of the English Definite Article the by Arabic Speakers of English* [MA Thesis]. Carbondale: Southern Illinois University.
- Alenizi, A. (2017). *Noun Countability Judgments by Arabic Speakers of English* [PhD Dissertation]. Ball, IN: Ball State University.
- Al-Jallad, A. (2012). *Ancient Levantine Arabic: A Reconstruction Based on the Earliest Sources and the Modern Dialects* [PhD Dissertation]. Cambridge, MA: Harvard University Press.
- Al-Malki, E.; Norazman, A.; Noor, A. (2014). "Generic Reference in English, Arabic and Malay: A Cross Linguistic Typology and Comparison". *English Language Teaching*, 7(11), 15-27.
- Al-Masri, M. (2015). *Colloquial Arabic (Levantine): The Complete Course for Beginners*. London: Routledge.
- Al-Rawi, M. (2005). *The Structure of Determiner Phrases in Standard Arabic* [PhD Dissertation]. Leeds: University of Leeds.
- Badawi, E.; Carter, M.; Gully, A. (2004). *Modern Written Arabic: A Comprehensive Grammar*. New York: Routledge.
- Behrens, L. (2000). *Typological Parameters of Genericity*. Köln: Institut für Sprachwissenschaft, Universität zu Köln, 1-110. Arbeitspapier 37.
- Bettega, S.; D'Anna, L. (2023). *Gender and Number Agreement in Arabic*. Leiden: Brill.
- Bohnenmeyer, J. et al. (2022). "Reference Frames in Language and Cognition: Cross-Population Mismatches". *Linguistics Vanguard*, 8(1), 175-89.
- Bloom, P. (1994). "Syntax-Semantics Mappings as an Explanation for Some Transitions in Language Development". Levy, Y. (ed.), *Other Children, Other Languages: Issues in the Theory of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 41-75.
- Bolinger, D. (1977). *Meaning and Form*. New York; London: Longman.
- Borg, A. (2000). "Some Observations on the *שִׁי יוֹם הַמֵּת* Syndrome on the Hebrew of the Dead Sea Scrolls". Muraoka, T.; John, E. (eds), *Diggers at the Well: Proceedings of a Third International Symposium on the Hebrew of the Dead Sea Scrolls and Ben Sira*. Leiden: Brill, 26-39.
- Bosch, P. (1983). *Agreement and Anaphora*. New York: Academic Press.
- Butler, Y. (2002). "Second Language Learners' Theories on the Use of English Articles". *Studies in Second Language Acquisition*, 24, 451-80.
- Carlson, G. (1977). *Reference to Kinds in English* [PhD Dissertation]. Amherst: University of Massachusetts.
- Cerqueglini, L. (2022). *Space and Time in aṣ-Ṣāni' Arabic. A Cross-Generational Study*. Leiden: Brill.
- Chen, P. (2004). "Identifiability and Definiteness in Chinese". *Linguistics*, 42(6), 1129-84.

- Chesterman, A. (2005). *On Definiteness: A Study with Special Reference to English and Finnish*. Cambridge: Cambridge University Press.
- Chierchia, G. (1998). "Plurality of Mass Nouns and the Notion of 'Semantic Parameter'". *Events and Grammar*, 70, 53-103.
- Cho, J.; Slabakova, R. (2014). "Interpreting Definiteness in a Second Language without Articles: The Case of L2 Russian". *Second Language Research*, 30(2), 159-90.
- Czardybon, A. (2017). *Definiteness in a Language without Articles. A Study on Polish*. Düsseldorf: Düsseldorf University Press.
- Dahl, Ö. (1975). "On Generics". Keenan, E. (ed.), *Formal Semantics of Natural Language*. London; New York: Cambridge University Press, 99-111.
- Danon, G. (2001). "Syntactic Definiteness in the Grammar of Modern Hebrew". *Linguistics*, 39(6), 1071-116.
- Danon, G. (2008). "Definiteness Spreading in the Hebrew Construct State". *Lingua*, 118(7), 872-906.
- Dickins, J. (2013). "Definiteness, Genitives and Two Kinds of Syntax in Standard Arabic". Kutay, R.; Seeger, U.; Talay, S. (eds), *Nicht nur mit Engelzungen: Beiträge zur semitischen Dialektologie. Festschrift für Werner Arnold zum 60. Geburtstag*. Wiesbaden: Harrassowitz, 59-72.
- Du Bois, J. (1980). "Beyond Definiteness: The Trace of Identity in Discourse". Chafe, W. (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex, 203-74.
- Fassi Fehri, A.; Vinet, M. (2008). "Nominal and Verbal Classes in Arabic and Chinese". *Revue linguistique de Vincennes*, 37, 55-83.
- Fiengo, R. (1987). "Definiteness, Specificity, and Familiarity". *Linguistic Inquiry*, 18, 163-6.
- Fromkin, V. (1973). *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
- Garrard, P. et al. (2004). "Dissociation of Lexical Syntax and Semantics: Evidence from Focal Cortical Degeneration". *Neurocase*, 10, 353-62.
- Garrett, M. (1992). "Lexical Retrieval Processes: Semantic Field Effects". Lehrer, A.; Kittay, E.; Lehrer, R. (eds), *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization*. Hillsdale, NJ: Lawrence Erlbaum Associates, 377-95.
- Gundel, J.; Hedberg, N.; Zacharski, R. (1993). "Cognitive Status and the Form of Referring Expressions in Discourse". *Language*, 69, 274-307.
- Harb, M. (2014). "A Closer Look at the English Article System". *International Journal of Linguistics*, 6, 87-101.
- Haspelmath, M. (1997). *Indefinite Pronouns*. Oxford: Oxford University Press.
- Haspelmath, M. (1999). "Explaining Article-Possessor Complementarity: Economic Motivation in Noun Phrase Syntax". *Language*, 75(2), 227-43.
- Hawas, H. (1986). *The Realization of Definiteness in English and Arabic: A Contrastive/Error Analysis Study* [PhD Dissertation]. Cardiff: University of Wales.
- Hawkins, J. (1978). *Definiteness and Indefiniteness*. San Diego: Academic Press.
- Heim, I. (1983). "File Change Semantics and the Familiarity Theory of Definiteness". Bäuerle, R.; Schwarze, C.; von Stechow, A. (eds), *Meaning, Use, and the Interpretation of Language*. Berlin: de Gruyter, 164-89.
- Heim, I. (2011). "Definiteness and indefiniteness". Von Heusinger, K.; Maienborn, C.; Portner, P. (eds), *Semantics. An International Handbook of Natural Language Meaning*. Vol. 2. Berlin: de Gruyter, 996-1025.
- Himmelmann, N. (1997). *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.
- Holes, C. (1995). *Modern Arabic. Structures, Functions and Varieties*. London-New York: Longmans.

- Husni, R.; Newman, D. (2015). *Arabic-English-Arabic translation: Issues and Strategies*. Abingdon-on-Thames: Routledge.
- Iwasaki, N.; Vinson, D.; Vigliocco, G. (2010). "Does the Grammatical Count/Mass Distinction Affect Semantic Representations? Evidence from Experiments in English and Japanese". *Language and Cognitive Processes*, 25(2), 189-223.
- Jaber, A. (2014). *On Genericity and Definiteness in Modern Standard Arabic* [PhD Dissertation]. Chapel Hill: University of North Carolina.
- Jarrah, M. (2016). "On Definiteness and Information Trigger in Arabic". *Advances in Language and Literary Studies*, 7(2), 55-67.
- Kashgary, A. (2015). "Meanings and Functions of the Arabic Article 'al-' 'the' with Reference to the Concept of Definiteness". *Journal of Semitic Studies*, 60(1), 93-110.
- Koga, K. (1992). "A Cognitive Approach to Definiteness: From Specific to Generic". *English Linguistics*, 9, 152-75.
- Krámský, J. (1972). *The Article and the Concept of Definiteness in Language*. The Hague: Mouton.
- Krifka, M. et al. (1995). *Genericity: An Introduction*. Gregory, C.; Pelletier, F. (eds), *The Generic Book*. Chicago: University of Chicago Press, 1-124.
- Levelt, W.; Roelofs, A.; Meyer, A. (1999). "A Theory of Lexical Access in Speech Production". *Behavioral and Brain Sciences*, 22, 1-75.
- Liu, D.; Gleason, J. (2002). "Acquisition of the Article 'the' by Nonnative Speakers of English. An Analysis of Four Nongeneric Uses". *Studies in Second Language Acquisition*, 24, 1-26.
- Lucy, J.; Gaskins, S. (2001). "Grammatical Categories and the Development of Classification Preferences: A Comparative Approach". Bowerman, M.; Levinson, S. (eds), *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 257-83.
- Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Master, P. (1987). *A Cross-Linguistic Interlanguage Analysis of the Acquisition of the English Article System* [PhD Dissertation]. Los Angeles: University of California.
- McLoughlin, L. (2009). *Colloquial Arabic (Levantine): Syria, Lebanon, Palestine, Jordan*. 2nd edition. London: Routledge.
- Middleton, E. et al. (2004). "Separating the Chaff from the Oats: Evidence for a Conceptual Distinction between Count Noun and Mass Noun Aggregates". *Journal of Memory and Language*, 50, 371-94.
- Milsark, G. (1977). "Toward an Explanation of Certain Peculiarities of the Existential Construction in English". *Linguistic Analysis*, 3, 1-29.
- Naim-Bader, F. (1988). *Analysis of Error Type, Source, and Gravity in the Writing of Arabic ESL students in USA*. Colleges [MA Thesis]. Portland, OR: Portland State University.
- Nariyama, S. (2003). *Ellipsis and Reference Tracking in Japanese*. Amsterdam: John Benjamins Publishing.
- Partee, B. (1972). "Opacity, Coreference, and Pronouns". Davidson, D.; Harman, G. (eds), *Semantics of Natural Language*. Dordrecht: Reidel, 415-41.
- Pat-El, N. (2009). "The Development of the Semitic Definite Article: A Syntactic Approach". *Journal of Semitic Studies*, 54, 19-50.
- Pavlou, N.; Fotiou, C. (2022). "The Role of Heritage and (Potential) New Speakers' Motivations to Learning a Language: The Case of Sanna". Georgiou, G. (ed.), *Mind and Second Language Acquisition: Experimental Approaches*. Cambridge: Cambridge University Press.

- Pavlou, N.; Grohmann, K. (2021). "Heritage Languages in Southeastern Europe". Montrul, S.; Polinsky, M. (eds), *The Cambridge Handbook of Heritage Languages and Linguistics*. Cambridge: Cambridge University Press, 69-90.
- Polinsky, M. (2011). "Reanalysis in Adult Heritage Language: A case for Attrition". *Studies in Second Language Acquisition*, 33, 305-28. <https://doi.org/10.1017/S027226311000077X>.
- Polinsky, M. (2018). *Heritage Languages and Their Speakers*. Cambridge: Cambridge University Press.
- Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Abingdon-on-Thames: Taylor and Francis.
- Retsö, J. (2010). "Status-eine vernachlässigte Kategorie der arabischen Grammatik". Talay, S.; Bobzin, H. (Hrsgg), *Grammatik, Dichtung und Dialekte. Beiträge einer Tagung in Erlangen zu Ehren von Wolfdietrich Fischer*. Wiesbaden: Reichert, 65-71.
- Roberts, C. (2003). "Uniqueness in Definite Noun Phrases". *Linguistics and Philosophy*, 26(3), 287-350.
- Russell, B. (1905). "On Denoting". *Mind*, 14, 479-93.
- Sartori, M. (2019). "Definition and Determination in Medieval Arabic Grammatical Thought". Gioffo, M.; Versteegh, K. (eds), *The Foundations of Arabic Linguistics IV*. Leiden: Brill, 253-72.
- Sakaedani, H. (2019). "What is Definiteness in Arabic? Focusing on Proper Nouns for Genera and *asmā muḥama* 'Ambiguous Nouns'". Gioffo, M.; Versteegh, K. (eds), *The Foundations of Arabic Linguistics IV*. Leiden: Brill, 233-52.
- Scontras, G. et al. (2017). "Cross-Linguistic Scope Ambiguity When Two Systems Meet". *Glossa: A Journal of General Linguistics*, 2(1), 1-28. <https://doi.org/10.5334/gjgl.198>.
- Shlonsky, U. (2004). "The Form of Semitic Noun Phrases". *Lingua*, 114(12), 1465-526.
- Slobin, D. (1992). *The Crosslinguistic Study of Language Acquisition*. Hillsdale: Erlbaum.
- Slobin, D. (2014). "Thinking for Speaking". *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 13, 435-45.
- Strik Lievers, F.; Bolognesi, M.; Winter, B. (2021). "The Linguistic Dimensions of Concrete and Abstract Concepts: Lexical Category, Morphological Structure, Countability, and Etymology". *Cognitive Linguistics*, 32(4), 641-70.
- Testen, D. (1998). "Literary Arabic and Early Hijazi Contrasts in the Marking of Definiteness". Haeri, N.; Eid, M.; Benmamoun, E. (eds), *Perspectives on Arabic Linguistics XI: Papers from the Eleventh Annual Symposium on Arabic Linguistics*. Amsterdam: John Benjamins, 207-25.
- von Heusinger, K. (1997a). "Salience and Definiteness". *Prague Bulletin of Mathematical Linguistics*, 67, 5-23.
- von Heusinger, K. (1997b). "Definite Descriptions and Choice Functions". Aka-ma, S. (ed.), *Logic, Language and Computation*. Springer: Dordrecht, 61-92. https://doi.org/10.1007/978-94-011-5638-7_4.
- von Heusinger, K. (2002). "Specificity and Definiteness in Sentence and Discourse Structure". *Journal of Semantics*, 19(3), 245-74.
- Wierzbicka, A. (1988). *The Semantics of Grammar*. Amsterdam: John Benjamin.
- Wisniewski, E.; Lamb, C.; Middleton, E. (2003). "On the Conceptual Basis for the Count and Mass Noun Distinction". *Language and Cognitive Processes*, 18, 583-624.
- Zucchi, A. (1995). "The Ingredients of Definiteness and the Definiteness Effect". *Natural Language Semantics*, 3(1), 33-78.