Behavior Research Methods (2018) 50:2408–2425
https://doi.org/10.3758/s13428-018-1019-6

Gulf Arabic nouns and verbs: A standardized set of 319 object pictures and 141 action pictures, with predictors of naming latencies

Tariq Khwaileh[1] · Eiman Mustafawi[1] · Ruth Herbert[2] · David Howard[3]

Published online: 15 February 2018

Abstract

Standardized pictorial stimuli and predictors of successful picture naming are not readily available for Gulf Arabic. On the basis of data obtained from Qatari Arabic, a variety of Gulf Arabic, the present study provides norms for a set of 319 object pictures and a set of 141 action pictures. Norms were collected from healthy speakers, using a picture-naming paradigm and rating tasks. Norms for naming latencies, name agreement, visual complexity, image agreement, imageability, age of acquisition, and familiarity were established. Furthermore, the database includes other intrinsic factors, such as syllable length and phoneme length. It also includes orthographic frequency values (extracted from Aralex; Boudelaa & Marslen-Wilson, 2010). These factors were then examined for their impact on picture-naming latencies in object- and action-naming tasks. The analysis showed that the primary determinants of naming latencies in both nouns and verbs are (in descending order) image agreement, name agreement, familiarity, age of acquisition, and imageability. These results indicate no evidence that noun- and verb-naming processes in Gulf Arabic are influenced in different ways by these variables. This is the first database for Gulf Arabic, and therefore the norms collected from the present study will be of paramount importance for researchers and clinicians working with speakers of this variety of Arabic. Due to the similarity of the Arabic varieties spoken in the Gulf, these different varieties are grouped together under the label "Gulf Arabic" in the literature. The normative databases and the standardized pictures from this study can be downloaded from http://qufaculty.qu.edu.qa/tariq-khwaileh/download-center/.

Keywords Gulf Arabic · Nouns · Verbs · Naming · Naming latency · Predictors · Name agreement · Age of acquisition · Imageability · Familiarity · Visual complexity · Image agreement

Picture naming refers to the process of describing a presented picture in no more than one word (Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Kosslyn & Chabris, 1990) and involves three broad levels of processing: visual analysis, semantic activation, and lexical retrieval (Barry, Morrison, & Ellis, 1997; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Levelt, Roelofs, & Meyer, 1999; Nickels & Howard, 1995). The picture-naming task is a widely used experimental paradigm to investigate lexical retrieval in both healthy and unhealthy participants. It is the elementary step toward using language. Since the publication of Snodgrass and Vanderwart's (1980) set of 260 pictures, researchers have been developing linguistically and culturally appropriate normative databases for pictures/words/concepts across different languages and varieties, to be utilized in experimental and clinical research fields. Furthermore, the developed normative databases include norms for factors influencing the lexical retrieval process at various levels (e.g., Barry et al., 1997; Bonin et al., 2003; Kosslyn & Chabris, 1990). These factors are referred to as *determinants* or *predictors* of lexical retrieval, and may include visual complexity of the pictures, name agreement, image agreement, imageability, age of acquisition, frequency, and familiarity. Bonin et al. (2003) stated that the lack of normative databases in a given language or variety results in hindering experimental and clinical research into language processing, leading researchers to develop picture sets that can be highly idiosyncratic, resulting in difficulties matching for relevant factors that could affect the conclusions drawn from these studies.

✉ Tariq Khwaileh
tariq.khwaileh@qu.edu.qa

[1] Department of English Literature and Linguistics, Qatar University, Doha, Qatar

[2] Department of Human Communication Sciences, University of Sheffield, Sheffield, UK

[3] School of Education Communication and Language Sciences, Newcastle University, Newcastle upon Tyne, UK

Normative databases

Cross-linguistic standardized pictures databases are commonly used in psycholinguistic research into language production and comprehension. The purpose of developing such databases is to provide readily available stimuli for use in both experimental linguistic research fields, and clinical fields. They are used to investigate how psycholinguistic variables such as name agreement, age of acquisition, frequency of use, and imageability affect the lexical retrieval process in terms of latency and accuracy in both typical (e.g., Khwaileh, Body, & Herbert, 2014) and atypical speakers (e.g., Khwaileh, Body, & Herbert, 2017). Developing a normative database for a specific geographical region or variation of language, ensures accuracy of results when used in academic and clinical research. Not all languages share the same linguistic features and cultural norms, and for this reason; normative databases for different languages are in demand. The first normative database for English was the Snodgrass and Vanderwart's (1980) set of 260 pictures in American English. This database was then extended to 400 pictures (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997). These two databases have been utilized across many studies into picture naming cross-linguistically (e.g., Bonin, Méot, Chalard, & Fayol, 2002; Bonin et al., 2003; Boukadi, Zouaidi, & Wilson, 2016). Normative databases exist for many languages, including Dutch (Shao, Roelofs, & Meyer, 2013), Portuguese (Cameirão & Vicente, 2010), Spanish (Alonso, Fernandez, & Díez, 2015), Russian (Akinina et al., 2015), French (Bonin et al., 2002; Bonin et al., 2003), Italian (Barca, Burani, & Arduino, 2002), and Turkish (Raman, Raman, & Mertan, 2014). However, the majority of the published normative databases in various languages are noun-based: English (Cycowicz et al., 1997), Dutch (Shao et al., 2013), French (Bonin et al., 2003), and Italian (Barca et al., 2002) to name a few. Noun-based normative databases are formulated for object-naming tasks to elicit verbal identification for pictures representing nouns. Verb-based databases are developed for the purpose of assessing action-naming. There are fewer verb-based than noun-based normative databases (e.g., Russian: Akinina et al., 2015; French: Schwitter, Boyer, Méot, Bonin, & Laganaro, 2004).

Nouns versus verbs processing

Processing of nouns and verbs has been the interest of many studies that aimed at finding whether grammatical class affects language processing. Two different assumptions emerged on processing of nouns and verbs. The first suggests that different grammatical classes may be processed differentially under the assumption that they are neurally separable (e.g., Pinker, 1994). This view has relied on double dissociations reported in aphasia case studies, in which patients showed an advantage of verbs over nouns (e.g., Miceli, Silveri, Villa, & Caramazza, 1984; Zingeser & Berndt, 1988), or patients showing greater impairment in verbs than in nouns (e.g., Caramazza & Hillis, 1991), which depends on the aphasia profile of the patient, leading researchers to conclude that nouns and verbs must be represented separately psychologically and neurally (e.g., Damasio & Tranel, 1993). Within this framework, it is hypothesized that verb processing is more difficult than noun processing, and that action-naming causes various and higher demands on language processing than object-naming, due to the more demanding nature of verb processing (Akinina et al., 2015). Per Mätzig, Druks, Masterson, and Vigliocco (2009), verbs may be less imageable but have more complex representations than nouns. Another factor to consider is the organizational features of nouns versus verbs. Masterson, Druks, and Gallienne (2008) explained that nouns may exist independently as objects in the world, whereas; verbs do not, on the contrary they bear reference to the nouns related to them in terms of instrumentality, location, and actor. Verbs have various argument structures; making it difficult to make generalizations from one verb to another, whereas it is easy to generalize rules from one noun to another, as in the case of plural marker "s" in English (Mätzig et al., 2009). Additionally, verbs are not as easily imageable as nouns.

The second view was first introduced by Sapir (1921) and later studied by functionalist specialists (e.g., Bates & MacWhinney, 1982). This view assumes that grammatical classes are neither behaviorally nor neurally separable. Rather, the perceived difference is an elusive byproduct of semantic/pragmatic distinctions dependent on frequency and co-occurrences within language. Vigliocco, Vinson, Druks, Barber, and Cappa (2011) carried out a comprehensive review of behavioral, electrophysiological, neuropsychological and imaging studies on nouns versus verbs distinctions and concluded that grammatical class is not an organizational principle of knowledge in the brain. They stated that the varying results reported in the literature can be attributed to different language typologies depending on semantic/pragmatic and distributional cues in different languages that distinguish nouns from verbs; different languages differentiate between nouns and verbs in different ways. For example, Arabic nouns and verbs select different vocalic patterns and CV skeletons at a morpho-phonological level. Vigliocco et al. further elaborated that grammatical class (noun–verb) distinction in processing is evident only when a word plays a role in phrase and sentence contexts, as opposed to single word processing. Studies investigating noun-verb distinction within sentence and phrase frames report differences between nouns and verbs, whereas single word processing studies report similarity in processing nouns and verbs (see Vigliocco et al., 2011, for a full review). In support of this view, Scott (2006) found that verbs and nouns actually share the same neural network that is activated upon encounter with nouns and verbs.

Nevertheless, Bird, Howard, and Franklin (2000) argued that imageability influences word retrieval more in nouns than it does in verbs, because the imageability of verbs is lower than the imageability of nouns. However, Berndt, Haengiges, Burton, and Mitchum (2001) reported that imageability is not the only factor that affects action-naming, but factors, such as instrumentality of the verb, name relation between an instrumental verb and the name of the instrument and argument structure, all these can influence word retrieval of verbs. In addition, Bastiaanse and Van Zonneveld (1998, 2005) reported that age of acquisition influences word retrieval for both nouns and verbs, in which the later the age of acquisition the lower the performance in word retrieval. The authors added that imageability plays a big role in word retrieval of nouns and verbs together; the more concrete they are the easier it is to retrieve them. As for the word class factor, it has been confirmed that the retrieval of verbs is more difficult than that of nouns. The authors attributed this difficulty to the grammatical encoder, in which verbs activate more information and lemma information than nouns, requiring a more complex grammatical encoding than nouns.

Previous studies developing verbs normative databases have investigated the predictors of verb retrieval. Akinina et al. (2015) examined the effect of name agreement, familiarity, subjective visual complexity, age of acquisition, imageability and image agreement on 414 black-and-white drawings of actions. They report a significant effect of name agreement and imageability on verb retrieval, suggesting that verbs that evoke images more easily tend to be named more uniformly. Another aspect that may affect latencies in action naming is the mode in which the material is presented; d'Honincthun and Pillon (2008) found that difficulty and latency in action naming was eradicated when a participant were shown video-taped and verbal stimuli rather than photographic stimuli. D'Honincthun and Pillon further argued that due to the fact that verbs tend to bear inflection more than nouns, processing takes longer due to the decisions that must be made on what verb to use in what context, and what inflection to use in a certain context; on top of the lexical retrieval process. However, it has also been suggested that there is no difference in the processing of nouns and verbs, as we reported above.

Predictors of picture-naming latencies

Previous studies developing normative databases, have investigated the impact of psycholinguistic factors on lexical retrieval. A number of factors have been found to influence lexical retrieval in healthy speakers cross-linguistically. These factors are properties of the stimuli and they contribute to the speed and accuracy of lexical retrieval. Variables such as visual complexity, word frequency, age of acquisition, name agreement, image agreement, imageability, familiarity, and word length have been investigated in research utilizing picture-naming tasks.

Visual complexity pertains to the complexity of the lining/ details of an image, and has been found to influence the naming latencies of picture naming (Ellis & Morrison, 1998). Findings from Shao, Roelofs, and Meyer's (2012) study indicate that action pictures that are less visually complex have higher imageability and image agreement, suggesting that the less visually complex an image is, the easier it is to evoke a mental image, and the more accurate the mental image is to the target. However, some studies have found that visual complexity in object naming does not robustly influence naming latency in healthy speakers, as per (Barry et al., 1997; Bonin et al., 2002; Bonin et al., 2003; Cuetos, Ellis, & Alvarez, 1999; Khwaileh et al., 2014; Snodgrass & Yuditsky, 1996). *Word frequency* refers to how frequent a word is used (spoken or written form) in a given language. Previous research suggests that the higher the word frequency, the faster the reaction and the higher the accuracy is in picture-naming tasks (Martein, 1995; Nickels, 1997). Furthermore, word frequency and age of acquisition have been found to be interrelated, per Meschyan and Hernandez (2002); words that are acquired earlier tend to be higher in frequency and they may have stronger lexical representation (Meschyan & Hernandez, 2002). Word frequency is often established through extracting frequency values from corpora or through rating tasks (e.g., Boukadi et al., 2016). *Age of acquisition* relates to the age at which certain words are learnt. The earlier a word is the learned, the faster and more accurately it is processed (e.g., Akinina et al., 2015). Age of acquisition has been reported to affect the latency and accuracy of word retrieval in previous studies (e.g., Akinina et al., 2015; Bonin et al., 2002; Bonin et al., 2003; Cameirão & Vicente, 2010). *Name agreement* refers to the degree to which participants produce the same name to a given picture. A picture may call to mind more than one name, and a given name can call to mind different pictorial representations (Khwaileh et al., 2014). Pictures with high name agreement have been found to have shorter naming latencies (Alario & Ferrand, 1999; Barry et al., 1997; Bonin et al., 2002; Bonin et al., 2003; Boukadi et al., 2016). *Image agreement* pertains to how accurate or close the mental image of a concept is to the presented stimulus. The higher the image agreement rating of an object is, the shorter the naming latency (Alario & Ferrand, 1999), conversely; items with low image agreement, take longer to retrieve due to competition at the visual recognition level (Barry et al., 1997). According to Alario and Ferrand, image agreement intercorrelates positively with name agreement; the higher the name agreement of a stimulus, the higher the image agreement. Alario and Ferrand attribute this to the number of competitors during the lexical retrieval process, in which items with high name agreement have fewer competitors, leading to a faster and more accurate response. *Imageability* refers to the ease of conjuring a mental image to correspond with a presented word (e.g., Akinina et al., 2015; Khwaileh et al., 2014). This variable is significant as the higher the imageability of a given word is, the higher the semantic richness and therefore the faster the response of picture naming (Akinina et al., 2015; Khwaileh et al., 2014; Nickels & Howard, 1995). This can be attributed to the assumption that words with high imageability may have stronger visual and verbal representation Previous studies report that words with high imageability are acquired earlier, and are more familiar, shorter, and have more tendency to have orthographic neighbors than words that are less imageable (e.g., Stadthagen-Gonzalez & Davis, 2006). *Familiarity* pertains to how familiar an object or word is within a specific language and sphere of experience (Boukadi et al., 2016). It has been found that concepts and words with high familiarity of a concept or word are retrieved faster in picture-naming tasks (Akinina et al., 2015; Barca et al., 2002; Boukadi et al., 2016). Furthermore, Boukadi et al. reported strong correlations between familiarity and frequency, suggesting that the names of the most familiar objects are more frequently used or heard in everyday communication. *Word length* concerns the number of syllables or phonemes present within a word. It is assumed that long words take longer time to process in production tasks (Akinina, et al., 2015). However, Alario et al. (2004) found that the number of phonemes in a word does not contribute significantly to naming latencies; they also found that shorter syllable length did not predict shorter latency. Instead, Alario et al. established that longer words caused shorter latencies; and trisyllabic words were processed faster than the monosyllabic and bisyllabic counterparts. They conclude that the effect of word length on naming latencies from healthy speakers is disputed, and therefore the issue warrants further investigation.

With regard to the Arabic language, there are two published normative databases for nouns: the Levantine-Arabic database (Khwaileh et al., 2014) and the Tunisian-Arabic database (Boukadi et al., 2016). Verbs and adjectives normative databases do not exist for any of the Arabic varieties. To the best of our knowledge, normative databases for Gulf Arabic are not readily available neither for nouns nor for verbs. The aim of the present study is to develop a set of standardized object and action pictures for Gulf Arabic, and to determine the predictors of successful retrieval from pictures of nouns and verbs in the variety under investigation.

Gulf Arabic

Although Modern Standard Arabic (MSA) is a variety of Arabic that is used and understood across the Arab region, its use is restricted to writing and formal settings. Instead, local and regional Arabic varieties are used for everyday communication. Contemporary Arabists generally classify modern spoken varieties into the following dialect groups: Egyptian Arabic, Meghrebi Arabic, Yemeni Arabic, Iraqi Arabic, Levantine Arabic, and Gulf Arabic (Holes, 2004, Mustafawi, 2018; Versteegh, 1997) due to linguistic and geographic considerations. Gulf Arabic is a label for the varieties of Arabic that are spoken by more than 26 million citizens in the area including the states of Kuwait, Bahrain, Qatar, Saudi Arabia, the United Arab Emirates, and Oman. This does not mean that Arabic speakers from the Gulf speak in a completely identical way as variation may exist even within the same country or city (Johnstone, 1967). However, certain linguistic attributes distinguish Gulf Arabic from other Arabic dialect groups. Since the present article is based on single words, we will restrict our illustration of the difference between Gulf Arabic and other Arabic dialect groups to aspects of the phonology and the lexicon of the language.

With respect to the phonology of dialect groups, a number of phonemes exist in some dialects or dialect groups but not in others. For example, the affricate /tš/ is part of the phonemic inventory of Gulf Arabic (GA) but is absent from Egyptian Arabic, and from most of the dialects of Levantine Arabic and Meghrebi Arabic. Similarly, some phonemes may exist in other dialect groups but not in GA. Examples of such phonemes are /ḍ/, /ẓ/, and /ž/ whose counterparts in GA are /đ/, /đ/, and /dž/, respectively. Also, the phoneme /g/ of GA is represented by the phoneme / / in Egyptian Arabic and most of the dialects in the Levant. Also, the number and quality of vowels differ in ways other than vowel length among various Arabic dialect groups (Ghazali, Hamdi, & Knis, 2007). In terms of syllable structure, GA and Iraqi Arabic permit more variation than the rest of the dialect groups. There are also differences among the dialect groups in terms of stress patterns and the application of certain phonological processes. For a detailed discussion of phonological differences among Arabic dialect groups, the reader is referred to Mustafawi (2018) and references therein.

As for the lexical differences among the dialect groups, the disagreements appear to be due to the existence of synonyms in the Arabic language in general, with each dialect adopting a specific form or forms. Adopting loanwords from other languages by certain dialects also contributes to the observed lexical disagreements. Table 1 provides a sample of such disagreements. The Gulf Arabic items were obtained from the present study, the Levantine Arabic nouns were obtained from Khwaileh, Body, and Herbert (2014), and the verbs from the first author, who is a native speaker of Jordanian Arabic. The Meghrebi items were obtained from a native speaker of Tunisian Arabic, and the Iraqi items were obtained from a native speaker of Baghdadi Arabic. Some of the listed items exemplify phonological differences among the dialect groups that were referred to above.

On the other hand, and as we indicated above, in most of the Gulf countries, two Arabic varieties are used: an urbanized

Table 1 Examples of noun and verb variations across spoken Arabic dialects

	Gulf Arabic	Levantine Arabic	Egyptian Arabic	Meghrebi Arabic	Iraqi Arabic
Nouns					
A ball	ku:ra	ṭa:be	ku:ṛa	ku:ra	ṭo:ba
A window	diri:ša	šubba:k	šibbæ:k	šibba:k	šubba:tš
An ashtray	ṭaffa:ya	makatte	ṭaffa:yit sagæ:yir	sandriya	Minfađa
A fish	smitša	samake	samaka	ħu:t	Simtša
A pillow	maxadda	wisa:de	maxadda	maxadda	Mxadda
A heater	daffa:ya	ṣo:be	daffæ:ya	saxxa:n	ṣo:pa
Verbs					
He cries	yṣi:ħ/yabtši:	yibki:	biy ayyaṭ	yibki:	yibtši:
He falls	yṭi:ħ	yu:ga	biyu a	yti:ħ	yo:ga
He pushes	ydizz	ydizz	biyzu	ydizz	yidfa
He vacuums	yxumm	ykannis	biyiknis	yuknus	Yiknus

variety and a Bedouin variety.[1] The main differences between these two varieties are a few disagreements in morpho-syntactic structure and very few phonological attributes. This made us ensure the inclusion of representative groups from each of the two varieties in Qatar, expecting we would create two databases, one for urbanized Gulf Arabic and one for Bedouin Gulf Arabic. However, after conducting the experiment we could not find significant difference in the outputs of the two groups; hence, we excluded this distinction from further analysis or reporting. We believe that the reason for this lack of differences between the outputs of the speakers of the two varieties is that the outputs that were sought in the picture-naming experiment consisted of single words. This automatically made the few morpho-syntactic differences between urbanized Qatari Arabic and Bedouin Qatari Arabic irrelevant, since these differences can only appear in longer strings (phrases and sentences). The only other difference between the two varieties is phonological, and this has, to a great extent, leveled over the years, partially due to the process of standardization (Al-Emadidhi, 1986), which was a result of the spread of formal education and mass media, and partially due to the constantly increasing opportunities for contact among the speakers of the two varieties.

Method

Participants

The participants were 170 (39% males; 61% females) native speakers of Qatari Arabic from three volunteer centers in Qatar, including both undergraduate and graduate students from Qatar University. They were informed beforehand that in order to participate, they had to be native speakers of Bedouin or Hadari (urbanized) Qatari Arabic, and should be above 18 years of age. All participants had normal or corrected-to-normal vision. A questionnaire was used to gather demographic information about the participants and their linguistic background. Of the 170 participants, 122 were speakers of urbanized Qatari Arabic, 35 were speakers of Bedouin Qatari Arabic, and 13 were speakers of a mixture of urbanized and Bedouin Qatari Arabic. The average age for participants was 31 years (range: 18 to 51 years old). All 170 participants had completed their secondary education, of which 66 held an undergraduate degree at the time of the experiment, and 104 were still studying for their undergraduate degree at the time of the experiment. The participants were asked to sign informed consent forms, and were provided with an information sheet to explain their role in the study. The study was ethically approved by the Qatar University IRB committee.

Design

Since the aim was to develop matched pictorial sets for use in research and clinical work, the design included a picture-naming task that was conducted to establish naming latency and name agreement. Two picture-rating tasks were undertaken to establish image agreement and visual complexity. Three word-rating tasks were carried out to establish familiarity, age of acquisition, and imageability norms. The apparatus used for the picture-naming tasks consisted of the Presentation software that is a response recorder. It controlled the presentation of the pictures, and it automatically recorded latencies in milliseconds from the time the picture was presented until the onset of the response. If the participant did not respond within 5 s, the software presented the next stimulus. The computer automatically saved the data to an excel sheet and saved sound files of the responses. All rating tasks were presented in separate booklets attached to individual answer sheets for the

[1] More variation exists in larger countries such as Saudi Arabia and Oman.

participant to write down ratings based on a scale of 1 to 5 (for image agreement, visual complexity, and familiarity) or 1 to 7 (for imageability and age of acquisition) next to each word stimulus presented in the answer sheet. The use of different scales for different variables is due to the nature of each variable in question. For example, age of acquisition requires a larger scale than visual complexity due to the high variability in age of acquisition ratings as opposed to visual complexity, which can be either complex or easy with less rating points in between (Alario et al., 2004; Biederman, 1987; Bonin, Boyer, Méot, Fayol, & Droit, 2004; Cuetos & Alija, 2003; Paivio, Clark, Digdon, & Bons, 1989; Schwitter et al., 2004; Shao et al., 2015; Snodgrass & Corwin, 1988; Snodgrass & Yuditsky, 1996).

For the image agreement and visual complexity tasks, pictures were projected onto a laptop screen for individuals, or onto a large white screen by an overhead projector for groups. All items were randomized using the randomizing function on Microsoft Office Excel. Four different lists were generated—that is, Lists A, B, C, and D. Randomizing the order was conducted to avoid the effect of word location in the set on picture naming. Each of the four different word lists was checked for semantic relatedness and initial phonemes of neighboring words, to ensure that successive items did not share semantic features or initial phonemes. The randomization process was repeated for all rating tasks in the present experiment. Each participant received different order of the stimuli in each task presented in the same session. A given participant would have done List A in the picture-naming task, List B in the visual complexity task, and List C in the age-of-acquisition task. In the second session, they would have done List D for the imageability task, List A for the familiarity task and List C for the image agreement task.

Materials

The materials used in the present study consisted of 334 line drawings representing concrete nouns, and 170 line drawings representing action verbs. The selection of the these nouns and verbs was based on most occurring nouns and verbs in Gulf drama and television programs and in everyday interactions within the Qatari society. The line drawings representing the nouns and verbs were drawn by three artists. These pictures were redrawn when found to be ambiguous or culturally inappropriate. An instance of this is a picture of a glass; which illustrated a drawing of a wine glass. This was not in line with cultural norms and did not represent the prototype of a glass in the Arab (Qatari) culture, which is a glass with no stem. An instance of actions is the verb "to fish," which illustrated a man using a fishing rod. This representation was not in line with the Qatari culture. However, sea activities have been part of the Qatari Hadari/urban culture for centuries, so a prototypical image of a man fishing would be a fisherman using a traditional net called the "ghazal" instead of a fishing rod. To maintain consistency of the style of drawings across the categories, two of the artists who used the same drawing software were assigned a list of nouns; and the third artist was allotted the list of action verbs that were to be drawn by hand on paper. The drawings were originally drawn at A4 size, and were then presented as digital files. Furthermore, the artists were given specific guidelines that emphasized that the objects and the actions must be drawn with respect to the local culture. Each picture was shown individually to three Bedouin speakers, and three Hadari/urbanized speakers (mean: 24 years old; two males and four females) who were not involved in the normative study. They had to assign a name to each object and action depicted by the drawings. They were asked to provide feedback about culture appropriateness and the name used to describe the drawing. Items agreed upon by the native speakers were kept for the normative study, and were used to collect norms for naming latencies, name agreement (through the picture-naming task), image agreement, and visual complexity (through rating tasks).

Procedure

The data were collected over four sessions with two weeks in between each session. In the first session, all participants completed the picture-naming tasks, the visual-complexity rating tasks, and the age-of-acquisition rating tasks. The average administration time for session one was 50 min per participant. In the second session, which was administered two weeks after the first one, 148 participants out of the 170 participants participated in the imageability rating task (22 participants were not available at the time when the second session was administered). The average administration time for Session 2 was 15 min per participant. Two weeks later the participants were invited to complete the familiarity rating task, 116/170 participants participated in this task. The image agreement task was carried out two weeks after the familiarity task, and 121/170 participants participated in this task. The rationale for separating the sessions was to prevent memory and priming effects in the imageability, familiarity, and image agreement rating tasks.

All sessions were conducted in a sound proofed room. At the beginning of each session participants were encouraged to respond carefully and consistently to each task. At the start of each task, participants were given instructions and were taken through practice items prior to commencing the task in question, followed by feedback. Instructions were given in Arabic; rating scales and other written materials were in Arabic script. A full description of each task conducted in the present experiment is reported below. The tasks below are presented according to their order of administration. The researcher controlled the presentation of all tasks, and participants were given the opportunity to take a break.

The picture-naming task was performed individually in isolated rooms, and all rating tasks were performed either individually or collectively, depending on the number of participants available during the same time. At the beginning of each rating task, instructions were provided in writing and verbally by the experimenter, along with each task's rating scale printed inside the task booklet. The experimenter explained to the participants that they were free to use any number on the scales, as long as it indicated their true judgment. A booklet for each of the five rating tasks was prepared with separate answer sheets. In the imageability, familiarity, and age of acquisition task booklets; a list of all the nouns and verbs appeared under two categories in writing. A list appeared under the Bedouin dialect, and a list appeared under the Hadari/urbanized dialect. Both varieties were listed in parallel inside each task booklet, in correspondence to the same item, and participants were asked to use the list that corresponds to their dialect. In the image agreement and visual complexity tasks, a list of the nouns and verbs corresponding to their projected pictures appeared under each category. Participants were asked to rate the list of words that appeared under the category of the dialect they speak as their mother tongue. In the case that the participant speaks both dialects as their mother tongue—that is, with each parent speaking a different dialect—they were asked to rate the list that corresponds to their mother's dialect.

During the picture-naming tasks, participants sat at a distance of 50 cm from a laptop screen. They were initially shown the line drawings of objects, and were asked to say out loud the first name that comes to mind, as quickly and as accurately as possible. The researcher explained that the task was to name the object in the picture using one word only, and to avoid describing it. The same instructions were applied for the second group of the action drawings, in which the focus was to name the action being carried out in the picture, rather than the object itself, using one word only. The software used for these tasks, presented a signal in a form of a cross (+), which appeared in the center of the screen for 1,000 ms. immediately followed by the picture. The cross served as a prompt to look at the center of the screen in preparation for the upcoming picture, which remained for 5 s before the next stimulus appeared. When the participant could not recognize the picture or did not know the name of the picture, they were asked to say out loud that they could not recognize the object/action, and the researcher would take a note of the item to revisit after the experiment and delete it's naming latency from the list. The average time of administration of the picture-naming task was 20 min. All sound files were exported to PRAAT (Version 6.0.08), and each sound file was revisited to make sure that the software did not include false triggering of noise or "em" or "err." False triggering and failures to press the response time key were noted, and were revisited at the end of each task. Responses were transcribed and coded by the first author using a numerical coding system (see the Appendix). Only pictures that were named accurately within the allotted time frame (5 s) were scored correct.

In the image agreement task, participants were asked to rate how closely each picture resembled their own mental image of the noun/action provided in writing in the answer sheet. They were first shown a section with nouns to rate, and then a section with verbs. For every word, they were given approximately 3 s to form a mental image of it, then were shown the corresponding picture on a screen and were asked to rate the degree of agreement between the picture and their mental image using a 5-point scale, where 1 indicated *low agreement* and 5 indicated *high agreement*. The average administration time for this task was 20 min.

During the visual complexity task, participants were asked to rate the degree of complexity of each drawing using a 5-point scale. They were first shown a section with nouns to rate, and then a section with verbs. They were informed that they should rate the complexity of the drawing, rather than the complexity of the real-life object/action it represented. "Complexity" was defined by the amount of details and lines in each drawing, where 1 corresponded to *very simple* and 5 corresponded to *very complex*. The average administration time for the visual complexity task was 20 min.

In the imageability task, the participants were asked to indicate whether each word evoked a mental image with great difficulty (rated 1) or very easily (rated 7). In the age of acquisition task, the participants were asked to estimate the age at which they thought they learned each word presented in the booklet. They were informed that the estimate should not only attribute to when they had first heard the word, or when they first learned to speak it, but to estimate the age at which they first understood the word when it was used in front of them. In this task, the values in the scale corresponded to 2-year age bands, with 1 corresponding to *0–2 years*, and 7 corresponding to *13 years or after*. In the familiarity task, the participants were asked to rate the degree of familiarity of the item in terms of how usual/unusual the word was in their realm of experience, regardless of its meaning. Participants were informed that the rating had to be attributed to how often they came across the word itself, rather than its concept, in either its heard, spoken, or written form. A word they came across very often should be rated 5, and a word they had never seen or heard should be rated 1. The average administration time for each of the three rating tasks was 20 min.

Frequency and intrinsic features

The frequency of the orthographic form for each item in the nouns and verbs sets were extracted from Aralex (Boudelaa & Marslen-Wilson, 2010). The frequency of orthographic form for each word was included as a compensatory measure for spoken frequency, due to the fact that frequency corpora for

Gulf Arabic are not readily available. Available frequency corpora on Arabic are drawn from Arabic written material (see the Buckwalter Arabic Corpus 1986–2003, or the An-Nahar Corpus, available from ELRA: http://catalog.elra.info/product_info.php?products_id=767), and Modern Standard Arabic (e.g., Aralex database; Boudelaa & Marslen-Wilson, 2010). Furthermore, other variables that are intrinsic features of words (can be determined directly from their surface structure) were included in the database. These are gender, animacy, rationality, pluralization type for nouns, and number of syllables and number of phonemes for both nouns and verbs.

Results

The original 334 object pictures and 170 action pictures yielded naming latencies for the nouns and verbs in question. All items in question were rated for imageability, image agreement, name agreement, age of acquisition, familiarity, and visual complexity. Intrinsic values (syllable length, phoneme length, and orthographic frequency) for the nouns and verbs were also extracted. The data was analyzed to establish norms for the various variables. Further analyses investigated the influence of the independent variables on naming latencies of nouns and verbs.

Picture-naming task data

Coding of the responses from the nouns and verbs picture-naming tasks was based on a 10-category coding system: *correct response*, *visual errors*, *semantic errors*, *phonological errors*, *morpho-syntactic errors*, *unrelated errors*, *tip-of-the-tongue*, *don't know name of* (the object/action), *don't recognize* (the object/action), and finally *no response* errors (i.e., no response produced within 5 s). For the noun picture naming, the coding issues were minimal as most responses were accurate, they mainly included the production of visually or semantically related items, however most of these items had low frequency values and familiarity ratings, for example, producing "screw" /s kru:b/ in Qatari Arabic for pictures of a "pin" or a "needle." The coding of verbs/actions picture naming was more challenging. Examples of such issues included instances of producing alternative masculine verb form instead of the feminine verb form (e.g., /jħ b/ [masculine] to /tħ b/ [feminine] "to kiss"). This was considered a morpho-syntactic error indicating a different gender to the target word. It could have been considered an acceptable alternative, since it shares the same consonantal roots, but this would have affected the sensitivity of detecting morpho-syntactic errors when the database is used with patients with agrammatism. Another instance of such issues was the production of a verb that intrinsically involved a doer instead of the target form, which rather involved the action being centered on the object itself (e.g., /j nz f/ "to bleed" to /j d rαħ/ "to hurt"). This was considered a visual/semantic error.

The picture-naming task yielded naming latencies and recorded responses for 334 nouns and 170 verbs/actions. Only latencies for accurate responses were included. All the naming latencies and responses were manually checked for false triggers using PRAAT (version 5.1.17; Boersma & Weenink 2009). Responses not produced within 5 s and responses that were coded as tip-of-the-tongue errors, "don't know the name," or "don't know the object/action" errors were removed from the database. The total number of items removed from the nouns" set was 15 items, and from the verbs" set 29 items. Removing these items from the database resulted in naming latencies for 319 nouns, 141verbs, and their pictorial representations. The name agreement ranges for nouns and verbs are shown in Table 2. Finally, the data were checked for outliers. To remove the effect of extreme outliers on the reaction time data, the 5% trimmed means procedure was performed (Pallant, 2005). This procedure replaced extreme outliers with values of the mean plus two standard deviations and recalculated a new mean for each item. Naming latencies of two standard deviations and above were deemed outliers, and were removed using the trimmed means procedure, prior to the analysis for both nouns and verbs.

Items with low name agreement were kept in the database to maintain a wide range of variance of the data for future research use in investigating effects of name agreement. Within clinical contexts, clinicians can select the items with high name agreement from the databases. Researchers may need more variance in name agreement values depending on the purpose of their research.

Rating task data

The rating tasks yielded visual complexity, imageability, image agreement, age of acquisition, and word familiarity for the 319 nouns, and 170 verbs. Participants with ratings falling more than three standard deviations away from the average mean were excluded, in line with Schock, Cortese, and Khanna (2012)

Table 2 Name agreement subsets for the noun and verbs

Name agreement percentage (%)	Number of nouns	Number of verbs
100–90	145	18
89–80	67	18
79–70	35	14
69–60	25	17
59–50	17	17
<50	30	57
Total number	319 items	141 items

and Bakhtiar and Weekes (2015). Cronbach's alpha revealed high internal consistency across nouns' ratings: visual complexity ($\alpha = .904$, $n = 334$), imageability ($\alpha = .821$, $n = 334$), image agreement ($\alpha = .912$, $n = 334$), age of acquisition ($\alpha = .781$, $n = 334$), and word familiarity ($\alpha = .793$, $n = 334$). Within verb ratings, Cronbach's alpha showed that the internal consistency for visual complexity ($\alpha = .741$, $n = 170$), imageability ($\alpha = .791$, $n = 170$), image agreement ($\alpha = .723$, $n = 170$), age of acquisition ($\alpha = .711$, $n = 170$), and word familiarity ($\alpha = .801$, $n = 170$) was high. This shows that the internal consistency of ratings was above moderate ($\alpha > .500$), indicating that participants were rating every item in the set consistently.

The means and standard deviations for naming latencies, ratings of visual complexity, imageability, image agreement, age of acquisition, and word familiarity were calculated to establish the norms for the nouns, verbs and their pictorial representations. The percentage of participants agreeing on a given name for the pictures representing each noun and verb was established as a measurement of name agreement. Variables that are intrinsic features of the nouns, and verbs were also included in the final database (e.g., phoneme number, syllable number and gender). The final database included norms for 319 object pictures and 141 action pictures, along with their ratings for the above mentioned variables. The databases and the standardized pictures can be downloaded from http://qufaculty.qu.edu.qa/tariq-khwaileh/download-center/. Table 3 summarizes the means and standard deviations for all the variables in the database.

Predictors of naming latencies in Gulf Arabic nouns and verbs

To determine the significant predictors of nouns and verbs retrieval, trimmed naming latencies underwent correlations, multiple regressions, and principal component analysis (PCA; i.e., factor analysis). This procedure was carried out for nouns only, verbs only, then the nouns and verbs combined. The dependent variable was the trimmed naming latency, and the independent variables were syllable length, phoneme length, initial phoneme (multiple regression only) frequency, imageability, image agreement, name agreement, age of acquisition, familiarity, and visual complexity. Word class (nouns versus verbs) was added as an independent variable for the analysis of nouns and verbs combined.

Analysis of nouns

In preparation for the analysis of the 319 nouns, a total of 27 items were removed from the analysis: eight nouns yielded compound nouns with no length data; eight nouns with no frequency data; and 11 nouns that had a name agreement value of less than 40%. The final set of nouns included 292 items. To explore the relationships between the variables in question, their strength and direction, the Pearson correlations were carried out. These relationships are demonstrated in Table 4.

We found significant correlations of the nouns' naming latencies and (in descending order) name agreement, age of acquisition, imageability, image agreement, familiarity, frequency, and visual complexity. All of these were in the expected directions. There were substantial correlations between the independent variables. All these correlations were also in the expected directions. For example syllables and phonemes correlated at .822; this makes it challenging to have an independent effect in predicting naming latency, since they are strongly related. Other notable significant correlations were in the .129 to .483 range, allowing the inclusion of those in the multiple regression model.

The standard multiple regression procedure was carried out to explore the predictive ability of the independent variables on naming latency. All variables included in the correlation table above were included as independent variables. The included data met the assumptions of normally distributed

Table 3 Summary of the database: Means and standard deviations

Variable	Nouns		Verbs	
	Mean	Standard deviation	Mean	Standard deviation
Naming latency	1,601.02 ms	416.3 ms	1,793.69	382.58
Name agreement (%)	0.86	0.17	0.73	0.21
Visual complexity	2.46	0.81	2.73	0.64
Image agreement	4.36	0.42	4.45	0.42
Imageability	6.10	0.36	5.93	0.36
Age of acquisition	3.63	0.68	3.91	0.67
Familiarity	3.71	0.51	3.96	0.39
Frequency	3.29	0.93	3.21	0.83
Phoneme length	5.23	1.29	6.07	0.93
Syllable length	2.17	0.73	2.31	0.46

Table 4 Correlation matrix for nouns only

	Syllable length	Phoneme length	Frequency	Name agreement	Visual complexity	Image agreement	Age of acquisition	Imageability	Familiarity	Naming latency
Syllable length		.842**	– .120*	– .036	.108	.048	.073	.009	– .040	.031
Phoneme length			– .142*	.024	.078	.032	.108	– .029	– .40	.089
Frequency				.010	.039	.039	– .066	.108	.188**	– .221**
Name agreement					.008	.289**	– .225**	.260**	.129*	– .589**
Visual complexity						– .167**	.094	– .269**	– .177**	.132**
Image agreement							– .142*	.275**	– .001	– .434**
Age of acquisition								– .483**	– .581**	.442**
Imageability									.480**	– .467**
Familiarity naming latency										– .299**

** Significant at the .01 level; * Significant at the .05 level.

residuals, homogeneity of variance and multicollinearity. The data contained no outliers. The model accounted for 57.1% (R^2 = .571) of the naming latency variance. The regression was significantly different from zero [F(9, 282) = 42.61, p < .001], suggesting that the model was appropriate for the investigated data. The regression analysis revealed that factors significantly predicting naming latency in descending order were: name agreement [Beta = – .456; t(116) = – 10.37, p < .05]; image agreement [Beta = – .264; t(48) = – 5.65, p < .05]; age of acquisition (Beta = .216; t(35) = 4.32, p < .05]; frequency [Beta = – .171; t(20) = – 4.12, p < .05]; familiarity [Beta = – .145; t(48) = – 3.42, p < .05]; and visual complexity [Beta = .101; t(24) = 3.01, p < 0.05]. Other variables did not show significant contribution to the naming latency variance: Initial phoneme [Beta = .354; t(27) = 0.101, n.s.], phoneme length [Beta = .173; t(27) = 2.01, n.s.], syllable length [Beta = – .141; t(47) = – 2.21, n.s.] and imageability [Beta = – .091; t(65) = – 1.67, n.s.]. Then, a factor analysis (the PCA with Varimax rotation) was carried out to explore the relatedness of the independent variables (all nine independent variables listed above), to condense them into a smaller number of factors, on the basis of the underlying patterns of the correlations among those variables. The sample size and the strength of inter-correlations were suitable, as recommended by Tabachnick and Fidell (2007). The KMO value was .583, and Bartlett's test was significant (p = .000).

The PCA with Varimax rotation showed that only four components recorded eigenvalues above 1 (2.332, 1.912, 1.214, and 1.219), explaining a total variance of 72.02%. This extracted four orthogonal factors: Familiarity (loading on imageability = .743, age of acquisition = – .759 and familiarity = .723), Length (loading on number of syllables = .892 and phonemes = .882), and Agreement (loading on image agreement = .709 and name agreement = .498). The fourth component was Visual Complexity, with a substantial loading only on visual complexity.

The four orthogonal factors extracted from the PCA were inserted into a multiple regression as independent variables to check their predictive power of naming latency for nouns. The model accounted for 54.9% (R^2 = .549) of the naming latency variance. The regression was significantly different from zero [F(4, 287) = 84.98, p < .000]. The regression analysis revealed that the Agreement factor, combining image agreement and name agreement, had the highest predictive power of naming latency (Beta = – .587; t = – 14.01, p < .000). The Familiarity factor, combining imageability, age of acquisition and familiarity, was the second significant predictor of naming latency (Beta = – .487; t = – 12.16, p < .000). The Length factor (syllable and phoneme numbers) did not show significant predictive power of nouns' naming latency (Beta = .065; t = 1.78, n.s.), nor did visual complexity.

Analysis of verbs only

Forty-six verbs were removed from the original set of 141 verbs: Four verbs yielded compounds with no length data; nine verbs with no frequency data; 33 items with name agreement less than 40%. Only 95 verbs entered the analysis. All naming latencies (trimmed) yielded by verb pictorial representations were inserted into a Pearson's correlation with the nine independent variables described above. The initial phoneme was included in the multiple regression analysis. Table 5 shows the strength, direction and significance of these correlations.

We found significant correlations between verbs' naming latencies and (in descending order) image agreement, name agreement, imageability, and age of acquisition. One variable showed just above a significant correlation: familiarity (r = – .202, p = .058). All of these were in the expected directions. All correlations between the independent variables were also in the expected directions.

The standard simultaneous multiple regression procedure was carried out to explore the predictive ability of the

Table 5 Pearson's correlation matrix for verbs only

	Syllable length	Phoneme length	Frequency	Name agreement	Visual complexity	Image agreement	Age of acquisition	Imageability	Familiarity	Naming latency
Syllable length		.806**	– .148	.172	.106	.117	.116	.017	– .053	.021
Phoneme length			– .297**	.124	.089	.108	.136	.047	– .085	.032
Frequency				.187	– .163	.038	– .093	.030	.238*	– .091
Name agreement					– .109	.358**	– .268**	.431**	.027	– .595**
Visual complexity						– .134	.217*	.060	– .060	.139
Image agreement							– .369**	.421**	.087	– .602**
Age of acquisition								– .381**	.065	.456**
Imageability									.243*	– .587**
Familiarity Naming latency										– .202

** Significant at the .01 level. * Significant at the .05 level.

independent variables on naming latency. The included data met the assumptions of normally distributed residuals, homogeneity of variance and multicollinearity. The model accounted for 59.6% ($R^2 = .596$) of the verbs' naming latency variance. The regression was significantly different from zero [$F(9, 85) = 14.24$, $p < .000$], suggesting that the model was appropriate for the investigated data. The regression analysis revealed that only two variables significantly predicted naming latency in descending order: Name agreement (Beta = – .425; $t = – 4.86$, $p < .05$); image agreement (Beta = – .387; $t = – 4.73$, $p < .05$). As in the correlation analysis, familiarity showed an effect that is just below significance (Beta = – .146; $t = – 1.82$, $p = .08$). None of the other variables showed significant contribution to the naming latency of the verbs in question.

The PCA with Varimax rotation was carried out to explore the relatedness of the independent variables (all nine listed above). The sample size and the strength of intercorrelations were suitable, as recommended by Tabachnick and Fidell (2007). The KMO value was .564, and the Bartlett's test is significant ($p = .000$).

The PCA with Varimax rotation showed that four components recorded eigenvalues above 1 (2.224, 2.048, 1.186, and 1.074), explaining a total variance of 71.91%. This extracted three orthogonal factors: Familiarity (loadings on frequency = .471, age of acquisition = .478, and familiarity = .801), Length (loadings on number of syllables = .889 and phonemes = .923), and Agreement (loadings on image agreement = .743, name agreement = .719, and imageability = .757). The fourth orthogonal factor, Visual Complexity, contained visual complexity (.791) and imageability (.463).

The four orthogonal factors extracted from the PCA were inserted into a multiple regression to check their predictive power of naming latency for nouns. The model accounted for 54.5% ($R^2 = .545$) of the naming latency variance. The regression was significantly different from zero [$F(4, 90) = 26.51$, $p < .0001$]. The regression analysis revealed that the Agreement factor, combining image agreement, imageability, and name agreement, was the only significant predictor of verbs' naming latencies (Beta = – .724; $t = – 10.13$, $p = .000$). The Familiarity factor, combining frequency, age of acquisition and familiarity, showed a smaller effect on verbs' naming latencies (Beta = – .167; $t = – 2.24$, $p = .038$). The Length (syllable and phoneme numbers) and Visual Complexity (visual complexity and imageability) orthogonal factors did not show significant predictive power of verbs' naming latencies.

Analysis of nouns and verbs combined

The Pearson correlations, multiple regression, and the PCA were repeated to explore whether a different pattern would emerge when nouns and verbs were taken together. The dependent variable was the naming latency for nouns and verbs taken together ($n = 387$). All nine variables mentioned above were included as independent variables. The Pearson correlation results are shown in Table 6.

There were significant correlations of the naming latencies of nouns and verbs combined. These correlations were with name agreement, age of acquisition, imageability, image agreement, familiarity, frequency, and visual complexity. All of these were in the expected directions. We found substantial correlations between the independent variables; for example, syllables and phonemes correlated at .822. Other notable correlations were those between imageability, familiarity, and age of acquisition (all in the .31 to .44 range).

A simultaneous regression was then carried out. The regression included all the independent variables combined for nouns and verbs (NV), and the combined naming latency (NV) was set as the dependent variable. The model accounted for 58% ($R^2 = .580$) of the naming latency variance. The regression was significantly different from zero [$F(19, 367)$

Table 6 Correlation matrix of nouns and verbs combined

	Phoneme length	Frequency	Name agreement	Visual complexity	Image agreement	Age of acquisition	Imageability	Familiarity	Naming latency combined
Syllable length	.822**	– .126*	– .013	.119*	.066	.093	.017	– .021	.042
Phoneme length		– .173**	– .024	.119*	.069	.157**	.014	.019	.083
Frequency			.064	– .005	.035	– .078	.085	.183**	– .201**
Name agreement				– .055	.277**	– .267**	.275**	.047	– .578**
Visual complexity					– .144**	.142**	– .182**	– .119*	.189*
Image agreement						– .177**	.317**	.035	– .456**
Age of acquisition							.432**	– .393**	.465**
Imageability								.439*	– .487**
Familiarity Naming latency									– .345**

** Significant at the .01 level. * Significant at the .05 level.

= 26.67, $p < .000$]. The regression analysis revealed that name agreement (NV) was the most significant predictor of naming latency (Beta = .103; $t = 2.41$, $p < .05$), then came frequency (NV) (Beta = .109; $t(16) = 2.42$, $p < .05$). The remaining variables did not show significant effects when combined: visual complexity (NV) [Beta = – .041; $t(19) = – 0.110$, n.s.]; image agreement (NV) [Beta = – .046; $t(18) = – 0.845$, n.s.]; age of acquisition (NV) [Beta = – .062; $t(19) = – 1.32$, n.s.]; imageability (NV) [Beta = – .023; $t(20) = – 0.447$, n.s.]; familiarity (NV) [Beta = – .011; $t(21) = – .167$, n.s.]; initial phoneme [Beta = – .049; $t(20) = – 0.479$, n.s.]. Word class (noun vs. verb) is not a significant predictor of performance.

Syllable length (NV), phoneme length (NV), frequency (NV), name agreement (NV), visual complexity (NV), image agreement (NV), age of acquisition (NV), imageability (NV), and familiarity (NV) were included in the PCA. The Kaiser–Meyer–Olkin value met the recommended value of .6 and Bartlett's test of sphericity reached statistical significance, supporting the factorability of the correlation matrix. The PCA with Varimax rotation showed that only three components recorded eigenvalues above 1 (2.287, 1.884, and 1.136), explaining a total variance of 58.67%. This extracted three orthogonal factors: Familiarity (loading on imageability = .642, age of acquisition = – .626 and familiarity = .867), Length (loading on syllable number = .954 and phoneme number = .967), and Agreement (loading on image agreement = .780 and name agreement = .708).

In the first block, the three orthogonal factors from the NV PCA (length, familiarity and agreement) were entered. The first block (model) accounted for 50.9% ($R^2 = .509$) of the variance in naming latencies. The model was significantly different from zero [$F(6, 380) = 65.48$, $p < .000$]. The coefficients show significant effects of the agreement (Beta = – .611; $t = – 16.621$, $p < .000$) and familiarity (Beta = – .378; $t = – 10.102$, $p < .000$) orthogonal factors but not length (Beta = .043; $t = 0.791$, n.s..) or word class-noun versus verb (Beta = .037; $t = 0.549$, n.s.).

The second block included the three factors and word class. The model accounted for 51% ($R^2 = .510$) of the variance in naming latencies. The model was significantly different from zero [$F(7, 379) = 55.83$, $p < .000$]. None of the orthogonal factors showed significant prediction of naming latencies when word class (NV) were combined. There was no significant effect of adding these variables.

Discussion

The present study was carried out to establish a database of line drawings for Gulf Arabic nouns and verbs. Norms for naming latencies, name agreement, visual complexity, image agreement, imageability, age of acquisition, and familiarity were established. In addition, the database includes other intrinsic factors, such as syllable length and phoneme length. It also includes orthographic frequency values (extracted from Aralex; Boudelaa and Marslen-Wilson, 2010). This normative database is the first linguistically and culturally appropriate dataset of its kind for Gulf Arabic. The stimuli for the present database were developed to accommodate the demand for a purposely developed normative database for both research and clinical fields within the Gulf region (e.g., Khwaileh, Mustafawi, Howard, & Herbert, 2016). Linguistic and cultural appropriateness is of utmost importance to consider when developing a normative database, precision of cultural context must be maintained to ensure accuracy in data collection, and to cater to the specific linguistic and cultural contexts.

The influence of the variables in question on naming latency was examined and compared between nouns and verbs. The present findings suggest that name agreement is a significant predictor of naming latency in picture naming in healthy Gulf Arabic speakers in both nouns and verbs. This finding is in line with various studies (Barry et al., 1997; Bonin et al., 2002; Boukadi et al., 2016, Lachman, Shaffer, & Hennrikus, 1974; Paivio et al., 1989; Snodgrass & Yuditsky, 1996; Vitkovitch & Tyrrell, 1995), all of which have found that name agreement significantly contributes to latency in spoken picture naming. Name agreement is a robust predictor of naming latency (Alario et al., 2004); name agreement is the degree to which a noun object is named with the same term. The higher the name agreement is, the fewer competing lexical items exist for an object, which significantly influences naming accuracy and naming latency. Per Mätzig et al. (2009), verbs are not as richly semantically represented and have more complex representations than nouns and are therefore more susceptible to name agreement. Furthermore, as opposed to nouns, verbs do not exist as independent objects in the world, instead, they refer to actions and states; and therefore tend to have more name agreement variance as evident from the verb name agreement results presented in this study.

The present results indicate that age of acquisition significantly influences naming latency in both nouns and verbs. This is in line with Bonin et al. (2003) and Meschyan and Hernandez (2002) who found a large contribution of age of acquisition in naming speed. According to Meschyan and Hernandez, words learnt at a later age have weaker lexical representations than earlier- learned words. An early explanation of the effects of age of acquisition was put forth by Brown and Watson (1987); their phonological completeness hypothesis posited that during early stages of acquisition, phonological output representations are stored in a complete form, whereas later acquired words are stored segmentally and are therefore more difficult, and take longer to assemble, causing a larger naming latency. Another interpretation of the effect of age of acquisition on verbs is that verbs have been found to be acquired later on in life than nouns as reported in Bird, Franklin, and Howard (2001). An explanation as to why verbs are acquired later than nouns is their morphologically complex nature; verbs must undergo processes of inflection and tend to be heavily conjugated. Furthermore, during the process of verb acquisition; generalizations are more difficult to be drawn from one verb to another (Gleitman, 1994). An example of this is inflection for tense in words such as *write/wrote/written* (Masterson et al., 2008). The impact of age of acquisition on verbs has been proven to influence native speakers of other Semitic languages. Berman (2003) found that Hebrew speakers aged 3–4 were less successful at verb innovation-that is; the coinage of new verbs through identification and isolation of the consonantal skeleton (which is that of nonconcatenative morphology), whereas school-age children were able to successfully do so. This suggests that the effect of age of acquisition on naming latency of nouns is a universal phenomenon, independent from language typology.

Image agreement is a predictor of naming latency in both verbs and nouns as indicated in the present study. Words that are rated with higher image agreement are named faster than those with lower ratings (cf. Alario & Ferrand, 1999). To account for this, Barry et al. (1997) found that pictures that had higher image agreement ratings had shorter latencies than those with lower ratings. Barry et al. posited that image agreement influences at the level of object recognition; that is, the more accurate the stimulus is to the mental image of that object, the faster and more accurate the naming. This is because processing at this level is faster when the pictured item is close to the stored mental description.

Processing association between image agreement and name agreement was found to be present in Arabic nouns and verbs as evident from the principal component analysis. This relationship amounts to the lesser competing lexical entries as opposed to a stimulus with low name agreement, which would have a larger amount of competing lexical entries, and would cause naming latency. In verb-/action-naming tasks, name agreement and image agreement also correlate (as found by Akinina et al., 2015; Bonin et al., 2004; Shao et al., 2012): Named actions that have a more uniform mental image tend to be given more uniform names, indicating that a conventional image exists for the verb in question. Thus, the more a verb action name is able to evoke a common mental image, the more able participants are to accurately name it. This suggests that verbs with higher image agreement and name agreement tend to have less competing lexical entries, and are therefore named more quickly and uniformly. The processing association between these two variables, can be attributed to the rich diversity in the linguistic arena in Qatar and the Gulf region. The region attracts people from all over the world including hundreds of thousands of speakers of other varieties of Arabic. Consequentially many lexical borrowings and different dialectal terms for the same words are introduced to the local varieties. The existence of various lexical items for a noun object creates competition and latency during object-naming tasks. This could be one of the reasons leading the name agreement and image agreement effects found in the present data.

Imageability is also found to be a significant predictor of naming latency in nouns and verbs, too. Nouns that are highly imageable have shorter naming latencies (Bonin et al., 2002). This faster reaction occurs because of the semantic richness and dual coding (visual and verbal) that highly imageable lexical items have (Akinina et al., 2015). Lexical items that are highly imageable tend to be highly concrete in evoking sensory images of their referents. Paivio and Yarmey (1966) found that the naming latency for image arousal was quicker for concrete nouns than abstract nouns. Verbs on the other hand, tend to have low imageability ratings per (see, e.g., Eviatar, Menn, &

Zaidel, 1990). Therefore, verbs take longer to name (e.g., Kauschke & von Frankenberg, 2008), which can be explained by the semantic representation of verbs as compared to nouns, and are more complex, as was explained by Huttenlocher and Lui (1979). However, despite this, verb stimuli's naming latencies are influenced by the same psycholinguistic variables as nouns.

Familiarity significantly contributes to naming latencies in both nouns and verbs in the present study. Studies have found that familiarity does have an effect on latency (Feyereisen, Van der Borght, & Seron, 1988; Snodgrass & Yuditsky 1996); in the sense that the higher the familiarity of the object being presented, the shorter the latency. However, a study has questioned the reliability of familiarity rating tasks due to factors that may influence what participants may consider as familiarity (Balota, Pilotti, & Cortese, 2001); participants may rate items for familiarity on the basis of their semantic meaningfulness, or the familiarity of the sub lexical spelling to sound correspondence instead of the frequency of exposure to the object in question. In the case of nouns, imageability, age of acquisition and familiarity intercorrelate, suggesting that words learned at an earlier age tend to be more imageable, and more familiar that is in line with Stadthagen-Gonzalez and Davis (2006). As we know, nouns are learned much earlier in life than verbs (Bird et al., 2001). In the case of verbs; frequency, age of acquisition and imageability correlate, this indicates that verbs that are highly imageable and are frequently used tend to be more familiar.

The processing association between familiarity and frequency in the present data could be understood under the assumption that familiarity could be a measure of spoken frequency. Previous literature assumed that word frequency correlate with word familiarity. Tanaka-Ishii and Terada (2011) define word familiarity as "the relative ease of perception attributed to every word" (p. 96). However, the processes that are involved when readers rate familiarity have been a matter of dispute. Some studies interpret familiarity ratings as a measure of exposure frequency (MRC Psycholinguistic Database, 2006), others view it as an underlying effect of frequency influencing perception (Dupoux & Mehler, 1990; Marslen-Wilson, 1990; Segui, Mehler, Frauenfelder, & Morton, 1982). In spite of this, some studies have advocated the use of familiarity acquired through ratings, as a better predictor of word processing than frequency (Gernsbacher, 1984; Gordon, 1985; Kreuz, 1987; Nusbaum, Pisoni, & Davis, 1984). In their in-depth analysis of frequency and familiarity correlations, Tanaka-Ishii and Terada (2011) report that although words with high familiarity are not necessarily frequent, words with high frequency are necessarily familiar. Their findings also suggest that familiarity ratings highly correlated to that of spoken rather than written language, which is in support of our assumption that familiarity may be an alternative measure of spoken frequency in the present data. The fact that familiarity was a more robust predictor of naming latency than frequency can be attributed to the use of orthographic (written) frequency data in the present dataset due to the lack of spoken frequency corpora for Arabic.

Visual complexity proved to only influence latency in nouns but at a very negligible level, this is in line with previous studies that have established that visual complexity in object naming does not robustly influence naming latency (e.g., Barry et al., 1997; Bonin et al., 2002; Bonin et al., 2003; Cuetos et al., 1999; Snodgrass & Yuditsky, 1996). Visual complexity did not significantly predict latency for verbs.

Furthermore, initial phoneme, syllable and phoneme length do not significantly predict naming latency in both sets of nouns and verbs. The lack of a length effect in the present study is in line with the findings of numerous other studies with healthy speakers (e.g., Alario et al. 2004; Biederman, 1987; Paivio et al., 1989; Snodgrass & Corwin, 1988; Snodgrass & Yuditsky, 1996). For the set of nouns, frequency had no significant effect, this is as in previous findings (e.g., Bonin et al., 2004; Cuetos & Alija, 2003; Schwitter et al., 2004; Shao et al., 2015).

Vigliocco et al. (2011) states that the noun–verb distinction should not be evident in single word processing. The differences between nouns and verbs observed in the present study were differences in psycholinguistic variables influencing single word retrieval, in absence of any higher linguistic structures (phrases or sentences). To be able to test Vigliocco et al.'s claim, an in-depth investigation into the differences between nouns and verbs would need to be carried out at multiple levels: the single-word level, the phrase level, and the clause and sentence level.

The present dataset shows that the primary determinants of naming latency in Gulf Arabic nouns and verbs are agreement (image agreement and name agreement), familiarity (age of acquisition, imageability and familiarity) but not length (syllable and phoneme numbers). Furthermore, the present data show that familiarity (a measure of spoken word frequency, probably) is a much better predictor of naming latency than frequency values extracted from Aralex (Boudelaa & Marslen-Wilson, 2010) which is based on Modern Standard Arabic written forms. There is very little evidence that the naming of verbs and the naming of nouns in Gulf Arabic are affected in different ways by the nine independent variables discussed above. Finally, the set of 319 object drawings and 141 action drawings and their norms are of principal importance for researchers and clinicians working with speakers of Gulf Arabic.

Funding This research was funded by a Qatar National Research Fund (National Priorities Research Program), Grant Number NPRP7-1506-3-390.

Appendix: The naming coding system

Correct: target response is produced.

Target response is produced with a different pronunciation using an alternative allophone; e.g., saying /z d ɑ:r / for /s gɑ:r / "cigarette"

Correct response in Standard Arabic (SA); e.g., saying /f tɑ:/ for /b nt/ "girl"

Correct response in English; e.g., saying /k rt n/ "curtain" for picture of a curtain

Alternative response: production of a response equal in meaning to the target word and can be used interchangeably; e.g., saying /3 rɑb / instead of /gɑ:r / "Baby carriage"

Visual error: Production of a response visually related to the target picture; e.g., saying /bɑ:b/ "door" for /d ri: / "window"

Visual error due to a visual distractor in the presented picture; e.g., saying /m xb / "pocket" for a picture of "trousers with pockets" or /ɣær æ/ "bottle" instead of /j tf h/ "to float" for an action picture of "bottle floating"

Semantic errors: Production of a response semantically related to the target picture. This included six subcategories:

3.1. Semantic super-ordinate error: production of a semantically related error that is super-ordinate to the target response; e.g., saying /h wɑ:n/ "animal" instead of /x r f/ "lamb," or /j nað f/ "to clean" instead of /y sg l/ "to polish"

3.2. Semantic coordinate error: Production of a semantically coordinate response to the target response; e.g., saying /ɣæzɑ:l/ "deer" instead of /z rɑ:fæ/ "giraffe," or /j sbæah/ "to swim" for /jɣ :s/ "to dive"

3.3. Semantic associate error: production of a response that is associated to the target response; e.g., saying /d xɑ:n/ "smoke" instead of /s gɑ:r / "cigarette," or /j xb z/ "to bake" for /j d n/ "to knead"

3.4. Semantic circumlocution error: production of a description of the target word form rather than producing the target word form itself. This included descriptions with a minimum of one content word form; e.g., /hæg- l- ðɑ:f r/ "for the nails," instead of /m bræd/ "nailfiler"

3.4.1 Sentential circumlocution: production of a complete sentence instead of producing the singular target response; e.g., saying /jæbi-j ng ð/ "he wants to rescue" instead of /j ng ð/ "to rescue," or /j hfær- l- ærð- æɑ :n- l-zrɑ: / "he is digging the ground for the plants," instead of "/j hfær/ "to dig"

3.4.2 Phrasal circumlocution: production of a noun/verb phrase by adding a doer/object to the target response; e.g., saying /lɑ: b-k r / "football player" instead of/lɑ: b/ "player," or /j d- l-hæbl/ "to pull the rope" instead of /j d/ or /j shæb/ "to pull"

3.4. Visual circumlocution within a syntactic frame: production of a visual description of the picture in a phrase or sentence; e.g., saying /ɣær æ-f l-mɑ:j/ "A bottle in the water" instead of /j tf h/ "to float" for an action picture of "bottle floating," or /s fi:n -ɣɑ:rg / "A ship sinking" instead of /tæɣr g/ "to sink" for an action picture of "ship sinking"

3.5. Semantic and visual error: production of an inaccurate response that shares semantic and visual features with the target word form such as producing /le m n/ "lemon" instead of /b rt qæl / "orange".

3.6. Semantic and phonological error: Production of an inaccurate response that shared semantic and phonological (share 50% or above of the phonemes of the target response) features with the target response such as producing /hmɑ:r/ "donkey" instead of /hsɑ:n/ "horse".

Phonological error: Production of an inaccurate response that shares 50% or more phonemes with the target response. This included two subcategories:

4.1. Phonological related real word form: when participants produced a phonological error that is a real word form, such as producing /kælb/ "dog" instead of /gælb/ "heart"

. Phonological related word form that is not real: when participants produced a phonological error that resulted in a word that does not exist; e.g., saying / ælɑ:gijæ/ for /z hlɑ:gijæ/ "slide"

4.3. Phonological circumlocution within a syntactic frame: when participant describes the sounds of the target word; e.g., saying /fihæ-hærf- l-gɑ:/ for the target word /wr gæ/ "leaf"

Morpho-syntactic error: production of the target consonantal root with a morpho-syntactic error. This included six subcategories:

5.1. Inflectional error: This subcategory was scored if a participant's inaccurate response was presented with an inflectional error. This was scored if the incorrect number, gender, or person inflections were present, such as producing /m lɑ:jkæ/ [plural noun] "angels" instead of /m lɑ:k/ [singular noun] "angel," or /g t / [masculine noun] "male cat" instead of /g twæ/ [feminine noun] "female cat," or /j ðr bæ/ "to hit him"[3rd person]

5.2 Tense error: production of inaccurate response with a tense error in producing the target response; e.g., saying /tɑ:h/ [past tense]"he fell" for /j ti:h/ "to fall"

5.3 Progressive/Non-progressive error: Production of inaccurate response in a progressive/non-progressive form of the target word; e.g., saying /j nɑ:b h/ [progressive] "barking" instead of /j nbæh/ "to bark," or /j g d/ [non-progressive] instead of /gɑ: d/

5.4 Production of the target word in an incorrect form that implies an object/agent the action is being carried out with, through adding the diacritic / æddæ/ /; e.g., saying /jɣ æss l/ "to wash (object)" instead of /jɣ s l/ "to wash"

5.5. Derivational error: this subcategory was scored if the participant's inaccurate response was presented with a derivational error, such as producing a noun/verb/adjective derived from the same consonantal root of the target response. An example of this would be producing /mhæd bæ/ [adjective] “hair-covered” instead of/h d ɑ:b/ [noun] “hair cover,” or /m bɑ:r zæ/ [noun] “Fencing” instead of /j bɑ:r z/ “to fence” [verb]

5.6. Passivization error: production of a passivized form of the target response; e.g., saying /j n n g/ [passive] “hanged” instead of /j n g/ [active] “to hang”

Unrelated word form: scored if participants produced a real word form that is visually, semantically and phonologically unrelated to the target response, such as producing /d æhhæ/ “watermelon” instead of /sf nd æ/ “sponge”

Tip of the Tongue error: this category included responses in which a participant indicated that they know the name of the object/action but have forgotten it

Don’t know name of object/action error: this category included responses in which a participant indicated that they recognize the object/action but do not know the name.

Don’t know object/action error: this category included responses in which a participant indicated that they do not recognize the object/action.

No Response: Failure to respond to the presented picture within 5 s.

References

Akinina, Y., Malyutina, S., Ivanova, M., Iskra, E., Mannova, E., & Dragoy, O. (2015). Russian normative data for 375 action pictures and verbs. *Behavior Research Methods*, *47*, 691–707. doi:https://doi.org/10.3758/s13428-014-0492-9

Alario, F.-X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*, 531–552. doi:https://doi.org/10.3758/BF03200732

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*, 140–155. doi:https://doi.org/10.3758/BF03195559

Al-Emadidhi, D. (1986). *Lexical and sociolinguistic variation in Qatari Arabic* (Unpublished Ph. D thesis). University of Edinburgh, UK.

Alonso, M. A., Fernandez, A., & Díez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior Research Methods*, *47*, 268–274. doi:https://doi.org/10.3758/s13428-014-0454-2

Bakhtiar, M., & Weekes, B. (2015). Lexico-semantic effects on word naming in Persian: Does age of acquisition have an effect? *Memory & Cognition*, *43*, 298–313.

Balota, D. A., Pilotti, M., & Cortese, J. M. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*, 639–647. doi:https://doi.org/10.3758/BF03200465

Barca, L., Burani, C., & Arduino, S. L. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers*, *34*, 424–434. doi:https://doi.org/10.3758/BF03195471

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology*, *50A*, 560–585. doi:https://doi.org/10.1080/783663595

Bastiaanse, R., & Van Zonneveld, R. (1998). On the relation between verb inflection and verb position in Dutch agrammatic aphasics. *Brain and Language*, *64*, 165–181.

Bastiaanse, R., & Van Zonneveld, R. (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca’s aphasia. *Journal of Neurolinguistics*, *18*, 57–68.

Bates, E. A., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), Language acquisition: The state of the art (pp. 173–218). Cambridge, UK: Cambridge University Press.

Berman, R. A. (2003). Children’s lexical innovations: Developmental perspectives on Hebrew verb structure. In J. Shimron (Ed.), *Language processing and acquisition in language of Semitic, root-based, morphology* (pp. 243–291). Benjamins.

Berndt, R. S., Haengiges, A., Burton, M., & Mitchum, C. (2001). Grammatical class and imageability in aphasic word production: Their effects are independent. *Journal of Neurolinguistics*, *15*, 353–371.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147. doi: https://doi.org/10.1037/0033-295X.94.2.115

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*, 73–79. doi:https://doi.org/10.3758/BF03195349

Bird, H., Howard, D., & Franklin, S. (2000). Why is a verb like an inanimate object? Grammatical category and semantic category deficits. *Brain and Language*, *72*, 246–309.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1.17) [Computer program]. Retrieved January, 2015, from www.praat.org

Bonin, P., Boyer, B., Méot, A., Fayol, M., & Droit, S. (2004). Psycholinguistic norms for action photographs in French and their relationships with spoken and written latencies. *Behavior Research Methods, Instruments, & Computers*, *36*, 127–139. doi:https://doi.org/10.3758/BF03195558

Bonin, P., Méot, A., Chalard, M., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*, 89–114.

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, *35*, 158–167. doi:https://doi.org/10.3758/BF03195507

Boudelaa, S., & Marslen-Wilson, W. (2010). Aralex: A lexical database for Modern Standard Arabic. *Behavior Research Methods*, *42*, 481–487. doi:https://doi.org/10.3758/BRM.42.2.481

Boukadi, M., Zouaidi, C., & Wilson, M. A. (2016). Norms for name agreement, familiarity, subjective frequency, and imageability for 348 object names in Tunisian Arabic. *Behavior Research Methods*, *48*, 585–99. doi:https://doi.org/10.3758/s13428-015-0602-3

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*, 208–216. doi: https://doi.org/10.3758/BF03197718

Cameirão, M. L., & Vicente, S. G. (2010). Age-of-acquisition norms for a set of 1,749 Portuguese words. *Behavior Research Methods*, *42*, 474–480. doi:https://doi.org/10.3758/BRM.42.2.474

Caramazza, A., & Hillis, A. (1991). Lexical organisation of nouns and verbs in the brain. *Nature*, 349, 788–790.

Cuetos, F., & Alija, M. (2003). Normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers*, *35*, 168–177. doi:https://doi.org/10.3758/BF03195508

Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research*

Methods, Instruments, & Computers, *31*, 650–658. doi:https://doi.org/10.3758/BF03200741

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*, 171–237. doi:https://doi.org/10.1006/jecp.1996.2356

Damasio, A. R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy for Sciences*, *90*, 4957–4960.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838. doi:https://doi.org/10.1037/0033-295X.104.4.801

d'Honincthun, P., & Pillon, A. (2008). Verb comprehension and naming in frontotemporal degeneration: The role of the static depiction of actions. *Cortex*, *44*, 834–847. doi:https://doi.org/10.1016/j.cortex.2007.04.003

Dupoux, E., & Mehler, J. (1990). Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Journal of Memory and Language*, *29*, 316–335

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 515–523. doi:https://doi.org/10.1037/0278-7393.24.2.515

Eviatar, Z., Menn, L., & Zaidel, E. (1990). Concreteness: Nouns, verbs, and hemispheres. *Cortex, 26,* 611–624.

Feyereisen, P., Van der Borght, F., & Seron, X. (1988). The operativity effect in naming: A re-analysis. *Neuropsychologia*, *26*, 791.

Gernsbacher, M. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256–281. doi:https://doi.org/10.1037/0096-3445.113.2.256

Ghazali, S., Hamdi, R., & Knis, K. (2007). Intonational and rhythmic patterns across the Arabic dialect continuum. In E. Benmamoun (Ed.), Perspectives on Arabic linguistics XIX: Current issues in linguistic theory (pp. 97–121). Amsterdam, The Netherlands: Benjamins.

Gleitman, L. (1994). The structural sources of verb meanings. In P. Bloom (Ed.), Language acquisition (pp. 174–221). Cambridge, MA: MIT Press.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language* 24, 631–645.

Holes, C. (2004). Quadriliteral verbs in the Arabic dialects of Eastern Arabia. In M. Haak, R. de Jong, & K. Versteegh (Eds.), Approaches to Arabic dialects: a collection of articles presented to Manfred Woidich on the occasion of his 60th birthday (pp. 97–116). Brill: Leiden.

Huttenlocher, J., & Lui, F. (1979). The semantic organization of some simple nouns and verbs. *Journal of Verbal Learning and Verbal Behavior*, *18*, 141–162.

Johnstone, T. M. (1967). Eastern Arabian dialect studies. London, UK: Oxford University Press.

Khwaileh, T., Body, R., & Herbert, R. (2014). A normative database and determinants of lexical retrieval for 186 Arabic nouns: Effects of psycholinguistic and morpho-syntactic variables on naming latency. *Journal of Psycholinguistic Research*, *43*, 749–769. doi:https://doi.org/10.1007/s10936-013-9277-z

Khwaileh, T., Body, R., & Herbert, R. (2017). Lexical retrieval after Arabic aphasia: Syntactic access and predictors of spoken naming. *Journal of Neurolinguistics*, *42*, 140–155.

Khwaileh, T. A., Mustafawi, E., Howard, D., & Herbert, R. (2016). An aphasia battery for Qatari/Gulf Arabic (Abstract). *Frontiers in Psychology 54th Annual Academy of Aphasia Meeting Conference Abstracts*, 26. doi:https://doi.org/10.3389/conf.fpsyg.2016.68.00026

Kosslyn, S. M., & Chabris, C. F. (1990). Naming pictures. *Journal of Visual Languages and Computing*, *1*, 77–95.

Kreuz, R. (1987). The subjective familiarity of English homophones. *Memory & Cognition*, *15*, 154–168.

Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning and Verbal Behavior*, *13*, 613–625.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38, disc. 38–75.

Marslen-Wilson, W. (1990). Activation, competition and frequency in lexical access. In G. Altmann (Ed.), Cognitive models of speech processing (pp. 148–172). Cambridge, MA: MIT Press.

Martein, R. (1995). Norms for name and concept agreement, familiarity, visual complexity and image agreement on a set of 216 pictures. *Psychologica Belgica*, *35*, 205–225.

Masterson, J., Druks, J., & Gallienne, D. (2008). Object and action picture naming in three- and five-year-old children. *Journal of Child Language*, *35*, 373–402.

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: Past studies and new evidence. *Cortex*, *45*, 738–758. doi:https://doi.org/10.1016/j.cortex.2008.10.003

Miceli, G., Silveri, M. C., Villa, G., & Caramazza, A. (1984). On the basis of the agrammatics' difficulty in producing main verbs. *Cortex*, *20*, 207–220.

Meschyan, G., & Hernandez, A. (2002). Age of acquisition and word frequency: Determinants of object-naming speed and accuracy. *Memory & Cognition*, *30*, 262–269. doi:https://doi.org/10.3758/BF03195287

MRC Psycholinguistic Database. (2006). Retrieved from www.psy.uwa.edu.au/mrcdatabase/uwamrc.htm

Mustafawi, E. (2018). The phonology of Arabic. In E. Benmamoun & R. Basyouni, *Handbook on Arabic linguistics*. Basingstoke: Routledge.

Nickels, L. (1997). Spoken word production and its breakdown in aphasia. Hove, UK: Psychology Press.

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, *33*, 1281–1303.

Nusbaum, H. C., Pisoni, D., & Davis, C. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words In Research on Speech Perception, Progress Report 10 (pp. 357–376). Bloomington, IN: Indiana University.

Paivio, A., Clark, J. M., Digdon, N., & Bons, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition*, *17*, 163–174. doi:https://doi.org/10.3758/BF03197066

Paivio, A., & Yarmey, A. D. (1966). Pictures versus words as stimuli and responses in paired-associate learning. *Psychonomic Science*, *5*, 235–236.

Pallant, J. (2005). SPSS survival manual. Milton Keynes, UK: Open University Press.

Pinker, S. (1994). The language instinct: how the mind creates language. New York, NY: Morrow.

Raman, I., Raman, E., & Mertan, B. (2014). A standardized set of 260 pictures for Turkish: Norms of name and image agreement, age of acquisition, visual complexity, and conceptual familiarity. *Behavior Research*, *46*, 588–595. doi:https://doi.org/10.3758/s13428-013-0376-4

Sapir, E. (1921). Language: An Introduction to the study of speech. New York, NY: Harcourt, Brace, & Co.

Schock, J., Cortese, M. J., & Khanna, M. M. (2012). Imageability estimates for 3,000 disyllabic words. *Behavior Research Methods*, *44*, 374–379. doi:https://doi.org/10.3758/s13428-011-0162-0

Schwitter, V., Boyer, B., Méot, A., Bonin, P., & Laganaro, M. (2004). French normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers*, *36*, 564–576. doi:https://doi.org/10.3758/BF03195603

Scott, K. S. (2006). Language processing: The neural basis of nouns and verbs. *Current Biology*, *16*, 295–296.

Segui, J., Mehler, J., Frauenfelder, U., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia*, *20*, 615–627.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *Quarterly Journal of Experimental Psychology*, *65*, 1927–1944.

Shao, Z., Roelofs, A., & Meyer, A. S. (2013). Predicting naming latencies for action pictures: Dutch norms. *Behavior Research Methods*, *46*, 274–283. doi:https://doi.org/10.3758/s13428-013-0358-6

Shao, Z., Roelofs, A., Martin, R. C., & Meyer A. S. (2015). Selective inhibition and naming performance in semantic blocking, picture-word interference, and color-word Stroop tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1806–20.

Snodgrass, J. G., & Corwin, J. (1988). Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. *Perceptual and Motor Skills*, *67*, 3–36.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215. doi:https://doi.org/10.1037/0278-7393.6.2.174

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, *28*, 516–536. doi:https://doi.org/10.3758/BF03200540

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, *38*, 598–605. doi:https://doi.org/10.3758/BF03193891

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York, NY: Allyn & Bacon.

Tanaka-Ishii, K., & Terada, H. (2011). Word familiarity and frequency. *Studia Linguistica*, *65*, 96–116.

Versteegh, K. (1997). The Arabic language. New York, NY: Columbia University Press.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, *35*, 407–426. doi:https://doi.org/10.1016/j.neubiorev.2010.04.007

Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *Quarterly Journal of Experimental Psychology*, *48A*, 822–848. doi:https://doi.org/10.1080/14640749508401419

Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*, *39*, 14–32.