Journal of Arabic and Human Sciences
Qassim University, Vol. 13, No. 1, PP 15-25. (September 2019/ muharam1439)

Frequency Effects on Auditory Lexical Decision in Modern Standard Arabic

Dr. Faisal m. Aljasser
Assistant Professor
Department of English & Translation, Qassim University
jasr@qu.edu.sa

Abstract: Previous studies have shown that Word Frequency strongly affects spoken word recognition in various ways. Most of these studies computed Word Frequency based on written databases for languages in which written words closely match spoken ones. The present study investigates whether the Word Frequency effect still stands in Arabic, a language with a diglossic situation where the formal written discourse and spoken one are far from close correspondence. Using an Auditory Lexical Decision task, Arabic native speakers were asked to decide as quickly and as accurately as possible whether an orally presented stimulus was a real word or a non-word in Arabic. Results showed that subjects were affected by Word Frequency. They decided that high frequency stimuli items are real words faster and more accurately than low frequency ones. Implications of these findings to spoken word recognition of diglossic Arabic are discussed.

Introduction and literature review

Word recognition is a process that is affected by multiple interacting factors. One such robust effect that has long been established in the literature is the word frequency (WF) effect. In other words, words that occur more frequently in the language are recognized faster and more accurately. For example, subjects' lexical decisions are faster when visually presented words have high frequency (e.g. Whaley, 1978). In the visual lexical decision (LD) task subjects are required to decide as quickly and accurately as possible if a visually presented letter string is a word or a nonword (see Goldinger (2017) for full description).

Doing an analysis on the LD reaction times (RTs) of the 40 thousand words from the English Lexicon Project (Balota et al., 2007), Brysbaert et al. (2011) found out that, surpassing all other variables, WF accounted for around 41% of the variance in latencies. Similarly, WF seems not only to affect LD but also spoken word production. It has been shown that subjects name visually presented high frequency words faster than low frequency ones (Balota & Chumbley, 1984; Forster & Chambers, 1973).

In addition to its effects on spoken word production, other researchers have found WF effects on written word production. Barry & Seymour (1988) for example have shown that WF affects spelling accuracy. Similarly, in spelling- to- dictation tasks, it has been shown that participants typically initiate the spelling of high frequency words faster than those of low frequency ones (Bonin, Fayol, & Chalard, 2001; Bonin, Fayol, & Gombert, 1998).

Despite the fact that much less research has been conducted to investigate WF effects on spoken word recognition, this research has shown that these effects are equally robust as those in visual word recognition. For example, Grosjean (1980) used the gating paradigm to investigate the effect of WF on spoken word recognition. In this paradigm, subjects get repeated presentations of a spoken stimulus with the duration of presentation time increasing at each successive pass. After each pass, the subject has to guess the stimulus being presented and to provide a confidence rating of each guess. (See Grosjean (1980) for full description). Grosjean (1980, p. 268) used the isolation point "acoustic-phonetic information needed from the onset of the word to the point at which it is isolated from other words" as a measure of the effect of frequency. What he found out was that high frequency words needed less sensory input and therefore were isolated faster than low frequency ones (see also Tyler (1984) for similar results using the same task).

Similarly, Taft & Hambly (Taft & Hambly, 1986) investigated the effect of WF on spoken word recognition. However, unlike Grosjean (1980) and Tyler (1984) who used the gating task, Taft & Hambly (Taft & Hambly, 1986) used the auditory (LD) task. They reasoned that in the case of the gating task, not being an online task, any observed effects of WF cannot be attributed directly to word recognition as such but perhaps to mere frequency biases in the guessing of the words. Their study will be discussed in more detail here as it is closely related to the current study.

In their experiment 4, Taft & Hambly (Taft & Hambly, 1986) used the frequency norms of Carroll, Davies & Richman (1971). The stimuli items were arranged in 24 pairs where each pair had a low frequency word and a high frequency one. Both words in each pair started with the same few phonemes and had the same uniqueness point (i.e. the phoneme at which the word becomes the only possible word candidate (e.g. /k/ in difficult /dɪfɪkəlt/). 20 pairs of nonword items were also used. Both nonwords in each pair were of the same length of the word items. Also, both nonwords deviated from real words at the same point (i.e. phoneme). The procedure followed a typical auditory LD task procedure. That is, their subjects (N= 15) listened to the randomly ordered stimuli (i.e. words and nonwords) separated by a two-second delay and they had to judge whether the item is a real word or not as quickly and accurately as possible by pressing a "yes" or a "no" button. RTs and error rate were recorded for analysis.

Taft & Hambly (Taft & Hambly, 1986) found out that RTs were significantly shorter and error rates were significantly lower for high frequency words than for low frequency ones. These results were taken as evidence that WF aided the recognition of spoken words. (See also Meunier & Segui (1999) for similar results in French).

How does repeated exposure to a particular word make it more quickly and/or accurately recognizable? Models of spoken word recognition (see Weber & Scharenborg (2012) for a review) posit that words compete for recognition in lexical access. Moreover, most of these models include the WF effect as a crucial part of the word recognition process. The Cohort Model (Marslen-Wilson & Welsh, 1978), for example, posits that a word is recognized at its "recognition point". The more phonemes are perceived, the more competing candidate words are reduced successively. Taft & Hambly (Taft & Hambly, 1986) provide the word "crocodile" as an example. When /krɒkə/ is perceived the word "crocodile" will still be competing with words that share the same phoneme string such as "crockery". However, once /krɒkəd/ is perceived the word "crocodile" will be recognized because it is the only remaining word in the cohort of possible word candidates. The more developed version of the Cohort Model (Marslen-Wilson, 1987) accounts for the frequency effects by suggesting that candidate words have resting activation values. High frequency words have higher values than low frequency ones and therefore can reach the threshold of recognition faster.

It is evident from the previous review that the effect of WF on different aspects of word recognition is very well established in the literature. However, all of the studies on WF discussed above and many others have dealt with first languages (e.g. English (Taft & Hambly, 1986), Dutch (Ghyselinck, Lewis, & Brysbaert, 2004) and French (Meunier & Segui, 1999).

Findings from studies on second language (L2) indicate that WF effects are stronger in the L2 than in first language (L1). These include studies investigating visual word production (e.g. Van Wijnendaele & Brysbaert, 2002) and visual word recognition (Duyck, Vanderelst, Desmet, & Hartsuiker, 2008; Gollan et al., 2011). For example, Van Wijnendaele & Brysbaert (2002) compared the performance of Dutch-French and French- Dutch bilinguals in naming words in their L1 and L2. Both groups showed generally slower naming times in the L2. However, both groups showed stronger WF effects in the L2 than in the L1. Similarly, Duyck et al. (2008) used a visual lexical decision task, in which written words were presented, to compare WF effects in the L1 and L2 in Dutch-English bilinguals. What they found out was that WF effect in the L2, as shown in RTs and error rate, was about twice as large as that in the L1.

These studies show that although WF effect is clearly evident in both L1 and L2, it is stronger in L2. However, all these did not investigate a language that has a diglossic situation (see Ferguson (1959) for a discussion of diglossia) where the standard variety of the language is radically different from the spoken one. Studying frequency effect on spoken word recognition in a language that has this diglossic situation will shed light on the interaction of these two varieties in the native speaker's mind.

One such language is Arabic. Arabic native speakers use two varieties; Modern Standard Arabic (MSA) and colloquial Arabic. MSA is mainly used for written and formal spoken communications. Arabic-speaking children start to be fully exposed to MSA when they start primary school where teaching is mainly in MSA. Before joining school, however, they typically get relatively repeated exposure to MSA watching animation programs or hearing TV and radio news broadcasts. Despite this pre-school exposure, some researchers argue that MSA is learned as a second language (e.g. Ayari, 1996). A colloquial variety, on the other hand, is the regional dialect spoken at home and used for everyday informal spoken communication and is acquired as a first language.

The interaction between the two varieties in the native speaker's mind has recently started to draw the interest of researchers (Boudelaa & Marslen-Wilson, 2013; Ibrahim & Aharon-Peretz, 2005). Given this diglossic nature of Arabic language, the current study endeavors to investigate the question of whether Arabic native speakers are sensitive to WF in MSA and if this sensitivity is manifested in spoken word recognition of MSA words. The magnitude of WF effects in spoken word recognition of MSA will shed light on the status of MSA in the native Arabic speaker's mind. Is MSA processed by native Arabic speakers as a second language or as effectively as a first language?

The experiment

Method

An auditory LD task was used in the current experiment to explore the effect of WF on Arabic language spoken word recognition. In the auditory LD task, subjects listen to a stimulus item and their task is to judge as quickly and as accurately as possible if the item is a real word in the target language or a nonsense word (non-word). This task is particularly suited for the current investigation. Firstly, it has been shown to be a good measure of the effect of different variables on lexical processing (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Secondly and more importantly, it has been shown that LD RTs are quite sensitive to WF both in visual word recognition (e.g. Balota et al., 2004) and spoken word recognition (Taft & Hambly, 1986).

Participants

Twenty-four native Arabic speakers, all students at a university in the united states, took part voluntarily in the experiment. All participants were born and studied elementary/secondary school in their home countries (in Arabic). Most of them had different Arabic dialects as their first language. Their mean age was 27 years old.

Materials

The first step in selecting the stimuli was to choose a large Arabic language corpus that includes a frequency of occurrence wordlist. Ar ten ten (Arts, Belinkov, Habash, Kilgarriff, & Suchomel, 2014), an Arabic corpus that contains 5.8 billion words was selected. Not only is it large, providing reliable frequency counts, but it is also quite recent. Importantly, a subset of Ar ten ten containing 115 million words was processed with MADA (Habash, Rambow, & Roth, 2009). MADA is an Arabic language-processing tool that uses a morphological analyzer for MSA that works at disambiguating undiacritized MSA words in context by reaching a preferred analysis. This is quite important for disambiguating Arabic homographs when diacritics for short vowels and gemination are not used. The MADA processed 115-million-word subset was loaded to a corpus manager called The Sketch Engine (Kilgarriff, Rychly, Smrz, & Tugwell, 2004)

A frequency wordlist containing 100000 words was purchased from Sketch Engine (https://www. sketchengine.co.uk/). The words in the frequency wordlist were analyzed in fully vowelized Buckwalter transliteration (Buckwalter, 2002).

The stimuli consisted of 48 MSA words varying in frequency. These 48 words were divided into two groups. Half of the words (24 items) are high frequency words (mean frequency = 136) and the other half are low frequency ones (mean frequency = 2.7). The frequency counts reported here are given per-million calculated on the 115-million-word subset corpus.

Both groups of 24 words were matched on word class and number of phonemes and syllables. In addition, they were all nouns and had five phonemes and two syllables and the form /CVCVC/. Moreover, in order to be able to confidently attribute any significant differences in RTs or error rates in the current LD task to frequency effects, all words were controlled on a number of variables that have been shown to affect spoken word recognition. The variables computed include acoustic duration (i.e. duration of the word stimulus sound recording in millisecond)[1], neighborhood density “the number of known words that sound similar to a given word” (Han, Storkel, Lee, & Yoshinaga-Itano, 2014, p. 243)[2], Phonotactic Probability "the frequency with which phonological segments and sequences of phonological segments occur in words in a given language" (Vitevitch & Luce, 2005, p. 193) [3] , and uniqueness point (i.e. the phoneme at which the word becomes the only possible word candidate (e.g. phoneme number five /dˀ/ in /maχaːdˀ/ (throe). Table 1 shows the means and standard deviations for all aforementioned variables.

[1] The duration of the initial silence was fixed to 50ms in all stimuli files.

[2] This is computed by counting the number of words in the database that differ from the target word by only one phoneme addition, deletion or substitution. I am thankful to Prof. Michael Vitevitch for computing Neighborhood Density for my stimuli items.

[3] Phonotactic probability was computed using (Aljasser & Vitevitch, 2018) phonotactic probability calculator for Arabic language.

Table 1 Means and standard deviations of frequency, duration, Neighborhood Density, positional and biophone phonotactic probability and uniqueness point for high vs. low frequency stimuli words

	Mean frequency (SD)	Mean duration of sound file in ms (SD)	Mean Neighborhood Density (SD)	Mean Positional phonotactic probability (SD)	Mean biophone phonotactic probability (SD)	Mean Uniqueness point (SD)
High Frequency words (N=24)	136 (150)	588.9 (71.1)	2.1 (1.8)	1.525633 (0.224011)	1.037613 (0.0191)	4.8 (1.1)
Low Frequency words (N= 24)	2.7 (1.2)	611.3 (54.36324)	1.9 (1.81579)	1.51195 (0.234666)	1.035592 (0.022699)	4.6 (0.71)

Analysis of Variance (ANOVA) showed that WF for high frequency words was significantly greater than for low frequency words $p < 0.0001$. However, the analysis showed that all the other lexical characteristics for high vs. low frequency words were not significantly different $p > 0.05$ showing that all these variables were very well-controlled.

A group of 48 nonwords were also created by replacing one phoneme from each of the 48 real words. This method ensured that the created nonwords closely resemble real words in phonological structure.

The stimuli items were spoken in isolation and recorded by a male native Arabic speaker in an anechoic chamber using a high-quality microphone on to digital-audio-tape at a sampling rate of 44.1 kHz. The recordings were then saved as digital 16- bit files on a computer disk.

Procedure

An iMac computer running PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) was used for the experiment presentation and data collection. PsyScope is "an integrated environment for designing and running psychology experiments on Macintosh computers" (Cohen et al., 1993, p. 257). Participants were tested individually and one at a time. Each participant was seated in a computer booth equipped with a PsyScope button box and a set of Beyerdynamic DT-100 headphones. Prior to running the experiment, the instructions appeared on the computer screen in Arabic. Participants were instructed that they will listen to stimuli items and that their task is to judge which ones are real Arabic words and which ones are pseudowords by pressing the relevant colored button on the button box as quickly and as accurately as possible. All participants' inquiries were answered prior to the start of the experiment.

Each trial started with the Arabic word مستعد "READY" appearing in the center of the computer screen for 500 ms to indicate the beginning of the trial. Prior to the experimental trials, each participant received 10 practice trials. These trials were used to familiarize the participants with the task and were not included in the final data analysis. The participants were then presented with one of the randomly selected stimuli at a comfortable listening level over the headphones. The left button on the response box was for nonword selection and the right button (i.e., that for the dominant hands of the participants as all were right-handed) was for word selection. The same color of the two buttons on the button box flashed on the screen after each response with the (word vs non-word) labels in Arabic " ليست كلمة \كلمة " to remind subjects what the relevant button is. The participants responded as quickly and accurately as possible by pushing the appropriately labeled button. After each response 1500 ms elapsed before the next token was played. RT was measured from the onset of the stimulus file to the onset of the response.

Results

RTs for correct responses measured from the stimulus onset and error rates for the real words were collected. An error was counted when the subject clicked “non-word”for a real word. Table 2 shows means and STDs of RTs in milliseconds and % error rates for High frequency vs. Low frequency words.

Table 2 Means and STDs of subjects’ (N=24) RTs in milliseconds and % error rates for High frequency vs. Low frequency words

Frequency type	Mean RT in ms (STD)	Mean % Error Rate (STD)
High Frequency words	902 (82)	1.26% (2.32)
Low Frequency words	1059 (94)	10.85% (8.48)

T tests on RTs were carried out and revealed that high frequency words were responded to significantly faster than low frequency ones $p < 0.0001$. Moreover, T test carried out on error rate showed that high frequency words were responded to significantly more accurately than low frequency ones $p < 0.0001$.

Discussion

The current study set out to investigate whether WF measured using a written corpus can affect spoken word recognition in Arabic. Recall that Arabic has a diglossic situation (Ferguson, 1959) where two different varieties are used. MSA is used for written and formal spoken communications. Colloquial Arabic on the other hand is used for all other informal communications. The current WF counts used were purely based on MSA. That is, the MADA processed 115-million-word subset used in the current study targets only words used in MSA contexts.

The results of the present study are clear. Native Arabic speakers’ spoken word recognition was affected by WF in MSA. This WF effect has manifested itself very clearly and strongly in the online task used (i.e. the LD task). High frequency words were perceived faster and more accurately than low frequency ones.

This sensitivity to WF in MSA cannot be simply attributed to matched frequency counts between MSA and the colloquial variety. Our subjects were from different geographical areas and therefore spoke different vernaculars. Ferguson (1959) observed that although most of the vocabulary of the standard variety and the vernacular are shared, they vary in form and differ in use and meaning. In other cases, the standard and the vernacular have paired vocabulary items for shared concepts(Ferguson, 1959, p. 334) This again rules out the possibility that our subjects’ sensitivity to WF in MSA is a mere sensitivity to matched frequency values for the same words in their vernaculars.

Similarly, the current findings cannot be attributed to other factors that can potentially confound the results. In the current study, all factors that have been shown to affect spoken word recognition were well-controlled. These included acoustic duration, phonotactic probability, neighborhood density, and isolation point.

A more plausible explanation for the current findings is that when orally presented MSA is processed as effectively as the dialect (i.e. the first language). Recall that the difference between high frequency words and low frequency ones in mean latencies (RTs) is 157 ms. This is comparable to the difference in means by native English speakers found in Taft & Hambly’s (Taft & Hambly, 1986) study discussed earlier. In their study they found that native English speakers were on average 174 ms faster when accepting high frequency words. Not only does the difference between RTs for high vs. low frequency words closely resemble native-like performance but also the overall RT is similar to that produced by other native speakers of other languages. In Meunier & Segui’s (1999) study, native French speakers RTs for high frequency words were 925 ms and RTs for low frequency words were 1018 ms which closely resembles the pattern in the current findings (902 ms and 1059 ms, respectively). This provides further evidence that MSA is processed as effectively as a native language.

Moreover, this is in line with the findings of Boudelaa & Marslen-Wilson (2013). In their study they set out to investigate whether roots and word pattern play a different role in the processing of MSA and the spoken dialect (Southern Tunisian Arabic (STA)). Using auditory priming experiments on subjects who spoke STA as a first language they found out that, in the two varieties, subjects RTs and Error rates were comparable in terms of relying on roots and word patterns as the main linguistic units of processing. They relied on this result to argue that describing MSA simply as a "second language" for native Arabic speakers may not be an accurate description. They maintained that the fact that the age at which MSA is acquired is typically within the critical period (before puberty) and this qualifies MSA to be processed in a native-like manner.

The current findings lend support to this claim using one of the most established effects in spoken word recognition (i.e. WF effect). Obviously, our subjects showed native-like processing capacity of MSA. The sociolinguistic context in which MSA is acquired characterized by early childhood exposure and continuous exposure later in life may be the leading cause of this native-like performance in spoken word recognition.

Conclusion

The current study has provided evidence that native Arabic speakers show native-like sensitivity to WF when recognizing spoken words in MSA. However, an important question remains for further research. Does this native-like processing of spoken words in MSA also transfer to spoken word production? Or does the latter entail a different mechanism that prevents native Arabic speakers from behaving native-like in spoken word production of MSA, given that they perform their daily life communications using the spoken dialect?

References

[1] Aljasser, F., & Vitevitch, M. S. (2018). A Web-based interface to calculate phonotactic probability for words and nonwords in Modern Standard Arabic. *Behavior Research Methods, 50*(1), 313–322. https://doi.org/10.3758/s13428-017-0872-z

[2] Arts, T., Belinkov, Y., Habash, N., Kilgarriff, A., & Suchomel, V. (2014). ArTenTen: Arabic corpus and word sketches. *Journal of King Saud University - Computer and Information Sciences*, *26*(4), 357–371. https://doi.org/10.1016/j.jksuci.2014.06.009

[3] Ayari, S. (1996). Diglossia and illiteracy in the arab world. *Language, Culture and Curriculum*. https://doi.org/10.1080/07908319609525233

[4] Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. Retrieved from http://psych.wustl.edu/coglab/Publications_files/BalotaChumbley1984.pdf

[5] Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

[6] Balota, D. a, Yap, M. J., Cortese, M. J., Hutchison, K. a, Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

[7] Barry, C., & Seymour, P. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *The Quarterly Journal of Experimental Psychology*, *40*(1), 5–40. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/14640748808402280?needAccess=true

[8] Bonin, P., Fayol, M., & Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *54*(2), 469–489. https://doi.org/10.1080/02724980042000219

[9] Bonin, P., Fayol, M., & Gombert, J. E. (1998). An experimental study of lexical access in the writing and naming of isolated words. *International Journal of Psychology*, *33*(4), 269–286. https://doi.org/10.1080/002075998400312

[10] Boudelaa, S., & Marslen-Wilson, W. D. (2013). Morphological structure in the Arabic mental lexicon: Parallels between standard and dialectal Arabic. *Language and Cognitive Processes*, *28*(10), 1453–1473. https://doi.org/10.1080/01690965.2012.719629

[11] Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

[12] Buckwalter, T. (2002). Arabic Transliteration. Retrieved from www.qamus.org/ transliteration.htm

[13] Carroll, J. B. (John B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. Houghton Mifflin.

[14] Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, *25*(2), 257–271. https://doi.org/10.3758/BF03204507

[15] Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review*, *15*(4), 850–855. https://doi.org/10.3758/PBR.15.4.850

[16] Ferguson, C. A. (1959). Diglossia. *Word*, *15*, 325–340. https://doi.org/10.1080/00437956.1959.11659702

[17] Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*, 62–635. Retrieved from http://ac.els-cdn.com/S0022537173800428/1-s2.0-S0022537173800428-main.pdf?_tid=78c53760-872d-11e7-a4d4-00000aab0f01&acdnat=1503401683_9bcbf480a3502d82d6bf6337473f283d

[18] Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*, 43–67. https://doi.org/10.1016/j.actpsy.2003.11.002

[19] Goldinger, S. D. (2017). Auditory lexical decision. *Language and Cognitive Processes*, *0965*(July). https://doi.org/10.1080/016909696386944

[20] Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

[21] Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*(4), 267–283. https://doi.org/10.3758/BF03204386

[22] Habash, N., Rambow, O., & Roth, R. (2009). MADA+TOKAN: A Toolkit for Arabic Tokenization, Diacritization, Morphological Disambiguation, POS Tagging, Stemming and Lemmatization. In *the 2nd International Conference on Arabic Language Resources and Tools (MEDAR)* (pp. 102–109). Cairo: MEDAR. Retrieved from https://pdfs.semanticscholar.org/8dcc/f1749b25bf21d9208237fd2e68299615e4e0.pdf

[23] Han, M. K., Storkel, H. L., Lee, J., & Yoshinaga-Itano, C. (2014). The influence of word characteristics on the vocabulary of children with cochlear implants. *Journal of Deaf Studies and Deaf Education*, *20*(3), 242–251. https://doi.org/10.1093/deafed/env006

[24] Ibrahim, R., & Aharon-Peretz, J. (2005). Is literary Arabic a second language for native Arab speakers?: Evidence from semantic priming study. *Journal of Psycholinguistic Research*, *34*(1), 51–70. https://doi.org/10.1007/s10936-005-3631-8

[25] Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). The Sketch Engine. In *Proceedings of Euralex*, (pp. 105–116). Lorient. Retrieved from http://www.itri.brighton.ac.uk

[26] Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*, 71–102. Retrieved from https://ac.els-cdn.com/0010027787900059/1-s2.0-0010027787900059-main.pdf?_tid=3316a4ca-eafa-11e7-b497-00000aab0f02&acdnat=1514374778_441c69241b5d96fa94e65b12c159afc3

[27] Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. https://doi.org/10.1016/0010-0285(78)90018-X

[28] Meunier, F., & Segui, J. (1999). Frequency effects in auditory word recognition: The case of suffixed words. *Journal of Memory and Language*, *41*, 327–344. Retrieved from https://ac.els-cdn.com/S0749596X99926426/1-s2.0-S0749596X99926426-main.pdf?_tid=55d059ac-a450-11e7-8669-00000aacb35d&acdnat=1506605248_23d05d64f4ffc48cd9d40875e4e057ff

[29] Taft, M., & Hambly, G. (1986). Exploring the Cohort Model of spoken word recognition*. *Cogniton*, *22*, 259–282. Retrieved from https://ac.els-cdn.com/001002778690017X/1-s2.0-001002778690017X-main.pdf?_tid=608893a8-4430-46e6-b2fd-5fc2c0386747&acdnat=1526997534_a578c1e85a572d27977c96f83edebb4a

[30] Tyler, L. K. (1984). The structure of the initial cohort: Evidence from gating. *Perception & Psychophysics*, *36*(5), 417–427. Retrieved from https://link-springer-com.ezproxy.qu.edu.sa/content/pdf/10.3758%2FBF03207496.pdf

[31] Van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(3), 616–627. https://doi.org/10.1037//0096-1523.28.3.616

[32] Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*(2), 193–204. https://doi.org/10.1016/j.jml.2004.10.003

[33] Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 387–401. https://doi.org/10.1002/wcs.1178

[34] Whaley, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 143–154. https://doi.org/10.1016/S0022-5371(78)90110-X

Appendix

High frequency words	HF words transcription	Low frequency words	LF words transcription
شيطان	$YTAn	شقاق	$iqAq
دولة	dWlah	ديدن	dYdan
عالم	EAlam	عجاب	EujAb
عذاب	Ea*Ab	عتاد	EatAd
فساد	fasAd	فراء	firA?
غاية	gAyah	غيار	giyAr
غياب	giyAb	غشاء	gi$A?
هدف	hadaf	هوس	hawas
جمال	jamAl	جزاف	juzAf
مكان	makAn	مخاض	maxAD
رسول	rasUl	مدين	madIn
سلام	salAm	صداق	SadAq
صورة	SUrah	سياج	siyAj
سوال	su?Al	سجال	sijAl
صراع	SirAE	نشوب	nu$Ub
سلاح	silAH	سعال	suEAl
تاريخ	tArIx	توطيد	tWTId
طاقة	Taqah	توبيخ	tWbIx
توحيد	tWHId	زخم	zaxam
وطن	waTan	ولوج	wulUj
خليج	xalIj	خناق	xinAq
خلاف	xilAf	خاطر	xATir
خروج	xurUj	نواة	nawAh
خطأ	xaTa?	خور	xawar

أثر تكرار الكلمة على سرعة ودقة التعرف عليها في اللغة العربية المعاصرة

د. فيصل بن محمد الجاسر

الأستاذ المساعد بقسم اللغة الانجليزية والترجمة بجامعة القصيم

ملخص البحث: أظهرت الدراسات السابقة أن تكرار أو شيوع الكلمة في اللغة يؤثر بشدة وبطرق مختلفة على القدرة على التعرف على الكلمة المنطوقة. ولقد قامت معظم هذه الدراسات بالاعتماد على قواعد بيانات مكتوبة للغات تتطابق فيها الكلمات المكتوبة مع الكلمات المنطوقة. وتبحث الدراسة الحالية ما إذا كان تأثير تكرار الكلمة لا يزال قائماً باللغة العربية ، وهي لغة تعتمد على الاستخدام المزدوج لشكلين مختلفين من أشكال اللغة المكتوبة والمنطوقة. ففي الدراسة الحالية طلب من الناطقين باللغة العربية أن يقرروا بسرعة وبدقة ما إذا كان الحافز المقدم شفهياً كلمة حقيقية أو غير حقيقية في اللغة العربية. أظهرت النتائج، والتي تم قياسها باستخدام برنامج حاسوبي يقيس ردة الفعل بالجزء من الألف من الثانية، أن سرعة ودقة اتخاذ القرار تتأثر بمدى تكرار الكلمة في اللغة حيث قررت العينة أن الكلمات كثيرة التكرار هي كلمات حقيقية بشكل أسرع وأكثر دقة من تلك الكلمات قليلة التكرار. وتناقش الدراسة مضامين هذه النتائج على فهمنا لألية التعرف على الكلمات في اللغة العربية.

الكلمات الافتتاحيّة: تكرار / الكلمة / سرعة / اللغة العربية / المعاصرة.