About the Neural Basis of Arabic Diglossia: Behavioral and Event- Related Potential Analysis of Word Processing in Spoken and Literary Arabic

Asaid Khateb and Raphiq Ibrahim

1 Introduction

The question of whether or not the language acquired first in life (L1) and a second language (L2) learnt later in life are represented (i.e., managed, processed etc…) in the same/or different brain regions in bilinguals has stimulated a huge amount of research during the last two decades (Mouthon et al., 2013; Perani & Abutalebi, 2005). Clinical aphasic manifestations in bilinguals following brain damage had initially suggested that L1 and L2 might be managed by different brain areas in the bilingual brain (Albert & Obler, 1978; Fabbro, 2001a, b; Ibrahim, 2009; Junque et al., 1995; Paradis, 1977, 1983, 1998). In addition, experimental observations based on intraoperative electro-cortical stimulations had suggested that while some left brain regions could be involved either in L1 or L2 processing, other areas were involved equally in both languages (Ojemann, 1983; Ojemann & Whitaker, 1978). In line with such views, some early functional imaging studies in bilinguals had concluded that some areas might contribute differently to the two languages processing while others are shared between languages (see, De Bleser et al., 2003; Klein et al., 1994; Marian & Kaushanskaya, 2007). Later on, other studies conducted with bilinguals who were more proficient in L2 failed to demonstrate differences in brain language networks, in particular when classical language areas were considered (e.g., Chee et al., 2001; Hernandez et al., 2001). Hence, recent views about the brain's represevntation of L1 and L2 tend to assume that the bilinguals' different languages rely on a common brain network for their processing but that differences in activation observed during functional studies are explainable by other

A. Khateb (✉) · R. Ibrahim
The Unit for the Study of Arabic Language, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Faculty of Education, University of Haifa, Haifa, Israel
e-mail: akhateb@edu.haifa.ac.il; raphiq@psy.haifa.ac.il

193

factors (Mouthon et al., 2013). Among these, the age of acquisition of L2, the level of proficiency in L2 and the exposure to/and the patterns of use of L2 appear to have cumulative effects together influencing the bilingual's general cognitive-linguistic functioning (see Abutalebi, 2008; Abutalebi & Green, 2007; Bloch et al., 2009; Hernandez, 2009; Hernandez & Li, 2007; Perani & Abutalebi, 2005; Perani et al., 2003).

In the Arabic language, the diglossic situation is seen by some authors as a particular form of bilingualism (see, Eviatar & Ibrahim, 2000; Ibrahim, 2009; Ibrahim & Aharon-Peretz, 2005) or as a context inducing processing patterns akin to those seen in bilinguals (Saiegh-Haddad, 2003; Saiegh-Haddad et al., 2011). Diglossia as first defined by Ferguson (1959) refers to a stable socio-linguistic state that includes different spoken dialects and a remarkably different, grammatically more complex standardized language version. In Arabic, diglossia is defined by the existence of two main varieties of Arabic: (i) a low form which is the spoken version that is acquired naturally and used in everyday conversation and informal communication purposes (referred to hereafter as Spoken Arabic or SA) and (ii) a highly codified form, referred to as Literary Arabic (hereafter LA, referred to also Modern Standard Arabic or MSA, see Saiegh-Haddad & Joshi, 2014). LA is acquired mainly formally at school for reading and writing[1] and used in official contexts, such as media (written media and news broadcasts), speeches, religious sermons and formal discourses (Saiegh-Haddad, 2012; Saiegh-Haddad & Henkin-Roitfarb, 2014). LA differs from SA in almost all its linguistic aspects including the grammatical, syntactic, morphological, phonological and lexical aspects. Due to the distance between SA and LA and the fact that the written Arabic does not represent the spoken language (where various structures are different from LA structures), researchers suggest that diglossia impacts significantly reading and writing acquisition (Mahfoudhi et al., 2011; Saiegh-Haddad, 2003, 2004, 2018; Saiegh-Haddad & Schiff, 2016). Different studies pointed to the difficulty that the Arabic children encounter in the construction of phonological representation and processing for words and sub-lexical units in LA (Saiegh-Hadda & Haj, 2018; Saiegh-Haddad et al., 2020). In fact, at the beginning of their learning to read process, children are practically asked to acquire two systems simultaneously: a linguistic-auditory system (that normally exists in pre-school children in non diglossic situations) and an orthographic-visual system which happens regularly at the start of the school life (Ibrahim et al., 2002; Saiegh-Haddad, 2004).

On the basis of a series of psycholinguistic studies, Saiegh-Haddad and colleagues have shown that the linguistic distance between SA and LA impacts a variety of linguistic processing skills in LA (For a review see, Saiegh-Haddad, 2018;

---

[1] Because SA serves strictly for oral communication, it did not exist in the written form, until the recent emergence of "electronically mediated communication". The ability to exchange messages that tend to be of an informal nature has resulted in the emergence of written messages in SA. In the beginning and due to technical limitations, electronic communication used Roman characters (e.g., Palfreyman & Khalil, 2003; Warschauer et al., 2002), a phenomenon, referred to as "Arabizi". In recent years, thanks to the advent of smart phones which enable writing messages using the Arabic keyboard, this phenomenon has almost completely disappeared.

Saiegh- Haddad in this collection). For instance, at the syntactic level, word order in LA sentences is usually VSO (verb-subject-object) while in SA the common word order is SVO. Also, despite a certain overlapping, the phonological systems of LA and SA are quite different, with some LA phonemes being absent in certain SA dialects. Finally, although SA and LA share many words in common (often with certain phonological nuances), SA and LA may also have different words for the same referents. In this regard, Saiegh-Haddad and Spolsky (2014) analyzed the lexicon of young 5 year children and reported that 40% of the words consisting of nonstandard words that have no conventional written form, another 40% consisting of SA-LA cognates and only 20% of the words had identical forms in SA and LA. Also, phonological distance between SA and LA had been suggested to be at the origin of the difficulties in reading acquisition among Arabic native children (Saiegh-Haddad, 2007). For instance, in one study (Saiegh-Haddad et al., 2011) that used a picture selection task for words beginning with the same phoneme, the authors reported that the children's recognition of LA phonemes was poorer than that of SA ones. This finding suggested difficulty in the phonological representations for LA words, to which children are generally exposed for the first time at the moment of their entry to school.[2] In an earlier study, the same author (Saiegh-Haddad, 2003) investigated reading processes in children (kindergarten and first grade) and compared their performance on phonemic awareness and word syllabic structure between LA and SA words. She suggested that diglossic variables influenced the children's performance in phoneme isolation and pseudoword decoding. In line with the assumption that diglossia might delay (or lead to difficulties in) reading acquisition among Arabic native children, different studies have also suggested that early exposure of native Arabic speaking children to LA might improve their reading abilities in the early grades (Abu-Rabia, 2000; Feitelson et al., 1993).

During the last two decades, several researchers have also sought to assess the extent to which SA and LA behave as real L1 and L2 in the cognitive system of literate Arabic speakers, as in more classical forms of bilingualism. To address this question, researchers compared the processing of SA and LA words in different language tasks using behavioral measures (reaction times, performance) or compared the performance of native Arabic speakers with the performance of bilinguals (Eviatar & Ibrahim, 2000; Ibrahim, 2009; Ibrahim & Aharon-Peretz, 2005; Ibrahim et al., 2007). For instance, Eviatar and Ibrahim (2000) assessed the metalinguistic abilities in Arabic speaking children (kindergarten and first grade) who were exposed to both SA and LA and compared their performance to Russian-Hebrew bilinguals and to Hebrew-speaking monolingual children. The results indicated that the Russian- Hebrew bilinguals displayed the classical pattern of early bilingualism (as attested by higher meta-linguistic abilities, but with lower vocabulary compared to monolinguals), and Arabic-speaking children's behavior mimicked that of the

---

[2]Although the formal exposure to the standard language occurs when children go to school, they are however exposed to various extent to LA through media and TV programs for children and through oral storytelling by parents at home and by educators in kindergartens (see a discussion on this issues in Saiegh-Haddad & Spolsky, 2014).

Russian-Hebrew bilinguals but differed from the Hebrew monolinguals (see also Ibrahim et al., 2007). Based on such results, the authors suggested that since Arabic native speakers behaved as bilinguals, they could be considered as bilinguals (Eviatar & Ibrahim, 2000). In another study, Ibrahim and Aharon-Peretz (2005) examined intra- and inter-language (semantic) priming effects in auditory lexical decision in 11th and 12th grade native Arabic speaking students, who were L2 speakers of Hebrew. Presentation of stimuli in Hebrew, in addition to SA and LA, enabled comparisons between SA and LA, the processing of which was in the focus of the studies, as well as comparisons of both language varieties to Hebrew, their formal second a language. In this first study (Ibrahim & Aharon-Peretz, 2005), the authors reported that priming effects were larger when prime words were in SA and target words were either in LA or in Hebrew than when presentation was the other way around (primes in LA or Hebrew and targets in SA). Further, the magnitude of the priming effects for LA and for Hebrew were indistinguishable, suggesting that both languages behaved as second languages in the diglossic situation. In another study (Ibrahim, 2009), primes in SA yielded greater and longer lasting priming effects on decisions regarding targets in SA than did primes in either LA or Hebrew. Here again, effects of primes in LA did not differ from those in Hebrew. The priming effects observed by Ibrahim and colleagues resembled to previous observations in bilinguals (Gollan et al., 1997; Keatley et al., 1994), where forward priming (from the dominant L1 to the less dominant L2) are larger than priming in the opposite direction (from L2 to L1: backward priming). This asymmetry has been taken to indicate that words in L1 more readily initiate conceptual processing than words in L2 (Kroll & Tokowicz, 2001). Based on such results, it was suggested that the two varieties of Arabic are represented in the cognitive system in two separately organised lexicons and that literate speakers of Arabic behave as bilinguals, with SA as their first language (L1) and LA as their second language. This conclusion seemed to hold at least as far as auditory stimuli were concerned. Actually, in a previous study using visual presentation of words from LA, SA and Hebrew, Bentin and Ibrahim (1996) using lexical decision and word naming (reading aloud) tasks showed that the processing of written SA words was slower than that of LA words, with SA ones being processed like LA low frequency words. Altogether, these behavioral data suggested that processing of SA and LA words depends on the modality of presentation of the stimuli with SA showing a pattern of response dominance in the auditory modality and LA words showing a pattern of response dominance in the visual written modality. In order to test this assumption and to shed light into the neural basis of diglossia, a series of studies have been conducted using electrophysiological (event-related potential: ERP), behavioral and functional magnetic resonance imaging (fMRI) measures.

Actually, until very recently little research has been conducted on written Arabic and on the Arabic language and diglossic situation more generally using functional brain imaging (Bourisly et al., 2013) or electrophysiological methods. Few studies have investigated word processing in Arabic in general (Al-Hamouri et al., 2005; Boudelaa et al., 2010; Mountaj et al., 2015; Pratt et al., 2013a, b; Simon et al., 2006; Taha et al., 2013; Taha & Khateb, 2013) with ERPs and only one addressed the

question of diglossia (Khamis Dakwar & Froud, 2007) in particular. To give some examples, Boudelaa and colleagues (Boudelaa et al., 2010) for instance conducted a study which focused mainly on written LA word to assess morphemic processing using ERP analysis. In another study, Al-Hamouri and colleagues (Al-Hamouri et al., 2005) examined the spatiotemporal pattern of brain activity during reading in Arabic and Spanish by means of magneto-encephalographic recordings (MEG). They found no difference between the two languages between 200 and 500 ms after stimulus onset, but found that Arabic enhanced right hemisphere activity beyond 500 ms. Simon and colleagues (Simon et al., 2006) used ERP measures to analyze orthographic transparency effects in Arabic and French subjects. They observed that the N320 component, which is related to phonological transcription, was elicited only in French subjects while reading their L1. In another study, Taha and colleagues (Taha et al., 2013) assessed the effects of word letters' connectedness and reported that fully connected words were processed more efficiently than non-connected ones as attested by RT and ERP measures. As for studies addressing the diglossia question, the only study found here assessed language code-switching between SA and LA (and semantic anomaly processing) using auditory sentence presentations in only 5 subjects (Khamis Dakwar & Froud, 2007). Although the results of this last study must be considered with caution due to the very limited experimental sample, the authors concluded that the diglossic switches in their experiment between the two varieties of Arabic elicited the pattern of ERP responses predicted from previous studies investigating code-switching between two different languages. The authors claimed that these results support the view that the two language varieties involve distinct and separate lexical stores.

In this chapter, we describe the beginnings and some conclusions of a series of studies that sought to shed light on the neural underpinnings of the diglossic situation in the Arabic language. In fact, diglossia is a complex sociolinguistic situation that had only poorly been studied using brain research methods. While being aware of the need to address the question of the brain basis of diglossia from various angles (word recognition, comprehension, production etc…), we first choose to rely on the previous findings within this research domain. Namely, we relied on tasks using single word recognition in the auditory and the visual modalities during lexical decision paradigms. Based on the hypothesis that the two varieties of Arabic might be processed in the brain of Arabic native literate speakers as two different languages, the objective was to characterize the neural responses differentiating SA and LA word processing by means of event-related potential (ERP) analysis in adult subjects. Contrary to previous investigations which used only behavioral analysis, the use of electrophysiological measures allow to investigate in real time the brain responses involved in word recognition in the two forms of Arabic and to compare them to Hebrew, the participants formal second language. The combination of ERP and behavioral analysis allows correlating brain activity with response time patterns and define time periods during the stream of information processing where the two varieties could converge and where they could diverge. Indexes of convergence and divergence were hypothesized to be reflected in the ERPs. Hence, on the basis of the assumption that SA might be processed as an L1 in the auditory modality, we

predicted that auditorily presented SA words will be processed faster than LA ones. We expected to find ERP differences between SA and LA that reflect the RT differences. Furthermore, in line with the assumption and previous literature that LA words might be processed as L1 in the visual written modality, it was predicted that written LA words will be processed faster than SA ones. Similarly, we expected to find ERP differences between LA and SA that reflect such RT differences. In all cases, and in both the auditory and visual modality, the processing of Hebrew words, the participants' formal second language will be used as a control condition. Also, this study relied on the fact that there are lexical items that differ completely between LA and SA (Saiegh-Haddad & Spolsky, 2014), but designate the same referent such as for instance the word "*dallo*" in LA and "*satel*" in SA which both refer to the object "*bucket*".[3] Hence, in order to enhance the putative differential effects in the processing of SA and LA words, we selected the words mainly from this last category, together with other words which share a minimal phonological overlapping between SA and LA. In addition, because Hebrew is a Semitic sister language of Arabic, a particular caution was paid to avoid Hebrew words which overlap phonologically either with SA and LA ones.

2 Material and Methods

2.1 Participant

Two different groups of participants were included in these studies. A total of 43 students (28 women and 15 men, mean age 22.8 ± 1.75, range from 18 to 28 years) underwent the auditory lexical decision task. Of these, 31 participants underwent the ERP experiment. Also, a total of 45 students (23 women and 22 men, mean age 22.7 ± 2.3, range from 19 to 29 years) participated in the visual lexical decision experiment. Of these, 30 participants underwent the ERP experiment. All participants were recruited from the University of Haifa. All were self-declared right-handers, native speakers of Arabic, whose SA is the colloquial Palestinian Arabic and who have acquired LA through their schooling in Arabic speaking schools since the age of 6. All had acquired Hebrew since the 2nd grade and were moderately to highly proficient in this language, to which they were highly exposed in their everyday life in the University. All participants had normal or corrected-to-normal vision, with no history of dyslexia, neurological or psychiatric diseases. They were all asked to provide an informed written consent before the participation to the experiment and were paid for their participation.

---

[3] There are also words shared by SA and LA, and there are others which are characterized by variable degrees of relatedness between the two forms that ranges from identical phonological representation in both varieties, to a phonological distance that alters both the phonemic and the syllabic structure of the words.

2.2 Stimuli and Procedure

The same stimulus set was used for the two experiments. This was composed of 180 words and 180 legal pseudowords in Arabic and Hebrew of which one third in SA (i.e. 60 words and 60 pseudowords), one third in LA and one third in Hebrew. Of note is the fact that SA and LA words were exclusive such that a word in one variety was never a word in the other variety (see examples in Appendix 1). All words were rated as highly familiar nouns in each language and pseudowords were created in each language condition by changing one or two letters in the word (consequently one or two sounds auditorily). It is worth noting here that the use of pseudowords was only intended to create a lexical decision task to assess the process of word recognition auditorily and visually. The possible effects of lexicality (i.e., difference between real words and pseudowords) were not in the focus of this study but all word and non-word conditions were analyzed at the behavioral level to test the validity of the used material in each language variety. For ERP analysis, only responses elicited by the real words were analyzed to test this work's predictions. In the selected word lists, the real word lists were equated on the average frequency/ familiarity between languages. Thus, for the initial selection of the words, a first list that contained 321 randomized words was constituted (107 in each language condition), of which each word was rated for its frequency (familiarity in the respective language or language variety) by 46 adult volunteers using a 5 points scale (1 for non-frequent/non familiar and 5 for highly frequent/familiar). The average frequency for each item in each language variety list was computed and this allowed the selection of the 60 most frequent items in each language condition. These values were statistically compared using a one way ANOVA with three language conditions. This analysis showed that the stimuli did not differ in terms of word frequency ($p = .88$) with an average frequency of 4.3 (± 0.32), 4.33 (± 0.34) and 4.30 (± 0.33) respectively in SA, LA and Heb.[4] Once selected, these items allowed the creation of the equivalent language lists of pseudowords. All the stimuli were then digitalized for the auditory lexical decision task using a male voice speaking the SA, LA and Hebrew. The digitalized words underwent computer processing, designed to equalize their volume, and their length as much as possible (with an average duration of ~1000 ms). In the average auditorily, SA words were of 0.89 s ($\pm$ 0.14), LA words of 0.91 s (± 0.18) and Hebrew words of 0.89 s (± 0.19). Written words in all conditions varied between 3 and 6 letters in length. In the average, SA words were of 4.27 letters (± 0.98), LA of 4.28 (± 0.78) and Hebrew words of 4.13 (± 1.04). In each experiment, the stimuli belonging to the different language and word conditions were then pseudo- randomized in a list that contained 360 stimuli. This list was then

---

[4] *Of note is the fact that we initially relied on Arabic speakers to rate also the Hebrew words, but because this is formally their second language, the average frequency for each of the words appeared a little low. Since it is a well-known fact the in the average second language words are of subjectively lower frequency than first language words, we passed the questionnaire to 10 Hebrew native students who rated them again and the frequency values reported here come from the Hebrew speakers and as seen indeed they compare to their equivalents in Arabic.*

divided into three equivalent sub-lists of 120 items each to form three experimental blocs, the order of which was balanced across subjects in each experiment. In addition, the order of the stimuli in each list/experimental block was randomized at each run for each participant.

In the auditory lexical decision experiment, the stimuli were presented to the subjects through earphones. Participants were instructed after the presentation of each stimulus (in the mixed list of SA, LA, and Hebrew spoken words and pseudowords) to respond using two button presses as quickly and accurately as possible whether each stimulus was a word or not (in Arabic or Hebrew). Each stimulation trial started by a fixation cross that appeared for 650 ms on the center of the screen in black over a white background, then the auditory stimulus was presented within an allowable response window of 2 s (with the fixation continuing to appear), and then a blank screen for about 1050 ms (varying between 950 and 1200 ms) as an inter-stimulus interval announcing the eminence of the next trial.

In the visual lexical decision task, participants in each trial saw a string of letters and were required to respond using two button presses as quickly and accurately as possible whether or not these letters constitute a word they know. Each trial started with a 500 ms fixation cross, followed by the stimulus during 150 ms. A blank screen appeared during 1850 to allow for the subject's response.

The participants were seated comfortably in front of a computer screen, approximately at 90 cm distance and were asked to perform a speeded lexical decision task (LDT) by pressing as quickly and accurately as possible using two keyboard keys. All participants responded with their dominant right hand. Half of the subjects in each experiment responded with their dominant major and index fingers for word vs non-word (pseudoword) and the other half responded using the inverse, major for non-words and index for words. All subjects in each experiment underwent the three experimental blocs (separated by a short break of ~3–5 min) the order of which was balanced across subjects. In addition, all underwent a short training session to familiarize with their task.

2.3   Electroencephalographic (EEG) Recordings and Analysis

The experiments were carried out in a quiet and sound attenuated room. EEG recordings were collected continuously using a 64 channel BioSemi Active Two system (www.biosemi.com) and the Active view recording software. 64 pin-type electrodes were mounted on a customized Biosemi head-cap (distributed all over the scalp according to the 10–20 international system) using an electrode gel. Additionally, two flat electrodes were placed on the sides of the eyes in order to monitor horizontal eye movements and a third electrode was placed below the left eye to monitor vertical movements and eye blinks. The EEG was collected reference free (i.e., the so called "Biosemi active electrodes") with a 0.25 high pass filter, amplified and digitized with a 24-bit AD converter, at 2048 HZ sampling rate.

The ERP epochs for trials with correct responses were averaged and analyzed off line for the two experiments using the Cartool software© (v.3.51; http://brainmapping.unige.ch). Briefly, the ERP epochs were filtered between 1 and 30 Hz and averaged separately for each word condition from −100 ms before the presentation of the auditory/visual stimulus to 700 ms post-stimulus. Before accepting the ERP epoch for each trial for which a correct answer was provided, the EEG data passed also a visual inspection to exclude trials with eye-movement artifacts and to exclude sweeps exceeding ±100 µV. After ERP averaging, the individual ERPs of each condition were down-sampled from 2048 Hz to 512 Hz, baseline corrected using the 100 ms pre-stimulus period, referenced to the average-reference (Lehmann & Skrandies, 1980) and averaged separately in each language to compute the grand-mean ERPs for SA, LA and Hebrew.

2.4   ERP Wave Shape Analysis

The individual ERPs were then subjected to a waveform analysis based either on exploratory statistical analysis or on the visual inspection of the superposition of the grand-mean waveforms. These analyses allowed determining the earliest time windows where reliable differences seemed to occur after stimulus onset. In order to assess statistically the data driven hypotheses based on the waveforms inspection, the signal for the period of interest and the electrodes of interest (see the Results section for details), both from subsets of left and right hemisphere recording sites, was extracted. This signal was then subjected to repeated measures analysis of variance (ANOVA) using language condition (SA, LA and Heb), hemisphere (left and right) and electrodes as within subjects' factors.

2.5   Behavioral Analysis

The median of the individual reaction times (RTs) for correct trials (>75% accuracy in all conditions in both experiments) was computed for each language condition separately for words and pseudowords conditions. This detailed analysis was done only for the purpose of verifying the validity of our tasks and the stimuli used here for the two experiments. In both experiments, we expected real words to be recognized faster than pseudowords as generally found in lexical decision tasks. For the RT measures, the response times below 250 ms were discarded from the individual computations. Individual values of the different RT measures were compared statistically between word conditions and language conditions using 2×3 ANOVA with word type (i.e., lexicality: word vs pseudoword) and language as within subjects' factors.

3   Results

In this section, behavioral and electrophysiological results will be presented first for the auditory lexical decision task and then for the visual lexical decision task.

3.1   Auditory Lexical Decision Task

Response Time Table 1 shows the mean RTs (±SD) over subjects for the different conditions in the three language conditions. The 2×3 repeated measures analysis of variance (ANOVA) performed on the subjects' individual median RTs showed a significant main effect of word type ($F(1, 42) = 282.39$, $p < .00001$) due to the fact that RTs were in the average faster for words (M = 1182 ms) than for pseudowords (M = 1392 ms). A significant main effect of language condition was also observed ($F(2, 84) = 69.85$, $p < .00001$) due to the fact that RTs increased gradually from SA through LA and Hebrew (in the average, M = 1238 ms, = 1298 ms and = 1326 ms respectively for SA, LA and Heb). A significant interaction between the factors was also found ($F(2, 84) = 7.37$, $p < .005$), due to the fact that the difference between word and pseudowords condition was not homogeneous across language conditions. Interestingly, the lexicality effect was larger here for SA (M = 229 ms) than for LA (M = 173 ms). More particularly for our purpose, the one-way ANOVA performed on RTs for the words only showed a highly significant language effect ($F(2, 84) = 45.68$, $p < .00001$). Post-hoc Fisher's LSD tests showed that RTs were shorter in SA than in LA ($p < .00001$) and in Hebrew ($p < .00001$), with the later two exhibiting no significant difference ($p = .89$, see Fig. 1).

Electrophysiological Results Due to a high number of eye movements and other artefacts in the auditory EEG data, 23 (16 women, 7 men) of the 31 participants were included in the following ERP analysis.[5] In order to identify the earliest ERP differences between language conditions, we first conducted an exploratory point-wise t-test analysis (see details of methods in Khateb et al., 2010; Taha & Khateb,

Table 1 Auditory lexical decision: Mean reaction times in ms. (±standard deviation, n = 43) for word and pseudoword conditions in the three languages: SA (Spoken Arabic), LA (Literary Arabic) and Hebrew (Hebrew)

	SA	LA	Heb
Words	1123	1211	1212
	(129)	(138)	(118)
Pseudowords	1352	1384	1438
	(125)	(133)	(143)

[5] *It is worth noting the re-analysis of the 23 subjects' behavioral data yielded very similar statistics on the RTs (not included here).*

Fig. 1 Graph depicting the mean RTs (in ms.) in the auditory lexical decision task for the words conditions in the three language varieties: SA (Spoken Arabic), LA (Literary Arabic) and Hebrew (Hebrew). Vertical bars denote 0.95 confidence intervals
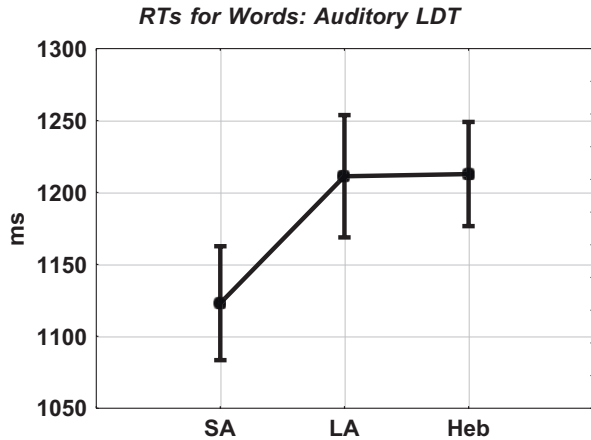
2013) on all electrodes and all time frames. This aimed at determining time points of reliable response differences between SA and LA, between SA and Hebrew and between LA and Hebrew after stimulus onset. This analysis (not illustrated here) showed that the earliest differences occurred at around 300 ms after word onset between SA vs LA and between SA vs Hebrew but not between LA vs Hebrew. Figure 2a illustrates the superposition of grand mean ERP traces from the different language conditions on a subset of left and right (anterior and posterior) electrodes which maximally depicted such differences. The traces on FC1 (upper left row) shows the P1-N1-P2 components' sequence as can be seen on the frontal electrodes. P1 occurred at around 130 ms, the N1 occurred at around 200 ms and the P2 occurred at around 300 ms. The posterior aspect of the P2 component showed the first reliable differences between SA and the other languages (see PO7 and PO8). The posterior aspect of the frontal P2 component was characterized by a negative response on the parieto-occipital electrodes (see blue shadow on Fig. 3a). In order to statistically assess these differences, we computed the mean signal in this time period between 280 and 330 ms from 3 left (P5 P7 PO7) and 3 right (P6 P8 PO8) posterior electrodes (see PO7 and PO8 in Fig. 3a, see inset in lower right panel). The $3 \times 2 \times 3$ ANOVA performed on the P2 mean amplitude using language condition (SA, LA and Heb), hemisphere (left vs right) and electrode (3 sites) showed a significant language effect ($F(2, 44) = 7.7$, $p < .005$), and an electrode effect ($F(2, 44) = 14.0$, $p < .00005$), with no significant interaction between the analysis factors. The language effect was due to the fact that ERP amplitude to SA was on the average more negative (mean $= -1.04$ mV) than to LA ($= -0.52$ mV, $p < .005$) and Hebrew ($= -0.48$ mV, $p < .001$), with the later two not differing ($p = .83$). This finding is illustrated in Fig. 3b which shows a more negative response in SA than in the other language conditions on all tested electrodes.

A similar $3 \times 2 \times 3$ ANOVA performed on the P2 mean amplitude on frontal electrodes (3 left, FC3, FC1, C1 and 3 right: FC4, FC2, C2, see examples of FC1 and FC2 in Fig. 2a) showed also a significant language effect ($F(2, 44) = 5.7$, $p < .01$),
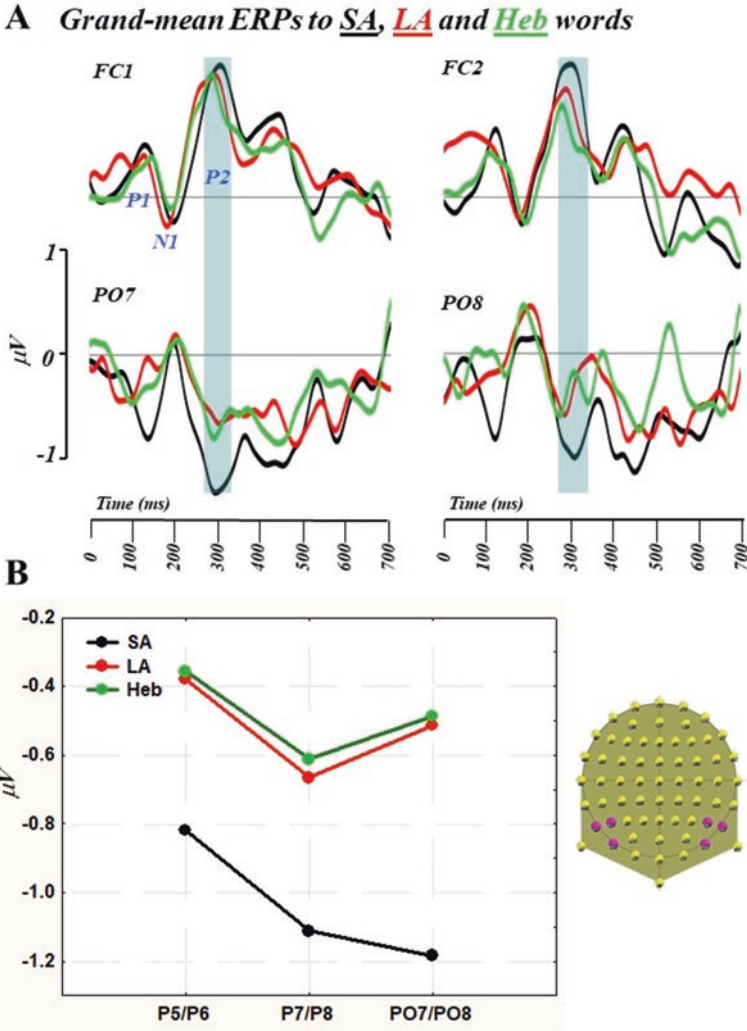
Fig. 2 (**a**) Superimposition of the grand mean ERP traces (from 0 to 700 ms post- stimulus) induced by SA words (black traces), LA words (red traces) and Hebrew words (green traces). The selected electrodes represent left (FC1) and right frontal (FC2) electrodes, left (PO7) and right (PO8) posterior electrodes where differences appeared maximally at the level of the P2 component. (**b**) Graph illustrating the mean signal for the P2 over left and right posterior electrodes as a function of language condition with SA (black) inducing responses being significantly different from LA (red) and Hebrew (Green, see text for statistics). Inset in the lower right shows the localization of the left and right posterior electrodes included in this analysis

and an electrode effect (F(2, 44) = 22.6, $p < .00001$), with no significant interaction between the analysis factors. Again here, the language effect was due to the fact that the frontal P2 response in SA was on the average more positive (mean = 1.03 mV)
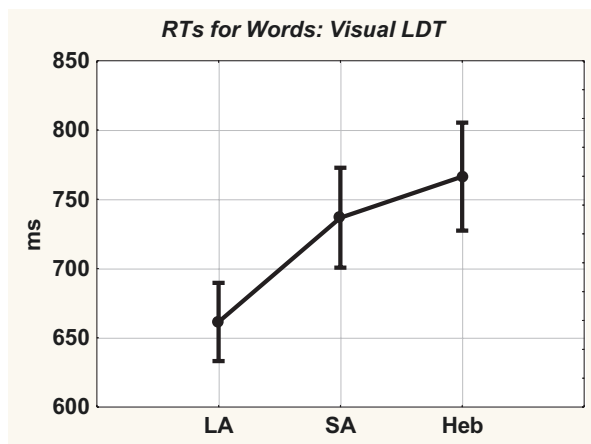
Fig. 3 Graph depicting the mean RTs (in ms.) in the visual lexical decision task for the for the words conditions in the three language varieties: LA (Literary Arabic), SA (Spoken Arabic) and Hebrew (Hebrew). Vertical bars denote 0.95 confidence intervals

than in LA (=0.82 mV, $p = .09$) and in Hebrew (=0.62 mV, $p < .005$), with the latter two not differing ($p = .11$).

To summarize, the results presented here showed that in terms of RTs, SA differed from both LA and Hebrew and behaved as the dominant language variety Electrophysiologically, the analysis of the participants' responses showed that the P2 component amplitude was larger in SA than in LA and Hebrew. Taken together, the results of this study in the auditory modality confirms that SA holds the status of the dominant language variety since SA and LA behaved differently during early the processing steps which seemingly strongly influenced word recognition speed and thus determined subjects' reaction times. During these early steps of information processing, LA which is acquired later in life and Hebrew which is the participants' formal second language behaved quite similarly.

3.2   *Visual Lexical Decision Task*

Response Times Table 2 summarizes the mean (±SD) of the subjects' median response time (RTs) for the different word types in the three language conditions. A 2×3 repeated measures analysis of variance (ANOVA) was conducted on the subjects' individual RTs with word type (2: word vs pseudoword, or lexicality) and language (3: LA, SA and Heb) as within subject factors. The analysis showed first a highly significant main effect of word type ($F(1, 44) = 65.48$, $p < .00001$) due to the fact that RTs were faster to words (M = 722 ms) than to pseudowords (M = 839 ms). This analysis showed also a highly significant main effect of language condition ($F(2, 88) = 25.22$, $p < .00001$) due to the fact that, in the average, RT augmented

Table 2   Visual lexical decision: Mean reaction times in ms. (±standard deviation, n = 45) for word and pseudoword conditions in the three language conditions: LA (Literary Arabic), SA (Spoken Arabic) and Hebrew (Hebrew)

	LA	SA	Heb
Words	661	737	766
	(94)	(120)	(130)
Pseudowords	825	817	874
	(133)	(134)	(192)

gradually from LA (M = 743 ms) to SA (M = 777 ms) to Hebrew (M = 820 ms). A significant interaction was also found between the two factors (F(2, 88) = 12.21, $p < .0001$) due to the fact that difference between words and pseudowords was again not homogenous. Interestingly, and contrary to the results in the auditory lexical decision, the lexicality effect was larger here for LA (M = 164 ms) than for SA (M = 80 ms). Of more interest for our purpose, the one-way ANOVA performed on median RTs for the words conditions only showed a highly significant language effect (F(2, 88) = 45.49, $p < .00001$). This was due to the fact that RTs were shorter in LA than in SA ($p < .00001$) and in Hebrew ($p < .00001$), and to the fact that RTs were also shorter in SA than in Hebrew ($p < .01$, see Table 2 and Fig. 3).

Electrophysiological Results   Due to the presence of a high amount of artefacts in the EEG data of one subject, the following analysis was performed on 29 out of 30 recorded subjects. In order to determine time periods of possible reliable response differences between the three language conditions, we first conducted a visual inspection of the grand-mean ERP traces of the different conditions. As shown in Fig. 4a which illustrates a superposition of the traces from a subset of frontal and posterior recording sites from LA (Black), SA (Red) and Hebrew (Green), the earliest response differences occurred around the N1-P2 component complex (see labelling of P1-N1-P2-N2 on electrode PO8). In order to statistically assess these differences, we computed the mean signal in two regions of interests that included four left (P5, P7, PO3 and PO7) and four right (P6, P8, PO4 and PO8) posterior electrodes (see inset in lower right panel, Fig. 4b). From the individual averaged left and right traces of each participant in each condition, we determined the time points of the successive P1-N1-P2-N2 components.[6] The 3 × 2 ANOVA performed on the amplitude of each of these early components using language condition (LA, SA and Heb), hemisphere (left vs right) showed no significant language effects for the P1 and N1 components. In contrast, a significant language effect (F(2, 56) = 4.74, $p < .02$) and a hemisphere effect (F(1, 28) = 74.76, $p < .00001$) was found for the P2 component. As shown in Fig. 4b, the language effect was due to the fact that ERP amplitude to LA was on the average more positive (mean = 0.51 mV) than to SA (= 0.05 mV) and Hebrew (= −0.97 mV). A similar finding was also found for the N2

---

[6] These points determined from around 100 ms onwards refer to the first highest positive value around the P100 component (~100 ms) and its time latency, then successively the most negative for the N170 (~170 ms), then for the P2 and the N2.
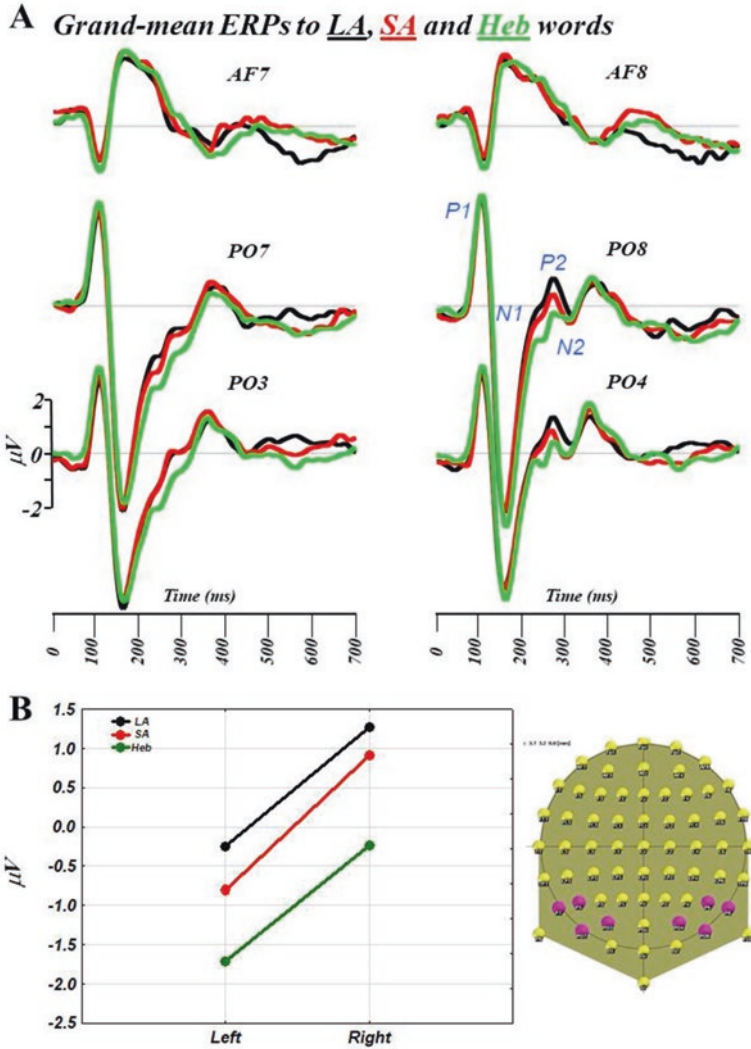
Fig. 4 (**a**) Superimposition of the grand mean ERP traces (from 0 to 700 ms post- stimulus) induced by LA words (black traces), SA words (red traces) and Hebrew words (green traces). The selected electrodes represent left (AF7) and right frontal (AF8) electrodes, left (PO7 and PO3) and right (PO8 and PO4) posterior electrodes where differences appeared maximally at the level of the P2/N2 components. (**b**) Graph illustrating the mean signal over the left and right regions of interest including 4 posterior electrodes showing a significant effect of language (and hemisphere) on the P2 component (black for LA, red for SA and Green for Hebrew, see text for statistics). Inset in the lower right shows the localization of the left and right posterior electrodes included in this region on interest analysis

component which demonstrated again significant language effect (F(2, 56) = 3.78, p < .03) and a hemisphere effect (F(1, 28) = 11.86, p < .002) due to a more positive signal in LA than in the other conditions. No effect of language was found for the time latency of either component. In the average in all conditions, the P1 occurred at ~105 ms, the N170 at ~173 ms, the P2 at ~250 ms and the N2 at ~300 ms.

To summarize, in terms of RTs, LA differed from both SA and Hebrew and behaved as the dominant language variety. The results of the P2-N2 complex showed a relation between the ease with which the words are recognized (as attested by RTs) and the amplitude of the response. Taken together, the results of this study in the visual modality indicated that LA holds the status of the dominant language variety both behaviorally and electrophysiologically during word recognition processes.

4   Discussion

Although some previous efforts have been devoted to investigate psycholinguistically the relationship between the two Arabic varieties, no previous research addressed the question of the brain basis of diglossia. In the diglossic Arabic-Hebrew bilinguals, previous investigations using behavioral measures only have suggested that the cognitive system treats LA differently than SA, which is the language variety acquired first by native Arabic speakers, and similarly to Hebrew, which is a formal L2 acquired later in life (Ibrahim, 2009). In particular, studies using auditory lexical decision assessing semantic priming suggested that SA behaved as the dominant language variety relative to LA and Hebrew as attested by the magnitude of the priming effects. Inversely, studies using visual presentation of LA, SA and Hebrew words (Bentin & Ibrahim, 1996) showed that LA behaved as the dominant language variety with SA ones behaving as LA low frequency words. Based on such previous evidence, we hypothesized that in the diglossic situation of Arabic, the status of SA and LA will be modality-dependent with SA functioning like an L1 and LA as L2 in the auditory modality and LA functioning as an L1 and SA as an L2 in the visual written modality. Because diglossia is a complex situation that must be tackled from the different angles of language production and comprehension, but still has not been investigated by means of functional brain measures, we first choose to rely on the these early findings related single word processing in the auditory and the visual modalities. Based on the hypothesis that the processing of the two varieties of Arabic in the brain of Arabic native literate speakers might mimic that of two different languages, the objective was to assess the neural responses differentiating SA and LA word processing by means of event-related potential (ERP) analysis in adult subjects. The reported studies aimed at providing for the first time both behavioral and electrophysiological evidence to test this prediction using the same type of task and the same linguistic material. In this lexical decision task, the participants' analysis of RT first showed a lexicality effect attested by the words' superiority effect in both varieties of Arabic and in Hebrew. This

expected effect is in line with previous data in the literature (Bentin et al., 1985; Coltheart et al., 2001; Forster & Chambers, 1973; Khateb et al., 2002) and confirmed the validity of the task and the used stimuli. The first prediction was that auditorily presented SA words would be processed faster than LA ones and that ERPs will show the correlates of this difference. The differences found here in RTs confirmed the first prediction. At the behavioral level, we found as expected that the response times were the shortest in SA and differed from both LA and Hebrew. These differences could be explained neither in terms of words' general familiarity in the Arabic language nor of word length since the words in all language conditions were equated in these respects. The fact that LA and Hebrew behaved similarly here in terms of RTs somewhat extends previous findings suggesting that, in certain instances, LA presents more similarities with Hebrew than with SA in the auditory modality (Ibrahim, 2009).

At the electrophysiological level, these first results were consistent with the behavioral findings presented above. The results showed that the earliest differences between language conditions appeared between words at around 300 ms after stimulus presentation, during a time period referred to here as to P2. The statistical comparisons of the ERP signals from SA, LA and Hebrew revealed reliable significant response differences between SA and both LA and Hebrew but not between LA and Hebrew. In a previous lexical decision task involving first and second language words (Sinai & Pratt, 2002), ERP analysis in Hebrew-English bilingual speakers reported significantly longer latencies for N1 and P2 components to word pairs including L2, and suggested that different processing of L1 and L2 words occurred as early as during the stages associated with activation of the auditory cortex, but also showed difference during N400 between the two languages.

Although a more detailed analysis of the time course of the ERPs is still needed in order to better assess processing steps where SA and LA diverge and converge, the direct results which arise from the data in connection with the goals of this study are very conclusive. The fact that differences between the two forms of Arabic appeared both in terms of response speed and brain response amplitude is compatible with the history of acquisition of the phonological representations and words (Saiegh-Haddad & Haj, 2018; Saiegh-Haddad et al., 2020) in these different language varieties, and with the frequency of exposure to them in the auditory modality in everyday life. These findings confirm the dominance of SA in the auditory modality and support results from previous studies that suggested that SA words behave as L1 ones (Ibrahim, 2009; Ibrahim & Aharon-Peretz, 2005; Ibrahim & Eviatar, 2009).

As for the visual lexical decision task, the results presented here also confirmed the study prediction. In fact, RT analysis showed that using the same stimuli in the visual modality led to faster recognition of LA words. ERP analysis in parallel showed a modulation of the P2-N2 components which reflected the ease with which words were identified in the different language conditions. Previous observations in ERP literature show a modulation by word frequency of ERP response during the 150–300 ms time period (Hauk et al., 2006; Hauk & Pulvermuller, 2004; Proverbio et al., 2008). Differences during this time period were also reported between L2 vs

L1 in ERP analysis (Khateb et al., 2016). Hence, it appears plausible to suggest that these differences are mainly due to the visual familiarity/frequency of exposure to words in LA. This interpretation is also compatible with the history of acquisition and the patterns of use of LA words which are more frequently used (and the participants are more often exposed to) in the visual written modality. In fact, it is generally assumed that written transliteration of SA words has no accepted upon standard form in Arabic. Hence, one can predict that recognition of frequent LA words proceed from print to semantics using the lexical-semantic route based on word patterns while recognition of written SA words will be realized through the slower non-lexical phonological route, through a process of grapheme to phoneme conversion (Coltheart, 2005; Taouk & Coltheart, 2004). In functional brain imaging literature, one of the explanations raised to account for the differences observed between the processing of L2 vs L1 written words was the difference in proficiency in L2 compared to L1, the age of acquisition of L2, or the difference in subjective frequency of L2 words (Abutalebi, 2008; Abutalebi & Green, 2007; Perani & Abutalebi, 2005). This interpretation is certainly true here for words in Hebrew which showed the larger difference with LA both in terms of RTs and ERP response amplitude. However, for SA written words, and due to the currently widespread use of SA in non-formal written communication in social media, it is possible to predict that this change would lead to a minimization of the differences between LA and SA, a process which is certainly occurring in our participants. In line with this prediction, the results presented here suggest that despite the fact that the study participants were supposed to be quite highly proficient in Hebrew, their formal second language and to which they are highly exposed in their everyday life as students, they still recognized and processed more efficiently SA than Hebrew words. Future research should examine the long terms effects of the use of SA words in written electronic communication. Altogether and more importantly to our purpose, the results of the visual lexical decision task confirmed the prediction that in the visual modality, due to the history of acquisition and patterns of use of the written language, LA holds the status of the dominant language variety.

The pattern of RTs and electrophysiological response differences in the auditory and visual lexical decision tasks using SA, LA and Hebrew words confirmed the prediction and previous findings in the literature that the status of SA and LA in the cognitive system of native literate Arabic speakers is modality-dependent. In particular, in the visual written modality, LA words, the language variety acquired later in life and used for reading and writing and for formal communication, were processed faster and more efficiently than SA ones. In line with these conclusions, two different other studies using ERP analysis during single word and sentence processing during semantic tasks provided similar results. In the first study (Shehadi, 2013), behavioral and ERP measures were analyzed during the processing of semantically related and semantically unrelated written word pairs in SA and LA. While RTs were faster to LA than SA word pairs, ERP showed a more negative N400 component and a delayed response peak latency in SA compared to LA words, mimicking other L2 vs L1 effects reported in the literature. In the study using sentence semantic judgement task in SA and LA (Khazen, 2016), both semantically incongruent

word endings in SA and LA elicited again a more negative N400 response and a delay in its peak latency in SA compared to LA words. As for the N400, it was shown that this component amplitude was globally more negative in SA than in LA. This effect, which observed after both semantically congruent and incongruent sentence endings in SA, was interpreted as reflecting the ease/difficulty with which semantic integration processes take place in written SA sentences. This interpretation is strengthened by the observation of a delayed peak latency in SA, consistent with the claim of LA functioning as the dominant language variety in the visual modality. Taken together, all these results again confirm that the history of acquisition and the patterns of use are clearly the factors that determine how the brain processes the different types of information it receive from whatever modality. Along these lines, a recent fMRI study that analyzed the processing of LA, SA and Hebrew words using a semantic categorization task confirmed these observations (Nevat et al., 2014). In this first functional study on Arabic diglossia, it was found that RTs were faster to LA than to SA and Hebrew. The comparison of brain responses between SA and LA revealed differences that mimicked activation patterns found in comparisons of L2 vs L1 word conditions. In particular, an increase of activation was found in SA relative to LA (and not the inverse) in several language and left hemisphere areas.

Because diglossia is a complex linguistic phenomenon that must be addressed through the different modalities and contexts of language use including comprehension and production, a first study was also conducted by means of functional MRI to investigate picture naming in SA and LA in a mixed diglossic context and compared SA and Hebrew in a mixed bilingual context. In this completely different linguistic register (Abou-Ghazaleh et al., 2018), this study showed that naming in SA was slightly easier than LA, but was considerably easier than in Hebrew. fMRI analysis showed no difference when comparing brain activation between SA and LA. In contrast, Hebrew compared to SA revealed activation differences that could be interpreted both in terms of recruitment of language control modules and of second-to- first language effects. In a subsequent study, the aim was to assess the extent to which language control modules are engaged during language switching between SA and LA, in comparison to switching between SA and Hebrew (Abou-Ghazaleh et al., 2020). For this purpose, naming in SA in the bilingual SA-Hebrew mixed context, and in the diglossic SA-LA mixed context was compared to the simple naming context. The comparison of picture naming in SA in different contexts was predicted to reveal differences related to language control processes. The analysis of fMRI revealed significant effect of context that involved four main areas sensitive to the naming contexts (namely the left inferior frontal gyrus, the precentral gyrus, the supplementary motor area (SMA) and the left inferior parietal lobule). Analysis of these areas, together with two other areas (the left caudate nucleus and the anterior cingulate cortex) hypothesized to participate to language control (Abutalebi et al., 2008) revealed very striking findings. The comparison of SA naming in the diglossic context relative to the simple pure SA naming revealed a higher activation in all areas. These results appeared to support Abutalebi and Green's (2016) adaptive control hypothesis that predicts changes to the control demands of

language use as a function of the context requirements. Also, the findings suggested that in order for language control areas to be recruited, a high level of lexical competition should exist. This was actually the case when SA and LA were mixed, hence no difference in activation was found when comparing the activation of SA and LA in the same context (Abou-Ghazaleh et al., 2018). The findings regarding SA and LA in fMRI analysis during production in picture naming task contrast with those reported by Nevat et al. (2014) using visual word stimuli, where more activation was found for SA relative to LA. These previous findings suggested that, for the unique diglossic population of native Arabic speakers, the first acquired SA could in the written modality 'look' like an L2. This same result for SA in the written modality (confirmed here in several ERP experiments at the level of single word processing) contrasts with the conclusions proposing that SA words and LA words are processed as L1 and L2 ones (Ibrahim & Aharon-Peretz, 2005; Ibrahim, 2009) and confirmed in the auditory lexical decision experiment reported here.

5   Conclusion

These apparent contradictions in the results strengthen the primary assumption that guided this work, according to which the place that SA and LA hold (as "L1" and "L2" or inversely) might change as a function of the language modality in use. Based on all the results discussed (and some presented here), we again propose, as previously claimed (Nevat et al., 2014), that literate native speakers of Arabic who master the use of both SA and LA in everyday life function with two first languages: One in the auditory modality (SA) and one in the visual written modality (LA). During a language production tasks, the available data suggest that SA and LA might behave very similarly, although competitively (since sharing different linguistic features including particularly at the phonological/articulatory and lexical-semantic levels) as two first languages. Actually, despite their ability to manage the use of these "two first languages"/two language varieties, it appears that, when they are pushed to perform a lexical selection (at the single word level) in a "forced mixed diglossic mode", naming in each variety becomes a very competitive process that requires the engagement of language control mechanisms. Back to the question of whether Arabic diglossia is a form of bilingualism, the response is neither direct nor unequivocal. The observation here that SA and LA exchange places as L1 and L2 according to the modality used do not allow to conclude that they represent two separate linguistic systems. The cognitive status of each of the Arabic varieties seems to depend on several parameters that include (among other things) the nature of the task's demands, the linguistic register, the individuals proficiency in both varieties, the modality of presentation of the stimuli (auditory vs visual) and the type of processing (reception vs production, etc.). Given that we are only just starting out in this area of research, the conclusions raised here might not seem warranted. Hence, future research directions should not only investigate this issue in a wide range of modalities at the level of single word processing, but also at the

sentence level, during reading, listening, and discourse production and control for individual language proficiency in both varieties and each modality. A better understanding of the representation of, and interactions between, the two language varieties of Arabic is not crucial only for a greater understanding of Arabic diglossia itself but also of the human cognition and language experience in general.

Appendix 1: Examples of Words Used for LA, SA and Hebrew and Their Phonetic Translation

LA	SA	Hebrew	Referent
/ʔanf/ أنف	/xuʃum/ خُشُم	אף /ʔaf/	Nose
/dalw/ دلو	/saṭˤel/ سَطِل	דלי /dli:/	Bucket
/na:fiða/ نافذة	/ʃubba:k/ شبّاك	חלון /ħalo:n/	window
/miˤtˤˤaf/ معطف	/kabbu:t/ كبّوت	מעיל /miˤiːl/	coat
/sari:r/ سرير	/taxet/ تَخت	מיטה /mita/	bed
/θiya:b/ ثياب	/ʔawaˤiː/ أواعي	בגדים /bgadi:m/	clothes

References

Abou-Ghazaleh, A., Khateb, A., & Nevat, M. (2018). Lexical competition between spoken and literary Arabic: A new look into the neural basis of diglossia using fMRI. *Neuroscience, 393*, 83–96.

Abou-Ghazaleh, A., Khateb, A., & Nevat, M. (2020). Language control in diglossic and bilingual contexts: An event-related fMRI study using picture naming tasks. *Brain Topography, 33*, 60–74.

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing: An Interdisciplinary Journal, 13*, 147–157.

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica, 128*(3), 466–478.

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242–275.

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition, 19*(4), 689–698.

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., et al. (2008). Language control and lexical competition in bilinguals: An event- related FMRI study. *Cerebral Cortex, 18*(7), 1496–1505.

Albert, M., & Obler, L. (1978). *The bilingual brain*. Academic.

Al-Hamouri, F., Maestu, F., Del Rio, D., Fernandez, S., Campo, P., Capilla, A., et al. (2005). Brain dynamics of Arabic reading: A magnetoencephalographic study. *Neuroreport, 16*, 1861–1864.

Bentin, S., & Ibrahim, R. (1996). New evidence for phonological processing during visual word recognition: The case of Arabic. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22*, 309–323.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60*, 343–355.

Bloch, C., Kaiser, A., Kuenzli, E., Zappatore, D., Haller, S., Franceschini, R., et al. (2009). The age of second language acquisition determines the variability in activation elicited by narration in three languages in Broca's and Wernicke's area. *Neuropsychologia, 47*, 625–633.

Boudelaa, S., Pulvermuller, F., Hauk, O., Shtyrov, Y., & Marslen-Wilson, W. (2010). Arabic morphology in the neural language system. *Journal of Cognitive Neuroscience, 22*, 998–1010.

Bourisly, A. K., Haynes, C., Bourisly, N., & Mody, M. (2013). Neural correlates of diacritics in Arabic: An fMRI study. *Journal of Neurolinguistics, 26*, 195–206.

Chee, M. W., Hon, N., Lee, H. L., & Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. Blood oxygen level dependent. *NeuroImage, 13*, 1155–1163.

Coltheart, M. (2005). Modelling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading*. Blackwells Publishing.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256.

De Bleser, R., Dupont, P., Postler, J., Bormans, G., Speelman, D., Mortelmans, L., et al. (2003). The organization of the bilingual lexicon: A PET study. *Journal of Neurolinguistics, 16*, 439–456.

Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied PsychoLinguistics, 21*, 451–471.

Fabbro, F. (2001a). The bilingual brain: Bilingual aphasia. *Brain and Language, 79*, 201–210.

Fabbro, F. (2001b). The bilingual brain: Cerebral representation of languages. *Brain and Language, 79*, 211–222.

Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly, 28*, 71–79.

Ferguson, C. A. (1959). Diglossia. *Word, 14*, 47–56.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627–635.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 23*, 1122–1139.

Hauk, O., & Pulvermuller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology, 115*, 1090–1103.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermuller, F., & Rogers, T. T. (2006). [Q:] When would you prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience, 18*, 818–832.

Hernandez, A. E. (2009). Language switching in the bilingual brain: What's next? *Brain and Language, 109*, 133–140.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*, 638–650.

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: An fMRI study. *NeuroImage, 14*, 510–520.

Ibrahim, R. (2009). The cognitive basis of diglossia in Arabic: Evidence from a repetition priming study within and between languages. *Psychology Research and Behavior Management, 12*, 95–105.

Ibrahim, R., & Aharon-Peretz, J. (2005). Is literary Arabic a second language for native Arab speakers?: Evidence from semantic priming study. *The Journal of Psycholinguistic Research, 34*, 51–70.

Ibrahim, R., & Eviatar, Z. (2009). Language status and hemispheric involvement in reading: Evidence from trilingual Arabic speakers tested in Arabic, Hebrew, and English. *Neuropsychology, 23*, 240–254.

Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2002). The characteristics of Arabic orthography slow its processing. *Neuropsychology, 16*(3), 322–326.

Ibrahim, R., Eviatar, Z., & Aharon Peretz, J. (2007). Metalinguistic awareness and reading performance: A cross language comparison. *The Journal of Psycholinguistic Research, 36*, 297–317.

Junque, C., Vendrell, P., & Vendrell, J. (1995). *Differential impairments and specific phenomena in 50 Catalan-Spanish bilingual aphasic patient*. Pergamon.

Keatley, C. W., Spinks, J. A., & de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition, 22*, 70–84.

Khamis Dakwar, R., & Froud, K. (2007). Lexical processing in two language varieties, an event-related brain potential study of Arabic native speaker. In M. Mughazy (Ed.), *Perspectives on Arabic linguistics XX* (pp. 153–166). John Benjamins.

Khateb, A., Pegna, A. J., Michel, C. M., Landis, T., & Annoni, J. M. (2002). Dynamics of brain activation during an explicit word and image recognition task: An electrophysiological study. *Brain Topography, 14*, 197–213.

Khateb, A., Pegna, A. J., Landis, T., Mouthon, M. S., & Annoni, J. M. (2010). On the origin of the N400 effects: An ERP waveform and source localization analysis in three matching tasks. *Brain Topography, 23*, 311–320.

Khateb, A., Pegna, A. J., Michel, C. M., Mouthon, M., & Annoni, J. M. (2016). Semantic relatedness and first-second language effects in the bilingual brain: A brain mapping study. *Bilingualism: Language and Cognition, 19*, 311–330.

Khazen, M. (2016). *Diglossia in Arabic: Event-related potentials during a visual sentence semantic judgment task*. University of Haifa.

Klein, D., Zatorre, R. J., Milner, B., Meyer, E., & Evans, A. C. (1994). Left putaminal activation when speaking a second language: Evidence from PET. *Neuroreport, 5*, 2295–2297.

Kroll, J. F., & Tokowicz, N. (2001). The development of conceptual representation for words in a second language. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing*. Blackwell Publishers.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of chekerboard-evoked multichannels potential fields. *Electroencephalography and Clinical Neurophysiology, 48*, 609–621.

Mahfoudhi, A., Everatt, J., & Elbeheri, G. (2011). Introduction to the special issue on literacy in Arabic. *Reading and Writing, 24*, 1011–1018.

Marian, V., & Kaushanskaya, M. (2007). Language context guides memory content. *Psychonomic Bulletin & Review, 14*, 925–933.

Mountaj, N., El Yagoubi, R., Himmi, M., Ghazal, F., Besson, M., & Boudelaa, S. (2015). Vowelling and semantic priming effects in Arabic. *International Journal of Psychophysiology, 95*, 46–55.

Mouthon, M., Annoni, J. M., & Khateb, A. (2013). The bilingual brain. Review article. *Swiss Archives of Neurology and Psychiatry, 164*, 266–273.

Nevat, M., Khateb, A., & Prior, A. (2014). When first language is not first: An functional magnetic resonance imaging investigation of the neural basis of diglossia in Arabic. *The European Journal of Neuroscience, 40*, 3387–3395.

Ojemann, G. A. (1983). Brain organization for language from the perspective of electrical stimulation mapping. *Behavioral and Brain Sciences, 2*, 189–230.

Ojemann, G. A., & Whitaker, H. A. (1978). The bilingual brain. *Archives of Neurology, 35*(7), 409–412.

Palfreyman, D., & Khalil, M. A. (2003). "A funky language for teenzz to use:" Representing Gulf Arabic in instant messaging. *Journal of Computer-Mediated Communication, 9*(1), JCMC917.

Paradis, M. (1977). Bilingualism and aphasia. In H. Whitaker & H. Whitaker (Eds.), *Studies in neurolinguistics* (Vol. 3, pp. 65–121). Academic.

Paradis, M. (1983). *Readings on aphasia in bilinguals and polyglots*. Marcel Didier.

Paradis, M. (1998). Aphasia in bilinguals: How atypical is it? In P. Coppens, Y. Lebrun, & A. Basso (Eds.), *Aphasia in atypical populations* (pp. 35–66). Lawrence Erlbaum Associates.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology, 15*, 202–206.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S., et al. (2003). The role of age of acquisition and language usage in early, high proficient bilinguals: A fMRI study during verbal fluency. *Human Brain Mapping, 19*, 170–182.

Pratt, H., Abbasi, D. A., Bleich, N., Mittelman, N., & Starr, A. (2013a). Spatiotemporal distribution of cortical processing of first and second languages in bilinguals. I. Effects of proficiency and linguistic setting. *Human Brain Mapping, 34*, 2863–2881.

Pratt, H., Abbasi, D. A., Bleich, N., Mittelman, N., & Starr, A. (2013b). Spatiotemporal distribution of cortical processing of first and second languages in bilinguals. II. Effects of phonologic and semantic priming. *Human Brain Mapping, 34*, 2882–2898.

Proverbio, A. M., Zani, A., & Adorni, R. (2008). The left fusiform area is affected by written frequency of words. *Neuropsychologia, 46*(9), 2292–2299.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied PsychoLinguistics, 24*(3), 431–451.

Saiegh-Haddad, E. (2004). The impact of phonemic and lexical distance on the phonological analysis of words and pseudo-words in a diglossic context. *Applied PsychoLinguistics, 25*, 495–512.

Saiegh-Haddad, E. (2007). Linguistic constraints on children's ability to isolate phonemes in Arabic. *Applied PsychoLinguistics, 28*, 605–625.

Saiegh-Haddad, E. (2012). Literacy reflexes of Arabic diglossia. In M. Leiken (Ed.), *Current issues in bilingualism: Cognitive and socio-linguistic perspectives* (Vol. 5, pp. 42–55). Springer.

Saiegh-Haddad, E. (2018). MAWRID: A model of Arabic word reading in development. *Journal of Learning Disabilities, 51*(5), 454–462.

Saiegh-Haddad, E., & Haj, L. (2018). Does phonological distance impact quality of phonological representations? Evidence from Arabic diglossia. *Journal of Child Language, 45*(6), 1377–1399.

Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. Handbook of Arabic Literacy. In E. Saiegh-Haddad & M. Joshi (Eds.), *Handbook of Arabic Literacy: Insights and perspectives* (pp. 3–28). Springer-Dordrecht.

Saiegh-Haddad, E., & Joshi, M. (2014). *Handbook of Arabic Literacy: Insights and perspectives*. Springer-Dordrecht.

Saiegh-Haddad, E., & Schiff, R. (2016). The impact of diglossia on voweled and unvoweled word reading in Arabic: A developmental study from childhood to adolescence. *Scientific Studies of Reading, 20*(4), 311–324.

Saiegh-Haddad, E., & Spolsky, B. (2014). Acquiring literacy in a diglossic context: Problems and prospects. In *Handbook of Arabic literacy* (pp. 225–240). Springer.

Saiegh-Haddad, E., Levin, I., Hende, N., & Ziv, M. (2011). The Linguistic Affiliation Constraint and phoneme recognition in diglossic Arabic. *Journal of Child Language, 38*, 297–315.

Saiegh-Haddad, E., Shahbari-Kassem, A., & Schiff, R. (2020). Phonological awareness in Arabic: The role of phonological distance, phonological-unit size, and SES. *Reading and Writing, 33*(6), 1649–1674.

Shehadi, M. (2013). *Semantic processing in literary and Spoken Arabic: An event- related potential study*. University of Haifa.

Simon, G., Bernard, C., Lalonde, R., & Rebai, M. (2006). Orthographic transparency and grapheme-phoneme conversion: An ERP study in Arabic and French readers. *Brain Research, 1104*, 141–152.

Sinai, A., & Pratt, H. (2002). Electrophysiological evidence for priming in response to words and pseudowords in first and second language. *Brain and Language, 80*, 240–252.

Taha, H., & Khateb, A. (2013). Resolving the orthographic ambiguity during visual word recognition in Arabic: An event-related potential investigation. *Front, 7*, 821. https://doi.org/10.3389/fnhum.2013.00821

Taha, H., Ibrahim, R., & Khateb, A. (2013). How does Arabic orthographic connectivity modulate brain activity during visual word recognition: An ERP study. *Brain Topography, 26*, 292–302.

Taouk, M., & Coltheart, M. (2004). The cognitive processes involved in learning to read in Arabic. *Reading and Writing: An Interdisciplinary Journal, 17*, 27–57.

Warschauer, M., Said, G. R. E., & Zohry, A. G. (2002). Language choice online: Globalization and identity in Egypt. *Journal of Computer-Mediated Communication, 7*(4), JCMC744.