Received: 28 February 2022 | Revised: 20 July 2022 | Accepted: 24 July 2022

DOI: 10.1111/ejn.15781

RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

# Behavioural and electrophysiological analyses of written word processing in spoken and literary Arabic: New insights into the diglossia question

**Samer Andria**[1,2] | **Bahaa Madi-Tarabya**[1,2] | **Asaid Khateb**[1,2]

[1]The Unit for the Study of Arabic Language, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel

[2]Department of Learning Disabilities, Faculty of Education, University of Haifa, Haifa, Israel

**Correspondence**
Asaid Khateb, The Unit for the study of Arabic language, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, 199 Abba Khoushy Ave., Mount Carmel, Haifa 3498838, Israel.
Email: akhateb@edu.haifa.ac.il

**Funding information**
Edmond J. Safra Brain Research Center for the Study of Learning Disabilities; Israeli Science Foundation, Grant/Award Number: 2695/19

Edited by: Edmund Lalor

**Abstract**

Diglossia in Arabic describes the existence and the use of two varieties of the same language: spoken Arabic (SA) and literary Arabic (LA). SA, the dialect first spoken by Arabic native speakers, is used in non-formal situations for everyday conversations, and varies from one region to another in the Arabic world. LA, acquired later in life when the children learn to read and write at school, is used for formal purposes such as media, speeches in public and religious sermons. Previous research showed that, in the auditory modality, SA words are processed faster than LA ones. In the visual modality, written LA words are processed faster than SA ones, the latter comparing with low-frequency words. This study analysed event-related potentials (ERPs) during the processing of high-frequency (LAHF), LA low-frequency (LALF) and SA high-frequency words (SAHF) in a visual lexical decision task. Faster reaction times were observed for LAHF, followed by SAHF and then by LALF. ERPs showed a modulation of the early components starting from the P100 component and of the late P600 component, supposedly related to memory processes. These findings, indicating that processing written SAHF words was largely comparable with processing of LALF, are discussed in the context of Arabic diglossia.

**KEYWORDS**

diglossia, event related potentials (ERPs), lexical decision task, literary Arabic (LA), N170, spoken Arabic (SA), visual word processing, word frequency

## 1 | INTRODUCTION

Psycholinguistic research has shown that reading accuracy and speed depend on a number of cognitive and linguistic factors across different languages (Coltheart, 2005, 2006). Among the linguistic factors, frequency (Forster & Chambers, 1973; Gordon, 1983; Rubenstein et al., 1970; Yap & Balota, 2007), and

---

**Abbreviations list:** EEG, electroencephalography; ERPs, event-related potentials; LA, literary Arabic; SA, spoken Arabic; LAHF, literary Arabic high-frequency words; LALF, literary Arabic low-frequency words; SAHF, spoken Arabic high-frequency words.

lexicality effects are by far the most frequently investigated (Balota et al., 2004; Coltheart et al., 1993, 2001). Word frequency is hypothesized, according to the dual-route model (Coltheart et al., 1993), to determine the route to be used for reading. Reading a non-frequent word (unknown or a non-word) necessitates translating the graphemes to phonemes (the non-lexical route). Reading and/or recognizing a frequent word passes through retrieval of the specific word orthographic/phonemic pattern stored in the mental lexicon, allowing thus more accurate and faster reading than the decoding process (see, e.g., Garlock et al., 2001; Gernsbacher, 1984; Perea et al., 2005). Duyck et al. (2008) have suggested that word frequency effect is the result of implicit learning and of the strengthening of lexical representations by words repetition (Perea et al., 2005). Therefore, high-frequency words are recognized faster than low-frequency ones because of the low surpass threshold for the high-frequency words (Besner & McCann, 1987; Monsell et al., 1989; Morton, 1970; Murray & Forster, 2004). Psychophysiological studies using event-related potential (ERP) analysis have suggested that effects of word frequency during visual word recognition could be found during the early stages of information processing (e.g., ∼132 ms [Sereno et al., 1998] or 150–190 ms [Hauk & Pulvermüller, 2004; Strijkers et al., 2010]) but also during later stages (e.g., ∼240–300 ms [Proverbio et al., 2008], ∼320–360 ms [Hauk & Pulvermüller, 2004] and ∼400 ms after word presentation [Van Petten & Kutas, 1990]). To give only few examples, a modulation by frequency was reported for the amplitude of the N170 component, with responses over the left scalp being larger for frequent than non-frequent words (Simon et al., 2007). Proverbio et al. (2008) reported differences between high-frequency and low-frequency words in the amplitude of N2 (240–300 ms) and N3 components (300–360 ms) which appeared larger for high-frequency words over left lateral occipital sites. In functional brain imaging studies, word frequency had also been shown to induce differences in brain activation. Keller et al. (2001) for instance, reported about higher activation for low-frequency words than for high-frequency words in the occipito-temporal brain regions. Kronbichler et al. (2004) showed negative correlation between word frequency and the activation of the left middle and posterior fusiform gyrus, suggesting less computational demands when encountering high-frequency words. A similar finding was reported by Hauk et al. (2008), who also suggested that frequency and brain activation were negatively correlated in several brain areas involved in visual word processing, including in the left fusiform gyrus and the bilateral inferior frontal gyri. Word frequency effects had been investigated in multiple languages including in English (Yap et al., 2012), German (Bronk et al., 2013; Brysbaert et al., 2011), French (Maïonchi-Pino et al., 2010), Italian (Proverbio et al., 2008, 2004), Hebrew (Koriat, 1985) and Chinese (Kuo et al., 2003) and also among bilinguals (Brysbaert et al., 2017; Diependaele et al., 2013). To date, very few studies have been conducted on the diglossic Arabic language, which is considered by some authors as a particular form of bilingualism (Ibrahim, 2009; Ibrahim et al., 2007).

The diglossic situation in the Arabic language refers to the existence of two varieties of the same language that are used in different situations (Ferguson, 1959). The spoken Arabic (SA) variety, considered as the low variety, is the dialect first spoken by Arabic native speakers, used at home and in non-formal situations in everyday conversations, and varies from one region to another in the Arabic world. The literary Arabic (LA), referred to by some authors as modern standard Arabic, MSA, or StA (Saiegh-Haddad & Henkin-Roitfarb, 2014), the high variety, is acquired later in life when the children learn to read and write at school and is used for formal purposes such as media, speeches in public and religious sermons (Saiegh-Haddad, 2012; Saiegh-Haddad & Schiff, 2016). In addition to the difference in the history of acquisition of SA and LA and to the patterns of use in everyday life, Saiegh-Haddad (2003) had suggested that differences between SA and LA exist in the different linguistic domains (phonology, morphology and semantics). At the phonological level for instance, there are multiple phonemes that exist in the LA but not in the SA in some dialects, for instance, the phoneme /ð/ which is pronounced in LA is replaced by the phoneme /d/ or /z/ in SA in some dialects (LA /ðawq/ [taste] vs. SA /dawq/ or /zawq/) (Saiegh-Haddad et al., 2020). At the lexico-semantic level, Saiegh-Haddad and Spolsky (2014) have analysed a corpus of words from 5-year old Arabic speakers and found that 40% of the children lexicon include unique SA words with non-classic written form, 40% of cognate words in SA and LA and ∼20% of words being identical—words that have the same phonological form and meaning in SA and LA (see also Saiegh-Haddad, 2018).

Based on empirical data, some authors have proposed that Arabic diglossia could be considered as a form of bilingualism. In one study, conducted with children in kindergarten and first grade (Eviatar & Ibrahim, 2000), metalinguistic skills of Arabic children who were exposed to both SA and LA were compared with the abilities of bilinguals (exposed to two different languages) and with those of monolingual children. The findings showed that Arabic children abilities compared with that of the Russian–Hebrew bilinguals, but differed from Hebrew monolinguals (see also other results in Ibrahim et al., 2007). The authors concluded that, because Arabic-native speakers behaved as bilinguals, they could be considered as bilinguals (Eviatar & Ibrahim, 2000). In another study, Ibrahim and Aharon-Peretz (2005) examined inter- and intra-language (semantic) priming effects in an auditory lexical decision among 11th and 12th grade native-Arabic speaking students, who also had Hebrew as their formal second language. Auditory presentation of SA, LA and Hebrew stimuli enabled comparisons between SA and LA varieties, as well as comparison of both varieties to Hebrew. The results indicated that priming effects were larger when prime words were in SA and target words were either in LA or in Hebrew, than when primes were in LA or Hebrew and target words in SA. In addition, they showed that priming effects for LA and Hebrew were very similar, indicating that both languages behaved as second languages (see similar results in Ibrahim (2009). The priming effects by SA observed in this study were interpreted in the light of previous observations in bilinguals (Gollan et al., 1997; Keatley et al., 1994) which show that forward priming (i.e., from the dominant first language [L1]) to the less dominant second language (L2) are larger than backward priming (i.e., from L2 to L1). Taken as evidence that L1 words more readily initiate conceptual processing than words in L2 (Kroll & Tokowicz, 2001), the authors suggested that the two varieties of Arabic are cognitively represented in two separately organized lexicons and that literate speakers of Arabic behave as bilinguals, with SA as their L1 and LA as their L2. Although this conclusion appeared to explain the results obtained in the auditory modality, other results have indicated that LA and SA status (mimicking L1 and L2) might depend on the modality of the presentation of the word stimuli: SA words showing a pattern of dominant responses in the auditory modality and LA words showing a pattern of dominant responses in the visual written modality. Indeed, in a prior study that used visual presentation of written LA and SA words, Bentin and Ibrahim (1996) showed that reading aloud SA words was slower than LA ones, and that SA words were processed as low-frequency LA ones. More recently, a study using functional magnetic resonance imaging (fMRI) during a semantic categorization task that included written words from LA and SA (and Hebrew) confirmed the dominant status of LA words during visual word processing (Nevat et al., 2014). Both in terms of behavioural (RTs and accuracy) and functional responses, the results showed that LA functioned as dominant language variety. Differences in response accuracy and response speed variance between LA and SA have also been reported in another recent study using a similar semantic categorization task (Tarabya et al., 2021).

Although at the behavioural level, the question of the dominance appears to have some reasonable responses, the question remains as to how and when LA and SA processing differs in the brain. In this study, the aim was to investigate, using event related potentials (ERPs), when LA and SA processing differs during the time course of word recognition. ERP analyses are widely used as a suitable tool for monitoring brain activity during visual word recognition. In language studies using the visual modality more particularly, several ERP components have been shown to be sensitive to linguistic factors. Among these components, the N170 component for instance (referred also in the ERP literature as N1 component, or word recognition potential; see Spironelli et al., 2020; Spironelli & Angrilli, 2009), which peaks around 170 ms after stimulus presentation, is a negative occipito-temporal response that was linked during the processing of visual words to the orthographic processing stage (Bentin et al., 1999). Previous studies have shown that this component's amplitude and peak latency could be modulated by various linguistic manipulations including among other things words' repetition, orthographic features and word frequency (Simon et al., 2004; Taha et al., 2013; Taha & Khateb, 2013). During the presentation of orthographic stimuli, this component is maximal at left occipito-temporal sites, while during the processing of non-orthographic such as objects (Khateb et al., 2002) and human faces (Bentin et al., 1996), this response is maximal at the right sites (Dundas et al., 2014; Proverbio et al., 2007; Simon et al., 2007). In this regard, Simon et al. (2007) suggested that the N170 is an occipito-temporal component which represents the first step where faces, objects and words processing are differentiated along the ventral system of the brain. In addition to the studies of healthy participants, the N170 was also described in clinical studies such as developmental dyslexia (DD), aphasia and schizophrenia (Blasi et al., 2002; Ibáñez et al., 2011; Ji et al., 2019; Spironelli et al., 2010). To give only few examples from the literature on dyslexia, the N170 amplitude was described to be reduced among dyslexic readers (Blau et al., 2007; Hasko et al., 2013). A relation between the N170 amplitude among DD and the severity of the dyslexia was also reported (Mahé et al., 2012). In the same vein, Korinth and Breznitz (2014) reported about a larger N170 amplitude among fast compared with slow readers. Finally, Spironelli et al. (2010) reported an increase of the component amplitude at left recording sites in DD following a phonological training, suggesting a cortical reorganization of the reading network. During the late steps of information processing, in addition to the N400 ERP component, which was described originally following semantic anomalies during sentence processing (Kutas &

Hillyard, 1980), a late posterior positive component (referred to as the P600 or the P6) with an onset at ~500 ms after stimulus presentation had often been reported in the context of lexical decision experiments. In some studies, the P600 was elicited as a result of misspellings and pseudo-homophones (Münte et al., 1998; Stites et al., 2016; Taha & Khateb, 2013; Van de Meerendonk et al., 2011). The P600 component was described also as a late waveform induced by syntactic anomalies during sentence processing as a reparation process (Friederici et al., 1996). In tasks involving word recognition this response was proposed to reflect late memory monitoring processes (Kaan & Swaab, 2003; Taha & Khateb, 2013).

The present study relied on previous results on Arabic in the visual modality (Bentin & Ibrahim, 1996; Nevat et al., 2014; Tarabya et al., 2021) and on other findings from bilinguals (Duyck et al., 2008) indicating that high-frequency L2 words were processed as low-frequency L1 words. Because differences between the processing of written LA and SA words were explained in terms of the subjects' lower visual familiarity with written SA word patterns, this study sought to examine the extent to which difference between the processing of SA versus LA words compares with the visual processing of low- versus high-frequency words in LA. For this aim, behavioural and electrophysiological (ERP) analyses were conducted during a visual lexical decision task in native Arabic speakers. The task included LA high-frequency (LAHF) words, LA low-frequency (LALF) words and SA high-frequency (SAHF) words. At the behavioural level, we predicted that LAHF will induce shorter RTs (and higher accuracy) compared with LALF and SAHF. In terms of ERPs, it was hypothesized that differences will be found between LAHF and LALF words due to frequency effects. Also, assuming that difference between SAHF and LAHF words is mainly due to difference in the visual familiarity (lower for SA words), it was predicted that processing of SAHF words will be comparable with the processing of LALF words. Based on the previous ERP studies documenting the modulation of the different components by word frequency, we expected to find differences starting already around the N170 component, previously shown to differentiate between words based on the their orthographic and frequency characteristics (Hasko et al., 2013; Maurer et al., 2005; Xue et al., 2019), and during the P600 component (Rugg, 1990; Taha & Khateb, 2013; Young & Rugg, 1992).

## 2 | MATERIALS AND METHODS

### 2.1 | Participants

Twenty-eight healthy young students (16 women and 12 men) aged between 18 and 24 years participated in the study ($M = 22.7$ years, $SD = 1.54$). All participants were native Arabic speakers, self-declared right-handed and had normal or corrected-to normal vision with no history of dyslexia or any other neurological or psychiatric diseases. The participants were recruited from the University of Haifa and were paid for their participation (35 NIS per hour), or received bonus points. Five subjects were excluded from behavioural analysis because of poor accuracy (accuracy rate below 60% in at least one experimental word condition), and other four subjects were excluded from the ERP analysis due to noisy signals.

### 2.2 | Stimuli and procedure

The stimuli consisted of three written word lists including 60 high-frequency LA (hereafter LAHF) words, 60 low-frequency LA (LALF) words and 60 high-frequency SA (SAHF) words. LA words were selected on the basis of a questionnaire that was filled by 26 Arabic speaking adult students (who did not participate in the experiment). They were asked to assess the familiarity/frequency (hereafter we will use frequency) of 136 written LA words using a 0–5 scale (0 for not frequent/not familiar, 5 for highly frequent/familiar). After this rating, the words were categorized on the basis of the mean frequency of each item into two groups of 60 words each, when LAHF words had a mean familiarity score >3 ($M = 4.46$, $SD = .27$; range = 3.5–4.9), and LALF words had a mean familiarity score <3 ($M = 1.75$, $SD = .75$; range = .88 = 2.63). In addition, the same participants were asked to assess the familiarity/frequency of 70 written SA words, which are highly familiar spoken words of which 60 words ranked above 4 (0–5 scale) were selected ($M = 4.8$, $SD = .17$; range = 3.6–5). The average word length was matched between word conditions (for LAHF $M = 4.3$, $SD = .7$; for LALF $M = 4.2$, $SD = .8$; and for SAHF $M = 4.25$, $SD = 1.01$). Also, three sub-lists of 60 pseudowords each were built on the basis of real words by replacing one or two letters keeping the pronunciation of the stimuli acceptable. Altogether, these provided a list of 360 words, which were divided randomly into two blocks and were distributed randomly in each block. The duration of each block was ~8 min; therefore, the whole experiment took approximately 15 min.

The experiment took place in an isolated, sound-shielded room. The words were presented on a computer screen using E-Prime software (v. 2.0.10.147, Psychology Software Tools, Inc., www.pstnet.com/, PA, USA). Each simulation trial lasted about ~2.5 s and started with a 500-ms fixation cross, which was followed by the stimulus during 150 ms. A blank screen then appeared for

1850 ms to allow for the participant's response. Participants were instructed to make lexical decisions on the visual stimuli by deciding (as quickly and accurately as possible) after each string of letters whether or not these letters constitute a word they know.

During the experiment, participants were seated at ~120 cm from the screen and asked to fixate at the central cross and to read silently the words. After the presentation of each stimulus, participants were required to respond by pressing one of two buttons with their right-hand middle and index fingers (half of the subjects responded for words with their index finger and for pseudowords with their middle finger and the other half responded the other way). They were also instructed to try not to blink during the visual presentation of the stimulus. All participants accomplished a training session consisting of 18 words (3 words from each condition and pseudowords) in order to ensure that they have understood the task demands and to familiarize with the button presses.

## 2.3 | EEG recordings and ERP pre-processing

Electroencephalographic (EEG) recordings were collected continuously using a 64 channel BioSemi Active Two system and the ActiveView recording software (www.biosemi.com). Pin-type electrodes were mounted on a customized BioSemi head-cap, using an electrode gel and arranged according to the 10–20 international system. Horizontal eye movements were monitored using two flat electrodes placed on the sides of the eyes. Two additional electrodes, placed underneath and above the right eye, monitored vertical eye movement and blinks. The EEG signals were collected reference-free at a 2048-Hz sampling rate (i.e., Biosemi active electrodes), with a .25 high-pass filter, amplified and digitized with a 24-bit AD converter. ERP epochs were averaged offline using the Brain Vision Analyzer software (Brain-products: http://www.brainproducts.com/). EEG data were first filtered (low pass: 30 Hz; and high pass: 1 Hz), ocular artefacts' corrected using the Gratton et al.'s (1983) method. This is an offline procedure, implemented in Brain Vision Analyzer software, which uses EOG channels and EEG recordings for each individual trial to estimate a factor, which is the relation between EOG and EEG data (called also a propagation factor). This factor is calculated separately for blinks and eye movements (for more information, see Gratton et al., 1983).

Artefacts also were rejected based on the next steps: (i) gradients maximum allowed voltage steps of 50.00 μV with a 200 ms before or after the event, (ii) maximal allowed absolute differences of 200 μV in intervals of 200 ms, (iii) amplitude higher than 70 μV and lower than −70 μV were rejected, and (v) the lowest activity allowed (max–min) was .5 μV within an interval length of 100 ms. After the artefact rejection, the ERP epochs were averaged separately for each word condition from −100 ms before stimulus onset to 900 ms post-stimulus only for trials with correct responses and re-referenced to the average-reference (Lehmann & Skrandies, 1980). The following analyses were conducted using the Cartool software© (v.3.51; http://brainmapping.unige.ch/ Cartool.php). The individual ERP of each word condition were first down-sampled to 512 Hz and baseline-corrected (using the 100 ms pre-stimulus interval). The resulting individual ERPs were then used to compute the grand mean ERP of each word condition and conduct component analysis.

## 2.4 | ERP analysis

Because the study focused on word recognition without any a priori hypothesis and question about pseudowords, the analysis presented here below aimed at determining ERP differences between LAHF, LALF and SAHF for word conditions only. For this purpose, we first compared the response amplitude between the conditions for the components P1, N170, P2, N2 and P3 (see Figure 1). For all these components, a sub-set of three left posterior electrodes (PO7, PO3 and O1) and three right electrodes (PO8, PO4 and O2) that best displayed these components were selected. This selection was based on the visual inspection of the superposition of the grand-mean ERPs of the three language conditions (LAHF, LALF and SAHF, Figure 1). The mean response amplitude was then computed from each of these electrodes: (i) for the P1 between ~115 and 125 ms, (ii) for the N170 between ~170 and 180 ms, (iii) for the P2 between ~280 and 300 ms, (iv) for the N2 between ~320 and 340 ms and (v) for the P3 between ~340 and 365 ms. The mean signals for each component from each of these electrodes were then subjected to statistical analysis using repeated measures analysis of variance (ANOVA). For the late responses, we analysed the late P600 component which best appeared at centro-parietal electrodes. For this, we computed the mean signal between 450 and 600 ms from each of three left (CP3, CP1 and P3) and three right (CP4, CP2 and P4) centro-parietal electrodes (see Figure 3).

The analysis of components' latency was conducted based on the averaged signal (as a region of interest) of the three left (PO7, PO3 and O1) and the averaged signal of three right (PO8, PO4 and O2). We then determined successively in each individual ERP of each condition the

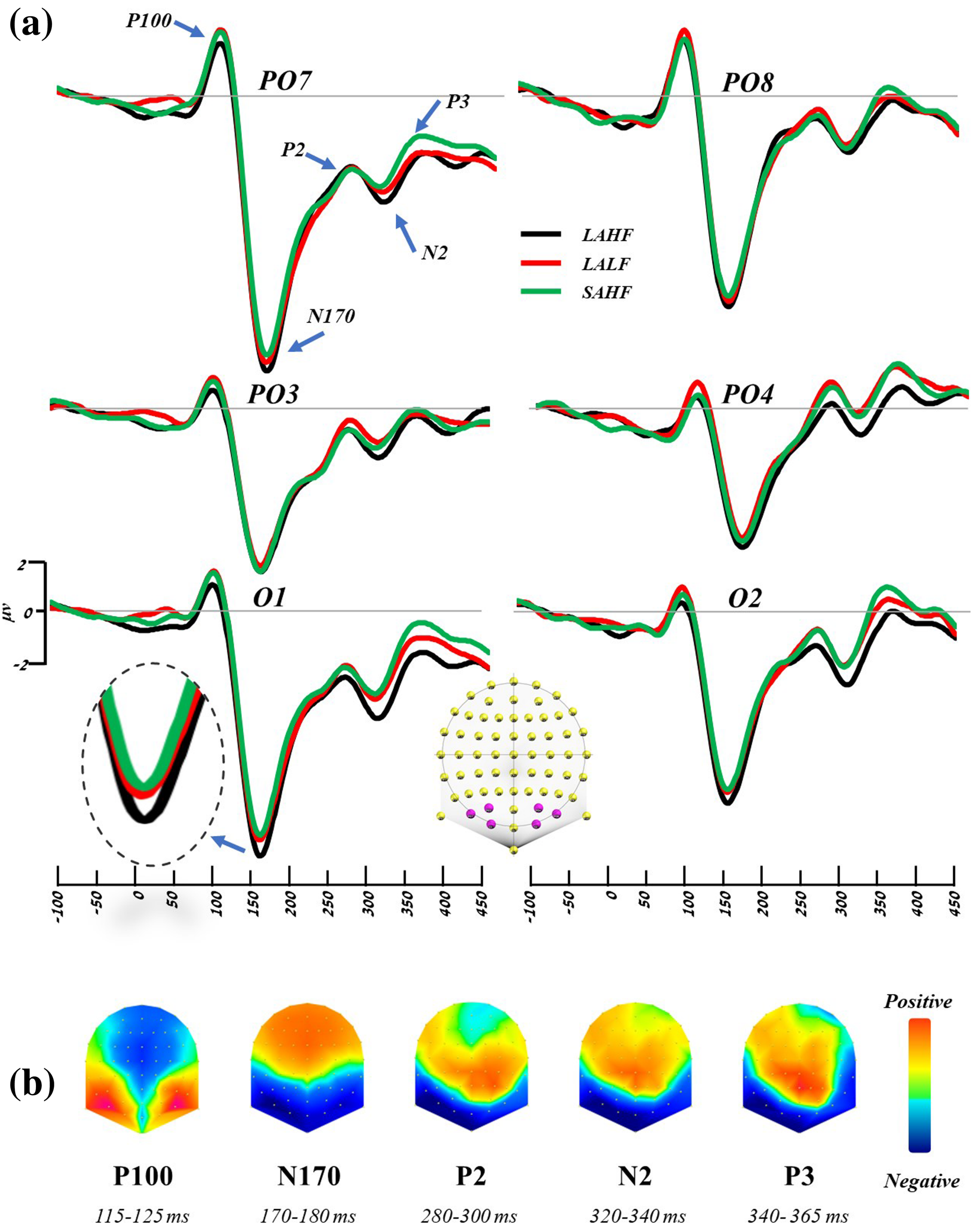

**FIGURE 1** (a) Superimposition of the grand mean event-related potential (ERP) traces for the three language condition (black: LAHF, red: LALF, green: SAHF) from −100 to 450 ms from three left posterior and three right posterior sites electrodes which best exhibited the succession of the early (P1 to P3) ERP components (see blue arrows in PO7). The inset in the middle of the figure illustrates the location of the selected electrodes. The inset below electrode O1 shows an enlargement of the difference between conditions during the N170. (b) Illustration of the topographic maps for the mean signal for the successive components from the P100 to the P3. Note the typical topography of the P100 component showing the characteristic posterior positivity and anterior negativity and the inversion in the N170 component map which shows a posterior negativity and anterior negativity. Numbers below the maps indicate the time period of analysis for each component. 2-D maps are scaled to their maxima (see color scale in the right: red for positive potentials and blue for negative potentials) and are presented with left ear left and right ear right, the nasion up and the inion down

latency of the P1 as the most positive time point occurring ~100 ms, then the next most negative time point ~N170 before the graph changed its slope (from negative to positive), and then again, the next most positive point for P2 before the slope has changed from positive to negative and so on for the N2 and P3. Statistical analyses were conducted on the latency values for each component using ANOVA. As for the P600 component, we first computed, as a region of interest in each participant, the average of 10 centro-parietal electrodes around CPz and Pz and which best displayed the P600 component (see Figure 3). The resulting individual centro-parietal 'P600' waves were then low-pass filtered at 5 Hz to avoid the selection of spurious peaks (see Khateb et al., 2010; Moreno & Kutas, 2005). The latency values for the P600 were then determined as the most positive time point between ~450 and 700 ms. Statistical analysis was then conducted on these values using one-way ANOVA with language condition as within-subject factor. Note that all repeated measures with more than one degree of freedom were first subjected to the Geisser–Greenhouse correction. In all statistical analyses, only adjusted significant main effects and interactions (at $p < .05$ corrected) were considered for follow up by post-hoc Bonferroni tests.

## 2.5 | Behavioural analysis

The individual median reaction time (RT) and accuracy (percentage of correct responses) were computed separately for each word condition. A one-way ANOVA with language condition as within-subject factor was then conducted on individual RTs and accuracy measures.

# 3 | RESULTS

## 3.1 | Behavioural results

Table 1 presents the descriptive statistics (means and standard deviation) for RTs and accuracy, across participants for words in each language conditions.

TABLE 1 Mean of reaction times (in ms) and mean accuracy (in %, ±standard deviation, $n = 19$) for words in the three language conditions

| | LA-HF | LA-LF | SA-HF |
|---|---|---|---|
| Reaction times (in ms) | 670 (81) | 789 (119) | 722 (91) |
| Accuracy (in %) | 95 (3.9) | 69 (12.4) | 92 (5.1) |

Abbreviations: LAHF, literary Arabic high frequency word; LALF, literary Arabic low frequency word; SAHF, spoken Arabic high frequency word.

## 3.2 | RTs analysis

The 1 × 3 repeated measures analysis of variance (ANOVA) performed on the individual median RTs using language condition (LAHF, LALF and SAHF) words showed a highly significant effect of language condition ($F$ [2, 36] = 41.38, $p =$ .000, $\eta_p^2 = .697$). Bonferroni post-hoc tests showed that RTs were significantly faster for LAHF words than for LALF ($p = .000$) and for SAHF ($p = .014$), with the difference between LALF and SAHF being also significant ($p = .000$).

### 3.2.1 | Accuracy analysis

The one-way ANOVA performed on the accuracy for words revealed a highly significant effect for language variety ($F$ [2, 36] = 80.79, $p =$ .000, $\eta_p^2 = .818$) due to a higher accuracy in LAHF than in LALF ($p = .000$), and higher in SAHF than in LALF ($p = .000$), with difference between LAHF and SAHF failing to reach significance ($p = .48$).

## 3.3 | ERP waveform analysis

Figure 1a displays the superimposition of the grand mean ERP traces (enlarged from −100 to 450 ms post-stimulus) from three left posterior electrodes and three right posterior electrodes (PO7, PO3 and O1 and PO8, PO4 and O2, respectively, see inset for electrodes' location), which best exhibited the successive P1, N170, P2, N2 and P3 components in the different conditions (see arrowheads on electrode PO7). As described in the methods, the mean signal amplitude was computed for each of the successive components from the three left and three right electrodes in the time period around the peak of each component in the grand mean ERPs. Figure 1b illustrates the mean topographic maps from the grand-mean ERP map series (all word conditions collapsed) characteristic of each of the successive components (see time periods below the maps). The mean amplitude values computed individually in each of these components' period in each language condition were subjected to 3 × 2 × 3 repeated measures analysis of variance using language condition (3: LAHF, LALF and SAHF), hemisphere (i.e., 2: left vs. right), and electrodes (for 3 left and 3 right) as within-subject factors. The results of these analyses are detailed here below and in Figures 2 and 3, which illustrate only significant language condition effects.

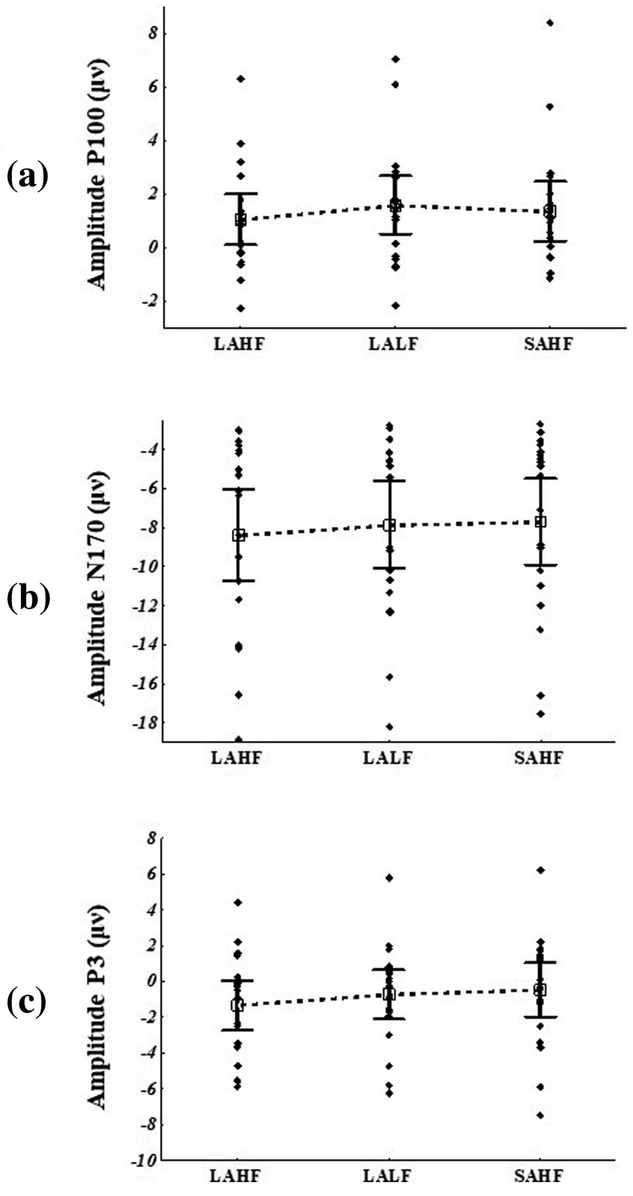

**FIGURE 2** (a–c) Graphs illustrating the significant language condition effects for the mean amplitude of the P1 (a), the N170 (b) and the P3 components (c) analysed from the same electrodes (and time periods) as in Figure 1 (see text for detailed statistics using Bonferroni post-hoc comparisons). In these graphs, the mean amplitude is shown by □, with bars denoting ±.95 confidence intervals and ◊ representing the individual values

### 3.4 | Amplitude and latency analysis of the P1 component

The 3 × 2 × 3 ANOVA performed on the mean amplitude of the P1 (computed between ~115 and125 ms) showed a significant main effect of language condition ($F[2,36] = 4.21$, $p = .023$, $\eta_p^2 = .19$). As illustrated in Figure 2a, amplitude varied over word conditions, with the highest amplitude for LALF, followed by SAHF and then by LAHF, with significant differences appearing only between LALF and LAHF ($p = .019$). A significant main effect was found also for electrode ($F[2,36] = 8.96$, $p = .0025$, $\eta_p^2 = .33$) due to varying P1 amplitude across the different electrodes with the highest amplitude for PO7 in the left (1.93 μv) and PO8 in the right (2.11 μv). No interaction was found between language condition and the analysis factors.

The 3 × 2 ANOVA performed on the latency of the P1 (see methods) showed only a significant main effect of language ($F[2,36] = 3.92$, $p = .043$, $\eta_p^2 = .179$). This effect was due to the fact that P1 in LAHF tended to peak earlier ($M = 109$ ms, $SD = 16$) in LALF ($M = 114$ ms, $SD = 15$, $p = .073$) and in SAHF ($M = 114$ ms, $SD = 14$, $p = .052$), with no difference between the two latter ($p < 1$). No interaction was found between the factors.

### 3.5 | Amplitude and latency analysis of the N170 component

The 3 × 2 × 3 ANOVA performed on the amplitude of the N170 (between ~170 and 180 ms) showed a main effect of electrode ($F[2,36] = 12.09$, $p = .0001$, $\eta_p^2 = .40$) and a significant interaction was found between language condition and electrode ($F[4,72] = 3.29$, $p = .029$, $\eta_p^2 = .154$), with no significant hemisphere effect. Bonferroni post-hoc test showed that significant language condition effects were observed at PO7/PO8 and O1/O2 electrode pairs. At electrodes PO7/PO8, language differences were found significant only between LAHF and SAHF ($p = .002$). The language condition differences at electrode O1/O2, illustrated in Figure 2b, were due to a higher N170 (negative) amplitude for LAHF than for LALF ($p = .002$) and for SAHF ($p = .0001$), with no difference between the latter two ($p = 1$). The 3 × 2 ANOVA performed on the latency of the N170 showed neither an effect of language condition ($p = .67$) nor an effect of hemisphere ($p = .60$).

### 3.6 | The P2-N2-P3 component complex analyses

Because the visual inspection of the grand-mean traces suggested the same pattern of response across the three language conditions and the three successive components (see electrodes O1 and O2 in Figure 1a), similar analyses were conducted on the mean amplitude of the components P2 (between ~270 and 300 ms), N2 (~310 and 340 ms) and P3 (between ~340 and 400 ms).

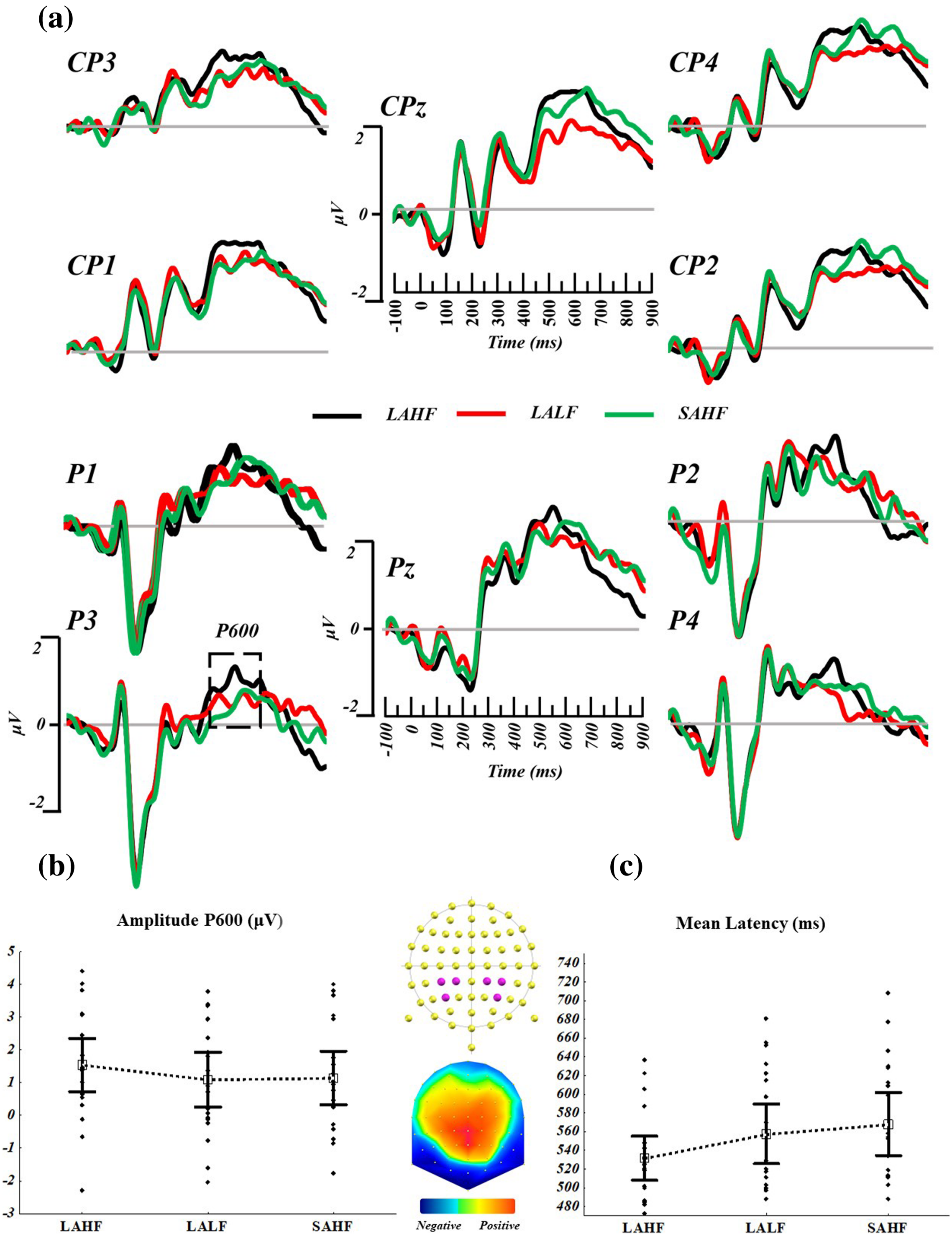

**FIGURE 3** (a) Superimposition of the grand mean ERP traces (from -100 to 900 ms) for the three language conditions (black: literary Arabic high-frequency [LAHF], red: literary Arabic low-frequency [LALF], green: spoken Arabic high-frequency [SAHF]) from four left, two midline and four right centro-parietal sites which best exhibited the ERP P600 component (see dashed box at P3). (b) Bar plot graph illustrating the significant language condition effect on the mean amplitude (with error bars) of the P600 computed on the basis of the mean signal from the left and right electrodes (see inset for these electrodes' location). The 2-D topographic map illustrates the average electric field configuration for the P6 period analysed (see colour scale in the right: red for positive and blue for negative potentials, left ear left and right ear right, the nasion up and the inion down). (c) Bar plot graph illustrating the significant language condition effect on the mean latency (with error bars) of the P600 computed on the basis of an average wave of the 10 electrodes (see text for details and statistics using Bonferroni post-hoc comparisons). In the two graphs (b,c), the mean is shown by □, with bars denoting ±.95 confidence intervals and ◊ representing the individual values

#### 3.6.1 | Amplitude and latency analysis of the P2

The ANOVA conducted on the mean amplitude of the P2 showed neither a language effect ($p = .27$), nor a significant interaction between this and the other factors (all $p > .05$). The ANOVA performed on the latency of the P2 showed a main effect of language condition ($F[2,36] = 3.85$, $p = .031$, $\eta_p^2 = .176$). In the average, LAHF ($M = 260$ ms, $SD = 28$) peaked slightly but not significantly earlier than LALF ($M = 267$ ms, $SD = 30$, $p = .39$) and significantly earlier than SAHF ($M = 272$ ms, $SD = 41$, $p = .027$), with the latter two showing no difference ($p = .69$). No hemisphere effect was found ($p = .65$) and no significant interaction was found between language and hemisphere ($p = .076$).

#### 3.6.2 | Amplitude and latency analysis of the N2

The ANOVA conducted on the mean amplitude of the N2 (computed between ~310 and 340 ms) again showed neither a significant effect of language condition ($p = .094$), nor a significant interaction between this and the other factors (all $p > .05$). The $3 \times 2$ ANOVA performed on the N2 time latency showed no effects of language condition ($p = .37$) or hemisphere ($p = .23$), and no interaction between the two factors ($p = .95$).

#### 3.6.3 | Amplitude and latency analysis of the P3

The ANOVA conducted on the mean amplitude of the P3 (computed between ~340 and 400 ms) showed a significant although small language condition effect ($F[2, 36] = 3.27$, $p = .049$, $\eta_p^2 = .154$), a hemisphere effect ($F[1, 18] = 7.07$, $p = .016$, $\eta_p^2 = .282$) and an electrode effect ($F[2, 36] = 8.40$, $p = .004$, $\eta_p^2 = .318$), but with no significant interaction between language condition and the other factors (hemisphere and electrodes). As illustrated in Figure 2c, which displays the average P3 amplitude (across electrodes and hemispheres), post-hoc comparison showed no significant difference between LAHF and LALF ($p = .37$) but with SAHF ($p = .048$). No significant difference between the latter two conditions ($p = 1$). Finally, the ANOVA performed on the P3 time latency showed neither an effect of language condition ($p = .82$) nor of hemisphere ($p = .82$), and no interaction between the two factors ($p = .55$).

### 3.7 | Amplitude and latency analysis of the P600 component

Figure 3a displays the superimposition of the grand mean ERP traces (enlarged from −100 to 900 ms post-stimulus) from 10 centro-parietal electrodes which best showed P6 component in the different conditions (see dashed boxes at electrode P3). The amplitude of this late component was analysed as the mean signal between 450 and 600 ms from three left (CP3, CP1 and P3) and three right (CP4, CP2 and P4) electrodes (see Figure 3a). The rational for choosing left and right electrodes (without central ones) aimed at testing a possible lateralization of the P6. The $3 \times 2 \times 3$ ANOVA performed on the P600 mean amplitude using language condition, hemisphere and electrode as within-subject factors showed a significant effect of language condition ($F[2,36] = 3.94$, $p = .03$, $\eta_p^2 = .180$), with no interaction between this and the other factors. As illustrated in Figure 3b (left graph), this effect was attributable to the higher P600 amplitude in LAHF than in LALF ($p = .044$) and SAHF ($p = .087$), with no difference between the latter two ($p = 1$). A significant electrode effect was observed ($F[2,36] = 7.19$, $p = .002$, $\eta_p^2 = .285$) due to varying amplitudes across electrodes. No effect of hemisphere was observed ($p = .85$) and no interaction was found between language condition and the other factors.

For determining the time latency of the P600, an average wave trace was computed on the basis of the signal of 10 centro-parietal electrodes (Figure 3a defined as a region of interest. From this wave, the P600 peak latency was measured as the most positive time point between 450 and 700 ms. The one-way ANOVA performed on this latency measure showed a significant effect of language ($F[2,36] = 4.43$, $p = .020$, $\eta_p^2 = .197$). Bonferroni post-hoc tests showed that P600 peaked earlier in LAHF ($M = 532$ ms) than in LALF (although non-significantly; $M = 558$ ms, $p = .14$) and in SAHF (M = 568 ms, $p = .019$), with no difference between the latter two ($p = 1$, see Figure 3b).

## 4 | DISCUSSION

This study aimed at examining how difference between the processing of SA versus LA written words compares with the visual processing of low- versus high-frequency words in LA. For this purpose, behavioural and electrophysiological (ERP) measures were collected from native Arabic speakers during a visual lexical decision task. This task included LAHF, LALF and SAHF words. ERPs were analysed in terms of components' amplitude and time latency. We expected to observe frequency effects both in behavioural (RTs and accuracy) and in brain electric measures. The prediction was that the frequency effects will affect the lexical decision task in a way that LAHF will induce shorter RTs than LALF and SAHF. Also, we expected higher accuracy for LA than SA because LA behaves like an L1 in the written modality (Bentin & Ibrahim, 1996; Ibrahim, 2009; Tarabya et al., 2021). Also, it was assumed that no major difference would be found in electric brain responses between LALF and SAHF words, while both will be differing from LAHF due to difference in the visual familiarity of the words.

At the behavioural level, we found a language variety effect, which was reflected in shorter RTs for LAHF than SAHF and word frequency effect, which appeared in shorter RTs for LAHF than LALF words. The language variety effect (SA vs. LA) mimicked the so-called second-to-first language effect, where RTs are generally slower for L2 than L1 words (de León Rodríguez et al., 2022; Khateb et al., 2016; von Studnitz & Green, 2002). These findings confirm previous observations (Khateb & Ibrahim, 2021; Nevat et al., 2014; Tarabya et al., 2021) showing that in the written modality, LA behaved as the dominant L1-like variety, being easier to recognize and read, while SA words seemed to behave as the less dominant L2-like ones. The difference found here in RTs is in line with our prediction, which proposed that LA words will be processed faster than SA, with difference in accuracy between LA and SA failing to reach significance. The frequency effect manifested in LAHF inducing faster RTs and higher accuracy than LALF is consistent with other findings from studies in various languages (Bronk et al., 2013; Coltheart, 2005; Forster & Chambers, 1973; Koriat, 1985; Kuo et al., 2003; Maïonchi-Pino et al., 2010; Proverbio et al., 2008; Schröter & Schroeder, 2018). Of note is the fact that the relatively low accuracy in LALF words cannot be explained by trade-offs effect between RT and accuracy because these measures did not correlate here (result not reported). Also, ad-hoc item-analysis of the words in LALF stimulus list revealed that only 20% of the words (i.e., 12 items) yielded accuracy <50%, but not to the extent that some words were completely not recognized by the participants. Altogether, the results suggest that processing low-frequency LA words was very demanding for the participants.

At the electrophysiological level, we expected differences at the N170 component due to the frequency effect, as suggested in various previous studies (Davis et al., 2019; Faísca et al., 2019; Hauk & Pulvermüller, 2004; Mahé et al., 2012; Maurer et al., 2005). Unexpectedly, this study showed first a modulation of the amplitude of the P100/P1 component by language condition/word frequency. Indeed, we found that the P1 amplitude was the lowest amplitude for LAHF words and the highest amplitude for LALF words, with SAHF words between the two formers. This observation indicated a relation between the ease with which the words were recognized and this component's amplitude. Previous studies suggested a relation between the P1 amplitude and word length: The longer the word, the larger the positivity (Davis et al., 2019; Hauk et al., 2006). Hauk et al. (2006) suggested that the P1 amplitude was also related to word frequency, lower amplitude for words with high frequency, as shown in our study here. Moreover, these authors argued that the P1 was related also to lexicality, with words and pseudowords with atypical orthography eliciting stronger brain activation than words with typical orthography. Although ERPs to pseudowords were not analysed in the present study, our results are in accordance with Hauk et al.' s interpretation because LAHF words (which are more typical to the readers and more frequent) had lower P1 amplitudes than LALF and SAHF, when there was no difference between LALF and SAHF words, which matches atypical representation in the written language. Furthermore, our results showed an earlier peak for LAHF than for LALF and SAHF. This observation indicates that a discrimination between low and high frequency of a word, begin already at a very early stages of word processing, with SAHF being, to a great extent, processed more as LALF ones (Bentin & Ibrahim, 1996), both in terms of amplitude and peak latency. A relation between the P100 amplitude and dyslexia had also been suggested. For instance, in the study investigating subgroups of dyslexia (Dujardin et al., 2011), it was reported that the P100 was of lower amplitude in dyslexics displaying poorer accuracy for infrequent and pseudowords than in dyslexics showing slower responses for pseudowords. A difference in the amplitude of the P100 was also reported between healthy adult controls and participants from various clinical (schizophrenia, bipolar disorder and major depression) samples (Spironelli et al., 2019).

Our results showed (on O1/O2 electrodes) a higher N170 amplitude for LAHF than for LALF and for SAHF (with no difference between the latter two), confirming thus the first and major prediction of study. Again, here as in the case of the P1, LALF and SAHF words behaved similarly and were both different from LAHF words. This observation is in line with other studies, which suggested that the N170 is related to the orthographic features of the stimuli rather than to their phonological features (Braun et al., 2009; Coch & Meade, 2016; Simon et al., 2004). Simon et al. (2004) suggested larger N170 waves for orthographic stimuli (relative to non-orthographic), and this amplitude was sensitive to word repetition. Other authors have proposed that the N170 was modulated by selective attention (Hillyard & Anllo-

Vento, 1998), such that the N170 appeared larger (i.e., enhanced N170) when the stimulus appeared in an expected location. Here, we showed that LAHF words showed a higher amplitude and this is probably because of the reader's higher familiarity with such stimuli. LAHF words, which are more frequent and thus easier to detect and recognize as orthographic patterns than the LALF and the SAHF ones, enhanced the N170. Our findings are consistent with results from Braun (see Braun et al., 2009) who suggested that the N170 reflects a prelexical orthographic processing which, as a holistic process, allows frequent words to be treated as a global visual pattern (Simon et al., 2007). Also, the modulation of the N170 seen here might also be interpreted in relation to selective attention processes because LAHF might be more predicted in the written modality than the SAHF and the LALF. A modulation of the amplitude of the N170 during word recognition was already observed in a previous study on Arabic where orthographic connectivity was manipulated (Taha et al., 2013). Connected (and more familiar) word patterns not only were processed more efficiently in terms of RTs and accuracy, they also showed a higher N170 amplitude and an earlier peak latency. These previous findings suggested, like here, that words which are easier to process induced a larger N170. These results and interpretation contradict others in previous studies, which found the opposite modulation of the N170 (Sereno et al., 1998). In this latter combined ERP and eye tracking study, the authors compared words (high and low frequency), pseudo words and non-words among 40 students in the ERP experiment and found lower N170 amplitudes for high-frequency words than low-frequency words. In another investigation on Arabic using real words and pseudo-homophones (words sharing the same phonology and semantics but with incorrect orthography), the N170 showed a higher amplitude for the orthographically incorrect words (Taha & Khateb, 2013). In this latter study, the differences were attributed to a more-in-depth orthographic analysis of the stimuli (Taha & Khateb, 2013). The results from these two previous studies suggested that the enhancement of the N170 amplitude cannot be interpreted unequivocally because in one case increased amplitude was observed in relation with an easier processing (Taha et al., 2013) while in the other increased amplitude was observed in relation to a more careful analysis of the orthographic pattern (Taha & Khateb, 2013). Finally, it should be noted that our analysis of the N170 showed a left hemisphere lateralization (larger over left than right hemisphere) for the N170. This result is in accordance with other studies using visual word processing tasks (Sereno et al., 1998; Xue et al., 2019). In bilinguals, Grossi et al. (2010) showed the same left hemisphere lateralization of N170 in semantic categorization task performed by early and late learners of English–Welsh bilinguals.

As for the time latency of the N170, we found no difference in latency between word conditions. This result is in accordance with previous results that showed a modulation of the N170 amplitude and not the latency by word frequency (Hauk & Pulvermüller, 2004). However, this finding contrasts with a previous observation on Arabic showing a modulation of both the amplitude and latency of the N170 with the ease with which words were processed as a function of the letters' connectedness (Taha et al., 2013). In this later study, the latency effect was assessed on three left and right electrodes. The absence of such effect in our present findings might simply be due to the sample size or to the analysis method used here which looked at the latency in the averaged region of interest.

The grand-mean ERPs globally indicated a quite similar pattern of differences between language conditions across the three successive *P2-N2-P3* components. The analysis conducted separately on the amplitude of each of these components revealed a significant language condition effect only for the P3. The amplitude of the P3 was the highest in SAHF, followed by LALF and then by LAHF word condition. Differences during similar time periods had been shown to differentiate between words and non-words (Khateb et al., 2002). Although the analysis conducted here did not compare words with non-words, a parallel can still be drawn between such previous findings and the differences found here between language conditions: with low visual familiarity words (LALF and SAHF) vs high visual familiarity (LAHF). In support of this interpretation, we also found an earlier P2 peak for LAHF than for LALF and SAHF. Also, this observation lines up with earlier results, which suggested a longer duration for this processing step occurring during this time period for L2 than for L1 words (Khateb et al., 2016). The differences between L2 and L1 in this latter study were interpreted in terms of word frequency. In fact, L2 words even in highly proficient bilinguals are subjectively of a lower frequency than L1 ones because bilinguals acquire their L2 later in life and are generally less exposed to L2 than to L1 words (Duyck et al., 2008; Gollan et al., 2008). In the present study, the fact that LALF and SAHF behaved similarly (and both differed from LAHF, which behaved as the dominant language condition) supports the previous claim that differences between L2 and L1 during word processing might simply be interpreted in terms of the visual familiarity/frequency of the words (Khateb et al., 2016; Nevat et al., 2014). In line with previous literature, the language condition effect during the P3 strongly supports this view. In fact, the ERP literature had repeatedly reported that the P300 component is best induced by rare and/or infrequent stimuli or by novelty (Comerchero & Polich, 1999; Sutton et al., 1965). The modulation of the response found here is thus compatible with the visual frequency of exposure to written SA words. Together with the history of acquisition and the patterns of use of SA, it is generally assumed that in the visual modality written transliteration of SA words has no accepted upon standard form in Arabic; hence, encountering a written SA words looks more like processing a novel/infrequent stimulus than LA words both of low and high frequency. Practically, their recognition will be realized through the slower non-lexical phonological route that is mediated by a process of grapheme to phoneme conversion (Coltheart, 2005; Taouk & Coltheart, 2004). Finally, the findings reported here showed that the P600 component exhibited its highest amplitude following LAHF words, and with no difference between LALF and SAHF ones, which both behaved similarly. Actually, this pattern of response difference (i.e., LAHF differing from the other two language conditions) was observed almost during the whole stream of word processing, as attested by the differences during the successive early components starting with the P1 and ending with the P3 (although with slight differences between LALF and SAHF during the P1 and P3). This component had been related to late memory monitoring before decision making (Kaan & Swaab, 2003; Savill & Thierry, 2011). Such component is usually described at centro-parietal electrodes (Osterhout & Holcomb, 1992). In our case, the higher amplitude reflected again the ease with which this monitoring stage of the lexical analysis of words with different frequencies was conducted to enable correct recognition of the real words. In support of this interpretation, the analysis also showed that this component occurred earlier in LAHF than in the other two conditions. Similar results have previously been obtained both in terms of P600 amplitude and latency in Arabic where this component differentiated real words from pseudo homophones during a speeded orthographic decision task (Taha & Khateb, 2013). Assuming that this component is related to the decision-making processing step (which in a lexical decision require differentiating pseudowords from words but with different frequencies), one would be tempted to conclude that the higher amplitude of this late response (and the earlier latency) to high-frequency words would explain the faster RTs to LAHF words in comparison with LALF and SA-LF ones. The analysis presented here for the successive components reflecting the different steps of information processing suggest that word recognition involves very dynamic sub-processes all over the stream. From the visual P100 up to the P600 component where the lexical items are retrieved from long-term memory and the decision process takes place, an interplay seems to exist between the different processing steps that finally predicts the participants' response speed. Such a dynamic interplay manifested here by the fact that the earlier the peak latency was for the P1 and P2 components, the earlier late responses occurred (i.e., P600) and the faster were participants' responses.

To summarize, from the results of this study, one might say that during word recognition in Arabic, the early stages of word processing differentiate LAHF from LALF and SAHF words. The difference between language conditions during the early components both in terms of response amplitude and latency re-appeared also later and determined to a great extent the time course of the P600 component and the participants' response times. These results highlight the importance of word frequency in word processing in Arabic language as in other languages. The results presented here support a previous proposition that differences during the early stages of processing written L1 and L2 words might mainly reflect difference in words' frequency (Khateb et al., 2016).

The results of this study suffer from the fact that the study sample size is relatively small for the conducted analysis. Because of this limitation, statistical effects, although significant and resisting to multiple comparison corrections, were generally small. This observation is particularly true for the language effect during N170 component. Hence, future studies should involve a larger sample size in order to draw stronger conclusions. Still, this study is to our knowledge the first of its kind to address the brain correlates of frequency effects in Arabic. This study showed that the LAHF behaved as the dominant language variety in the written modality (because it is the one with highest written frequency), and LALF and SAHF together behaved similarly (both with a lower written frequency). We thus conclude that the dichotomy L1 versus L2 to represent respectively SA and LA in Arabic, which was initially proposed on the basis of the history of acquisition of the two Arabic varieties and on empirical data obtained in the auditory modality, does not really hold in the visual written modality (see also Khateb & Ibrahim, 2021). These findings are in line with a previous fMRI study (Nevat et al., 2014) using LA and SA written words. Also, they align with previous findings reported in another fMRI picture naming study using SA and LA names for the same pictures and where no difference was found between the two varieties during production (Abou-Ghazaleh et al., 2018). Indeed, we showed here that SA written words, even when of high frequency of use in the oral/auditory modality, function as low-frequency words in the written modality. As for the question of how SA words compare with low frequent LA words, this study is the first to provide an answer and strengthen the view that response to the question of dominance in Arabic diglossia is neither direct nor unequivocal. We have previously proposed that the status of each of the varieties of Arabic depends on various parameters that among other include the nature of the task and its demands, the linguistic register involved, the individual's proficiency in each of the varieties, the modality of stimuli presentation (auditory/oral vs. visual) and the type of processing (reception vs. production, etc.) (Khateb & Ibrahim, 2021). Accordingly, future research should continue investigating the diglossia question in the different modalities not only at the level of single words but also at the sentence level, during reading, listening and discourse production. A better understanding of the interactions between the two Arabic varieties will be important for deciphering not only the cognitive underpinning of the diglossia phenomenon but also the human language experience more generally.

## ACKNOWLEDGEMENTS

This work was supported by the Israeli Science Foundation (grant 2695/19) and by the Edmond J. Safra Brain Research Center for the Study of Learning Disabilities. We thank all the participants for their participation in this study.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

**Samer Andria** participated in the study design, data collection and analysis and wrote the first draft of the paper and participated in the final revision. **Bahaa Madi-Tarabya** participated in data collection and analysis and participated in the writing and revision of the manuscript. **Asaid Khateb** supervised the whole process from the research design to data analysis, revised the manuscript through the start of the process of writing to the final revision of the paper.

## PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15781.

## DATA AVAILABILITY STATEMENT

Grand averaged EEG and individual ERP files, together with the data for behavioural analysis can be found on OSF (https://osf.io/6kv2r/?view_only=a3233aa67bf14b508a0cb1907c147ddf). Due to the General Data Protection Regulation (GDPR), raw EEG data and raw behavioural files (E-prime) are only available from the corresponding author upon request.

## ORCID

*Asaid Khateb* https://orcid.org/0000-0003-4639-1564

## 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. https://doi.org/10.1016/j.neuroscience.2018.09.045

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

Bentin, S., Allison, T., Puce, A., Perez, A., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565. https://doi.org/10.1162/jocn.1996.8.6.551

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*(2), 309–323. http://europepmc.org/abstract/MED/8901339, https://doi.org/10.1037/0278-7393.22.2.309

Bentin, S., Mouchetant-Rostaing, Y., Giard, M., Echallier, J., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260. https://doi.org/10.1162/089892999563373

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 201–219). Lawrence Erlbaum Associates, Inc.

Blasi, V., Young, A. C., Tansy, A. P., Petersen, S. E., Snyder, A. Z., & Corbetta, M. (2002). Word retrieval learning modulates right frontal cortex in patients with left frontal damage. *Neuron*, *36*(1), 159–170. https://doi.org/10.1016/S0896-6273(02)00936-4

Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, *3*(1), 1–13. https://doi.org/10.1186/1744-9081-3-7

Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, *30*(7), 1977–1989. https://doi.org/10.1002/hbm.20643

Bronk, M., Zwitserlood, P., & Bölte, J. (2013). Manipulations of word frequency reveal differences in the processing of morphologically complex and simple words in German. *Frontiers in Psychology*, *4*, 546. https://doi.org/10.3389/fpsyg.2013.00546

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect. *Experimental Psychology*, *58*, 412–424. https://doi.org/10.1027/1618-3169/a000123

Brysbaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, *20*(3), 530–548. https://doi.org/10.1017/S1366728916000353

Coch, D., & Meade, G. (2016). N1 and P2 to words and wordlike stimuli in late elementary school children and adults. *Psychophysiology*, *53*(2), 115–128. https://doi.org/10.1111/psyp.12567

Coltheart, M. (2005). *Modeling reading: The dual-route approach* (Vol. 6). Blackwells Publishing.

Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, *4*(1), 5–17. https://doi.org/10.1080/13603110600574322

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589

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*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24–30. https://doi.org/10.1016/S0168-5597(98)00033-1

Davis, C. P., Libben, G., & Segalowitz, S. J. (2019). Compounding matters: Event-related potential evidence for early semantic access to compound words. *Cognition*, *184*, 44–52. https://doi.org/10.1016/j.cognition.2018.12.006

de León Rodríguez, D., Mouthon, M., Annoni, J.-M., & Khateb, A. (2022). Current exposure to a second language modulates bilingual visual word recognition: An EEG study. *Neuropsychologia*, *164*, 108109. https://doi.org/10.1016/j.neuropsychologia.2021.108109

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, *66*(5), 843–863. https://doi.org/10.1080/17470218.2012.720994

Dujardin, T., Etienne, Y., Contentin, C., Bernard, C., Largy, P., Mellier, D., Lalonde, R., & Rebaï, M. (2011). Behavioral performances in participants with phonological dyslexia and different patterns on the N170 component. *Brain and Cognition*, *75*(2), 91–100. https://doi.org/10.1016/j.bandc.2010.10.006

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2014). An ERP investigation of the co-development of hemispheric lateralization of face and word recognition. *Neuropsychologia*, *61*, 315–323. https://doi.org/10.1016/j.neuropsychologia.2014.05.006

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

Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied PsychoLinguistics*, *21*(4), 451–471. https://doi.org/10.1017/S0142716400004021

Faísca, L., Reis, A., & Araújo, S. (2019). Early brain sensitivity to word frequency and lexicality during reading aloud and implicit reading. *Frontiers in Psychology*, *10*, 830. https://doi.org/10.3389/fpsyg.2019.00830

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

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627–635. https://doi.org/10.1016/S0022-5371(73)80042-8

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1219–1248.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, *45*(3), 468–492. http://www.sciencedirect.com/science/article/pii/S0749596X00927840, https://doi.org/10.1006/jmla.2000.2784

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281. https://doi.org/10.1037/0096-3445.113.2.256

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*(5), 1122–1139. https://doi.org/10.1037/0278-7393.23.5.1122

Gollan, T. H., Montoya, R.-I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gordon, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior*, *22*(1), 24–44. https://doi.org/10.1016/S0022-5371(83)80004-8

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Grossi, G., Savill, N., Thomas, E., & Thierry, G. (2010). Posterior N1 asymmetry to English and Welsh words in Early and Late English–Welsh bilinguals. *Biological Psychology*, *85*(1), 124–133.

Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2013). The time course of reading processes in children with and without dyslexia: An ERP study. *Frontiers in Human Neuroscience*, *7*, 570. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3791381/pdf/fnhum-07-00570.pdf. https://doi.org/10.3389/fnhum.2013.00570

Hauk, O., Davis, M. H., & Pulvermüller, F. (2008). Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. *NeuroImage*, *42*(3), 1185–1195. https://doi.org/10.1016/j.neuroimage.2008.05.054

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermüller, F., & Rogers, T. (2006). ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience*, *18*(5), 818–832. https://doi.org/10.1162/jocn.2006.18.5.818

Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, *115*(5), 1090–1103. https://doi.org/10.1016/j.clinph.2003.12.020

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of* *the National Academy of Sciences*, *95*(3), 781–787. https://doi.org/10.1073/pnas.95.3.781

Ibáñez, A., Petroni, A., Urquina, H., Torrente, F., Torralva, T., Hurtado, E., Guex, R., Blenkmann, A., Beltrachini, L., & Muravchik, C. (2011). Cortical deficits of emotional face processing in adults with ADHD: Its relation to social cognition and executive function. *Social Neuroscience*, *6*(5–6), 464–481. https://doi.org/10.1080/17470919.2011.620769

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. https://doi.org/10.2147/PRBM.S5138

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

Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2007). Metalinguistic awareness and reading performance: A cross language comparison. *Journal of Psycholinguistic Research*, *36*(4), 297–317. https://doi.org/10.1007/s10936-006-9046-3

Ji, Y., Liu, J., Zhu, X.-Q., Zhao, J., Wang, J., Du, Y.-C., & Bi, H.-Y. (2019). Atypical N170 lateralization of face and word recognition in Chinese children with autism spectrum disorder. *Journal of Neurolinguistics*, *52*, 100858. https://doi.org/10.1016/j.jneuroling.2019.100858

Kaan, E., & Swaab, T. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Cognitive Neuroscience, Journal of*, *15*(1), 98–110. https://doi.org/10.1162/089892903321107855

Keatley, C. W., Spinks, J. A., & de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, *22*(1), 70–84. https://doi.org/10.3758/BF03202763

Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cerebral Cortex*, *11*(3), 223–237. https://doi.org/10.1093/cercor/11.3.223

Khateb, A., & Ibrahim, R. (2021). About the neural basis of Arabic diglossia: Behavioral and event-related potential analysis of word processing in spoken and literary Arabic. In E. Saiegh-Haddad, L. Laks, & C. McBride (Eds.), *Handbook of literacy in diglossia and dialectal contexts: Psycholinguistic and educational perspectives*. Springer.

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*(3), 311–320. https://doi.org/10.1007/s10548-010-0149-7

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*(3), 197–213. https://doi.org/10.1023/A:1014502925003

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*(2), 311–330.

Koriat, A. (1985). Lexical access for low-and high-frequency words in Hebrew. *Memory & Cognition*, *13*(1), 37–44. https://doi.org/10.3758/BF03198441

Korinth, S. P., & Breznitz, Z. (2014). Fast and slow readers of the Hebrew language show divergence in brain response ∼200 ms post stimulus: An ERP study. *PLoS ONE*, *9*(7), e103139. https://doi.org/10.1371/journal.pone.0103139

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.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage*, *21*(3), 946–953. https://doi.org/10.1016/j.neuroimage.2003.10.021

Kuo, W.-J., Yeh, T.-C., Lee, C.-Y., Wu, Y.-T., Chou, C.-C., Ho, L.-T., Hung, D. L., Tzeng, O. J., & Hsieh, J.-C. (2003). Frequency effects of Chinese character processing in the brain: An event-related fMRI study. *NeuroImage*, *18*(3), 720–730. https://doi.org/10.1016/S1053-8119(03)00015-6

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of chekerboard-evoked multichannels potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609–621. https://doi.org/10.1016/0013-4694(80)90419-8

Mahé, G., Bonnefond, A., Gavens, N., Dufour, A., & Doignon-Camus, N. (2012). Impaired visual expertise for print in French adults with dyslexia as shown by N170 tuning. *Neuropsychologia*, *50*(14), 3200–3206. https://doi.org/10.1016/j.neuropsychologia.2012.10.013

Maïonchi-Pino, N., Magnan, A., & Écalle, J. (2010). Syllable frequency effects in visual word recognition: Developmental approach in French children. *Journal of Applied Developmental Psychology*, *31*(1), 70–82. https://doi.org/10.1016/j.appdev.2009.08.003

Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, *1*. https://doi.org/10.1186/1744-9081-1-13

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*(1), 43–71. https://doi.org/10.1037/0096-3445.118.1.43

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research*, *22*(2), 205–220. https://doi.org/10.1016/j.cogbrainres.2004.08.010

Morton, J. (1970). A functional model for memory. In *Models of human memory* (pp. 203–254). Academic Press.

Münte, T. F., Heinze, H.-J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia*, *36*(3), 217–226. https://doi.org/10.1016/S0028-3932(97)00119-X

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, *111*(3), 721–756. https://doi.org/10.1037/0033-295X.111.3.721

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. *European Journal of Neuroscience*, *40*(9), 3387–3395. https://doi.org/10.1111/ejn.12673

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics*, *67*(2), 301–314. https://doi.org/10.3758/BF03206493

Proverbio, A., Zani, A., & Adorni, R. (2008). The left fusiform area is affected by written frequency of words. *Neuropsychologia*, *46*(9), 2292–2299. http://f1000.com/1108855. https://doi.org/10.1016/j.neuropsychologia.2008.03.024

Proverbio, A. M., Vecchi, L., & Zani, A. (2004). From orthography to phonetics: ERP measures of grapheme-to-phoneme conversion mechanisms in reading. *Journal of Cognitive Neuroscience*, *16*(2), 301–317. https://doi.org/10.1162/089892904322984580

Proverbio, A. M., Wiedemann, F., Adorni, R., Rossi, V., Del Zotto, M., & Zani, A. (2007). Dissociating object familiarity from linguistic properties in mirror word reading. *Behavioral and Brain Functions*, *3*(1), 1–22. https://doi.org/10.1186/1744-9081-3-43

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 487–494. https://doi.org/10.1016/S0022-5371(70)80091-3

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, *18*, 367–379. https://doi.org/10.3758/BF03197126

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied PsychoLinguistics*, *24*(3), 431–451. https://doi.org/10.1017/S0142716403000225

Saiegh-Haddad, E. (2012). Literacy reflexes of Arabic diglossia. In *Current issues in bilingualism* (pp. 43–55). Springer. https://doi.org/10.1007/978-94-007-2327-6_3

Saiegh-Haddad, E. (2018). MAWRID: A model of Arabic word reading in development. *Journal of Learning Disabilities*, *51*(5), 454–462. https://doi.org/10.1177/0022219417720460

Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. In *Handbook of Arabic literacy* (pp. 3–28). Springer. https://doi.org/10.1007/978-94-017-8545-7_1

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. https://doi.org/10.1080/10888438.2016.1180526

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*, 1–26. https://doi.org/10.1007/s11145-020-10019-3

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

Savill, N. J., & Thierry, G. (2011). Reading for sound with dyslexia: Evidence for early orthographic and late phonological integration deficits. *Brain Research*, *1385*, 192–205. https://doi.org/10.1016/j.brainres.2011.02.012

Schröter, P., & Schroeder, S. (2018). Differences in visual word recognition between L1 and L2 speakers: The impact of length, frequency, and orthographic neighborhood size in German children. *Studies in Second Language Acquisition*, *40*(2), 319–339. https://doi.org/10.1017/S0272263117000201

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *Neuroreport*, *9*(10), 2195–2200. https://doi.org/10.1097/00001756-199807130-00009

Simon, G., Bernard, C., Largy, P., Lalonde, R., & Rebai, M. (2004). Chronometry of visual word recognition during passive and lexical decision tasks: An ERP investigation. *International Journal of Neuroscience*, *114*(11), 1401–1432. https://doi.org/10.1080/00207450490476057

Simon, G., Petit, L., Bernard, C., & Rebaï, M. (2007). N170 ERPs could represent a logographic processing strategy in visual word recognition. *Behavioral and Brain Functions*, *3*(1), 1–11. https://doi.org/10.1186/1744-9081-3-21

Spironelli, C., & Angrilli, A. (2009). Developmental aspects of automatic word processing: Language lateralization of early ERP components in children, young adults and middle-aged subjects. *Biological Psychology*, *80*(1), 35–45. https://doi.org/10.1016/j.biopsycho.2008.01.012

Spironelli, C., Carbone, E., & Borella, E. (2020). Electrophysiological correlates of the categorization working memory span task in older adults. *Behavioural Brain Research*, *393*, 112809. https://doi.org/10.1016/j.bbr.2020.112809

Spironelli, C., Penolazzi, B., Vio, C., & Angrilli, A. (2010). Cortical reorganization in dyslexic children after phonological training: Evidence from early evoked potentials. *Brain*, *133*(11), 3385–3395. https://doi.org/10.1093/brain/awq199

Spironelli, C., Romeo, Z., Maffei, A., & Angrilli, A. (2019). Comparison of automatic visual attention in schizophrenia, bipolar disorder, and major depression: Evidence from P1 event-related component. *Psychiatry and Clinical Neurosciences*, *73*(6), 331–339. https://doi.org/10.1111/pcn.12840

Stites, M. C., Federmeier, K. D., & Christianson, K. (2016). Do morphemes matter when reading compound words with transposed letters? Evidence from eye-tracking and event-related potentials. *Language, Cognition and Neuroscience*, *31*(10), 1299–1319. https://doi.org/10.1080/23273798.2016.1212082

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928. https://doi.org/10.1093/cercor/bhp153

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188. https://doi.org/10.1126/science.150.3700.1187

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*(2), 292–302. https://doi.org/10.1007/s10548-012-0241-2

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

Taouk, M., & Coltheart, M. (2004). The cognitive processes involved in learning to read in Arabic. *Reading and Writing*, *17*(1), 27–57. https://doi.org/10.1023/B:READ.0000013831.91795.ec

Tarabya, B. M., Khateb, A., & Andria, S. (2021). Processing printed words in literary Arabic and spoken Arabic: An fNIRS study. *Open Journal of Modern Linguistics*, *11*(3), 461–483. https://doi.org/10.4236/ojml.2021.113035

Van de Meerendonk, N., Indefrey, P., Chwilla, D. J., & Kolk, H. H. (2011). Monitoring in language perception: Electrophysiological and hemodynamic responses to spelling violations. *NeuroImage*, *54*(3), 2350–2363. https://doi.org/10.1016/j.neuroimage.2010.10.022

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequencyinevent-related brain potentials. *Memory & Cognition*, *18*(4), 380–393. https://doi.org/10.3758/BF03197127

von Studnitz, R. E., & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, *5*(01), 1–23. https://doi.org/10.1017/S1366728902000111

Xue, L., Maurer, U., Weng, X., & Zhao, J. (2019). Familiarity with visual forms contributes to a left-lateralized and increased N170 response for Chinese characters. *Neuropsychologia*, *134*, 107194. https://doi.org/10.1016/j.neuropsychologia.2019.107194

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 274–296.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English lexicon project. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 53–79. https://doi.org/10.1037/a0024177

Young, M. P., & Rugg, M. D. (1992). Word frequency and multiple repetition as determinants of the modulation of event-related potentials in a semantic classification task. *Psychophysiology*, *29*(6), 664–676. https://doi.org/10.1111/j.1469-8986.1992.tb02044.x

**How to cite this article:** Andria, S., Madi-Tarabya, B., & Khateb, A. (2022). Behavioural and electrophysiological analyses of written word processing in spoken and literary Arabic: New insights into the diglossia question. *European Journal of Neuroscience*, *56*(6), 4819–4836. https://doi.org/10.1111/ejn.15781