informa
healthcare

Original Article

International Journal of Audiology 2010; 00: 1–6

Hanin Karawani
Karen Banai

Department of Communication Sciences and Disorders, Faculty of Social Welfare and Health Sciences, University of Haifa, Israel

Speech-evoked brainstem responses in Arabic and Hebrew speakers

Keywords

Speech-sound encoding
auditory brainstem response (ABR)
speech ABR
entral auditory processing.

Abbreviations

ABR: Auditory brainstem response
BioMARK: Biological marker of auditory processing
VOT: Voice onset time

Abstract

Based on studies in English speakers, it has been proposed that speech-evoked brainstem responses (ABRs) may be used clinically to assess central auditory function. Whether the same procedure can be used across speakers of different languages remains unclear, because recent findings suggest that language experience affects even subcortical processing of speech. The goal of the present study was to characterize brainstem responses to the syllable /da/ in Arabic and Hebrew speakers using the US developed BioMARK procedure. To that end, ABRs evoked by clicks and the syllable /da/ were collected from 37 normal-hearing students from the University of Haifa. Neither the transient nor the sustained components of the brainstem response differed significantly between Arabic and Hebrew speakers. Across the two groups, timing of the major components of the speech-evoked response as well as the correlations between the speech- and click indices were well within the US norms. Therefore, brainstem processing of the syllable /da/ does not differ between speakers of English and speakers of Semitic languages such as Arabic and Hebrew.

Sumario

Con base en los estudios realizados en hablantes del inglés se ha propuesto que las respuestas del tallo cerebral evocadas por lenguaje (ABR) pueden utilizarse clínicamente para evaluar la función auditiva central. Aun no queda claro si el mismo procedimiento puede utilizarse (transversalmente) entre los hablantes de diferentes lenguas ya que algunos hallazgos recientes sugieren que la experiencia lingüística afecta incluso el procesamiento subcortical del lenguaje. El objetivo del presente estudio fue caracterizar las respuestas del tallo cerebral ante la sílaba /da/ en hablantes del árabe y el hebreo utilizando el procedimiento BioMARK desarrollado en EUA. Para tal fin, se colectaron respuestas de ABR evocadas por clicks y por la sílaba /da/ de 37 estudiantes normoyentes de la Universidad de Haifa. Ni los componentes basales ni los transitorios de las respuestas evocadas difirieron significativamente entre los hablantes del árabe y del hebreo. A través de los dos grupos la temporalidad de los componentes mayores de las respuestas evocadas por lenguaje así como la correlación entre los índices con lenguaje y con clicks estuvieron dentro de las normas de los EUA. Por lo tanto, el procesamiento del tallo cerebral de la sílaba /da/ no difiere entre los hablantes del inglés y los hablantes de las lenguas semíticas como el árabe y el hebreo.

Even though the majority of individuals referred for audiological evaluation complain about difficulties in understanding speech (e.g. Tremblay et al, 2003), traditionally, audiological assessment made extensive use of non-speech materials. In recent years, recognition of this discrepancy has lead to the realization that using speech based testing procedures is desirable (Cacace & McFarland, 2005; Johnson et al, 2005; Kraus & Nicol, 2005). One such procedure, which is the focus of this report, is the BioMARK (Biological Marker of Auditory Processing, Natus Medical Inc., Mundelein, USA) procedure that was developed to supplement the commonly used ABR procedure with the evaluation of brainstem encoding of the speech syllable /da/ (Johnson et al, 2005).

In response to acoustic stimulation, the ascending auditory pathway responds with a characteristic series of neural discharges that are reflected in electric potential fluctuations that can be measured on the scalp and provide information about the synchronous firing of structures along the pathway, which include the auditory nerve, cochlear nuclei, superior olivary nuclei, lateral lemnisci, and inferior colliculi (Møller & Jannetta, 1985). This orderly pattern, known as the auditory brainstem response (ABR), has been used as a non-invasive measure of far-field representation of stimulus-locked, synchronous electrical events. In the clinic, the ABR is most often elicited with click stimuli and is used to evaluate hearing and the integrity of the ascending auditory pathway (Hall, 1992; Jacobsen, 1985; Josey, 1985; Musiek, 1991; Starr & Don, 1988). In the present study, the click ABR was used to ensure normal waveform morphology in participants.

In addition to clicks, ABRs can be evoked using speech sounds (Chandrasekaran & Kraus, 2009; Galbraith et al, 1995; Russo et al, 2004). The speech sound /da/ is present in multiple languages, making it a nearly universal option (Maddieson, 1984). Over the last decade, brainstem encoding of the speech sound /da/ has been investigated in English speakers. Much like click evoked ABRs, the ABR to the syllable /da/ also has a characteristic morphology that varies little among individuals with normal hearing and typical development. Furthermore, it appears that this morphology reflects some of the acoustic-phonetic characteristics of speech with remarkable precision (Chandrasekaran & Kraus, 2009; Johnson et al, 2005, 2008, 2005; Russo et al, 2004). Therefore, it has been suggested that the speech-ABR evoked with the syllable /da/ can be used clinically for the assessment of central auditory function. Two lines of evidence provide tentative support for the clinical use of /da/. First, several studies have shown that the encoding of /da/ is compromised in individuals with language, reading, and learning problems and that the degree of deficit is correlated with the severity of reading problems as well as with cortical processing of speech (Abrams et al, 2006; Banai et al, 2005, 2009; King et al, 2002; Wible et al, 2005). This abnormal processing of speech is in contrast to the normal encoding of clicks in these populations (e.g. Song et al, 2006). Therefore, the speech-ABR can potentially provide additional information about central auditory function in these groups. In addition, there is evidence that language training is more beneficial to children with language learning problems with deficient subcortical encoding of speech than to children with language problems but no subcortical deficit, and that in this population

ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.3109/14992027.2010.495083

Received:
January 14, 2010
Accepted:
May 15, 2010

Karen Banai
Department of Communication Sciences and Disorders, University of Haifa, Mt. Carmel, Haifa 31905, Israel.
E-mail: kbanai@research.haifa.ac.il

training can enhance subcortical encoding (King et al, 2002; Russo et al, 2005). These plastic changes suggest that the waveform complex may have clinical applications as an objective physiologic correlate of speech-sound representation associated with speech-sound training.

The BioMARK (biological marker of auditory processing) procedure has been developed based on the above mentioned studies and introduced in the United States (US) as a tool for the objective assessment of speech sound processing in the brainstem in clinical settings such as the diagnosis of auditory processing disorder and learning problems.

Whether this procedure can be used in non English speaking populations without alteration remains unclear, because even though the constituents of the syllable /da/ are characteristic of multiple languages (Maddieson, 1984), recent studies have shown that even subcortical encoding of speech is subject to the long term effects of experience with language (Krishnan et al, 2005) and music (Wong et al, 2007).

The goal of the present study was, therefore, to characterize brainstem encoding of speech sounds among native speakers of Hebrew and Arabic using the BioMARK procedure and compare the responses to the US normative data. This question is of interest not only from a clinical perspective, but also theoretically because the effects of language experience on subcortical auditory function have been demonstrated mostly by comparing English speakers to speakers of a tonal language, e.g. Mandarin (Krishnan et al, 2005, 2009), and it is unclear whether they generalize to non-tonal languages or to acoustic parameters other than pitch, such as voice onset time (VOT), which varies in a language specific pattern. When stop consonants such as /da/ are produced by English speakers, the vibration of the vocal folds starts at a short lag after the release of the consonant by the articulators (VOT = 5 ms in the case of the stimuli used by BioMARK). In contrast, in Arabic (Khateeb, 2000) and Hebrew (Raphael et al, 1995) voicing starts earlier, prior to the release of the consonants (for example, in naturally produced Hebrew, VOT = −135 to −40 ms, Raphael et al, 1995). If long term experience with the specific acoustic rendition of the syllable results in changes to brainstem encoding of stop consonants, we would expect that ABRs evoked by the English syllable /da/ would differ between native speakers of Arabic and Hebrew on the one hand and native English speakers on the other. Alternatively, if long term experience results in changes that reflect perceptual, rather than acoustic experience, as has been suggested by animal work (Polley et al, 2006), we would expect that speech-ABRs of Arabic and Hebrew speakers would not differ from those of English speakers because perceptually, the /da/ syllable used is linguistically meaningful in all three languages. Either way, no differences are expected between native speakers of Arabic and Hebrew.

Methods

Participants

A total of 37 volunteer students from Haifa University, native Hebrew and Arabic speakers, (27 females), between the ages of 18–28 (M = 23.5, SD = 1.67), participated in this study. All subjects reported normal audiological and neurological histories and were never diagnosed with any form of learning problem. They all had normal hearing thresholds at or below 20 dB hearing level for octaves from 250 to 8000 Hz (ANSI, 1989). All subjects gave their consent to participate in the study in accordance with the guidelines of the Institutional Review Board of the University of Haifa, and filled in a questionnaire to acquire more information concerning language background and musical experience. The data of three participants was excluded from the analysis because it was noisy (SNR < 1.5) Therefore, data from 34 participants (20 Arabic; 14 Hebrew speakers) is reported below.

Both Hebrew and Arabic speakers learned English as a foreign language at school, starting around 7–8 years of age, and used it academically (mainly for reading academic materials). Arabic speakers started learning Hebrew as a second language for academic education starting around seven years of age. Thus, at the time of the study Arabic speaking participants used Hebrew intensively for a number of years (usually starting at high school) for academic purposes, but spoke Arabic at home and socially. None of the Hebrew-speaking participants reported any experience with Arabic.

Stimuli and electrophysiological recordings

Brainstem responses were elicited by an acoustic click and a speech syllable, /da/, and both brainstem responses were collected in the same manner and during the same recording session in a quiet (though not sound proof) room. All responses were elicited and collected using a Biologic Navigator Pro (Natus Medical Inc., Mundelein, USA). The Navigator's BioMARK (biological marker of auditory processing) module was used to collect the /da/-evoked responses.

For the click-evoked response, the stimuli were 100 μs acoustic clicks. For the speech evoked response, the 40-ms synthesized /da/ provided with the BioMARK module was used. This syllable has an initial noise burst and formant transition between the consonant and a steady-state vowel, and was synthesized with a fundamental frequency (F0) that linearly rises from 103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 ms. The first formant (F1) rose from 220 to 720 Hz while the second and third formants (F2 and F3) decreased from 1700 to 1240 Hz and 2580 to 2500 Hz, respectively, over the duration of the stimulus. The fourth and fifth formants (F4 and F5) were constant at 3600 and 4500 Hz, respectively.

Responses were collected with Ag–AgCl electrodes, and differentially recorded from Cz (active)-to right earlobe (reference), with the left earlobe as ground. Impedance was ≤ 5 kΩ for each electrode and the difference between the electrodes was ≤ 3 kΩ.

The stimuli were presented to the right ear through Bio-Logic insert earphones (580-SINSER) at an intensity of 80 dB SPL while the left ear was unoccluded. Two blocks of speech stimuli (3000 alternating sweeps per block) two click ABR blocks (2000 rarefaction sweeps each), were collected. The click responses were recorded before and after the recording of the speech responses. Both BioMARK blocks were averaged after each recording session to yield a final waveform.

Clicks were presented at a rate of 13.3/s, and were recorded with a 10.66 ms recording window. Responses were on-line filtered from 100 to 1500 Hz. For the /da/, presentation rate was 10.9/s and a 85.33 ms recording window (including a 15-ms pre-stimulus period) was used. Responses were filtered on-line from 100 to 2000 Hz. For both stimuli, sweeps with activity exceeding +23.8 μV were rejected from the average and data collection continued until the target number of artifact-free responses was obtained (2000 for click blocks, 3000 for speech blocks).

Data analysis

The latencies of all waves of interest were identified by the first author, and confirmed by the second author. The second author was blind to the first language of the participants at the time the waveforms were examined.

Click-ABR

Wave V was identified and marked for each subject as data point on the waveform before the negative slope that follows the wave. A normal click evoked response latency is defined as occurring within two standard deviations of the normal population (Hall, 1992). Click-evoked wave V latencies were identified and measured to ensure that no alternations in recording parameters occurred during the recording session such as electrode position and impedance.

Speech-ABR

The electrophysiological brainstem response to a speech sound is a complex waveform that includes transient peaks that reflect the encoding of rapid temporal changes inherent in consonants, which were compared by means of latencies and amplitudes, as well as sustained elements that comprise the frequency following response (FFR), that encodes the harmonic and periodic sound structure of vowels, which were compared using the RMS amplitude in the time domain, and the magnitude of the and spectral components corresponding to F0 and harmonics from 11.4–40.6 ms obtained with a fast FFT in the frequency domain. The characteristic response to the speech stimulus /da/ includes a positive peak (wave V), likely analogous to the wave V elicited by click stimuli, followed immediately by a negative trough (wave A). Following the onset response, peaks C to F are present in the FFR period and the offset (wave O) indicates the cessation of the stimulus (see Figure 1 for a typical response).

Data analysis

Data analysis was automated using routines coded in Matlab 7.7 (The MathWorks, Inc., Natick, USA). Measures that were considered for evaluation were calculated from the averaged final waveform in each subject, which were marked manually based on the markings provided in the BioMARK default unit program (v. 6.3).

Statistical analysis

Responses of Arabic and Hebrew speakers were compared using two tailed t-tests. The data of the Israeli sample was than compared to the 95% confidence interval around the mean computed based on the US norms for the 18–28 years age group.

Results

All 34 participants exhibited normal brainstem responses to click stimuli with wave V latency ranging from 5.28 to 5.95 ms, as used in clinical setting; wave V latency to an 80-dB nHL click, typically occurring at about 5.47 ms for adults (Hood, 1998). Furthermore, the mean latency of peak V of the click-evoked brainstem response was not significantly different between Arabic (5.6 $\pm$ 0.16 ms) and Hebrew (5.5 $\pm$ 0.18 ms) speakers ($p = 0.29$).

Further analysis of ABR data involved assessment of speech-evoked responses in our participants. Measures of transient (see Table1 and Figure 2) and sustained components (see Figure 3, all $p \geq 0.20$) of the brainstem response to speech syllables were compared between the Arabic and Hebrew speaking groups. No

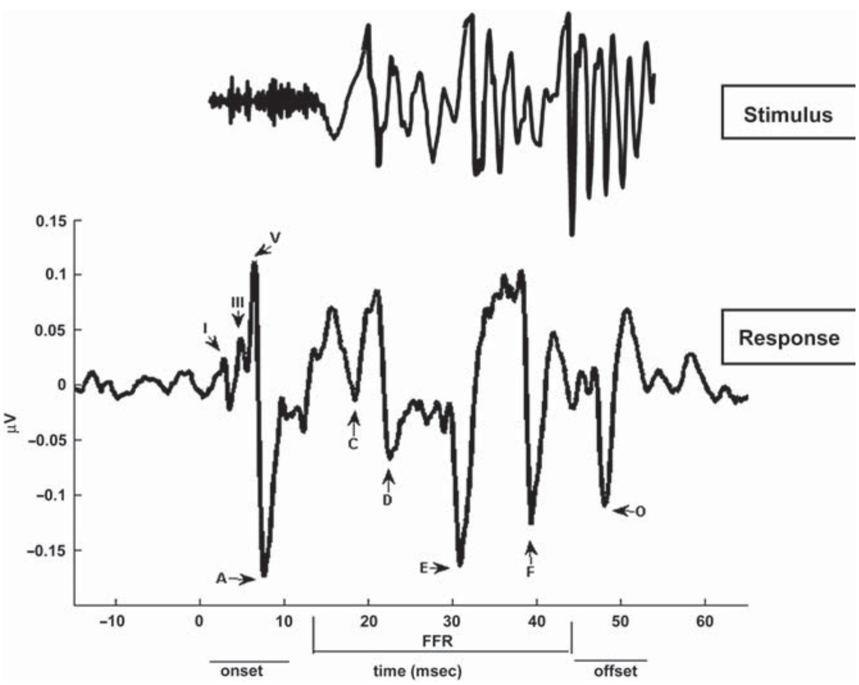

Figure 1. Top: the /da/ stimulus used in the study. Bottom: A typical response in the time domain to the speech syllable /da/ used in the current investigation. Morphologically, the response shows characteristic deflections roughly corresponding to the transient and sustained portions of the evoking stimulus. The offset is the cessation of the stimulus.

Table 1. Means, (SDs), and p values for transient measures in Arabic and Hebrew speaking groups (see Results for details).

	Latency (ms)			*Amplitude (μV)*		
Discrete peaks	*Arabic*	*Hebrew*	*p*	*Arabic*	*Hebrew*	*p*
V	6.608 (0.243)	6.569 (0.209)	0.63	0.123 (0.053)	0.155 (0.065)	0.12
A	7.550 (0.325)	7.605 (0.394)	0.66	−0.196 (0.077)	−0.211 (0.059)	0.56
C	18.481 (0.197)	18.364 (0.264)	0.16	−0.044 (0.091)	−0.012 (0.050)	0.26
D	22.708 (0.851)	22.496 (0.338)	0.38	−0.180 (0.227)	−0.136 (0.056)	0.48
E	31.052 (0.290)	30.952 (0.399)	0.40	−0.174 (0.101)	−0.199 (0.080)	0.46
F	39.481 (0.228)	39.363 (0.150)	0.10	−0.138 (0.090)	−0.131 (0.076)	0.81
O	48.168 (0.390)	48.274 (0.401)	0.45	−0.115 (0.076)	−0.162 (0.062)	0.07
VAslope ($\mu V/ms$)	−0.349 (0.125)	−0.384 (0.164)	0.47	–		

significant between group differences were found in either the transient (Table 1) or in the sustained (Figure 3) components of the response. Figure 2 shows the grand average speech-evoked brainstem response for the Arabic and Hebrew speaking groups. The waveforms are similar to each other and to grand average data from the English speaking population.

Given the lack of differences between the two language groups, the data from Arabic and Hebrew speakers were collapsed and compared to United States (US) English norms for the corresponding age range (18–28). The data from the entire test sample was not statistically distinguishable from the US norms, as evident from the overlap between the US and Israeli confidence intervals (see Table 2). As in Russo et al (2004), we also found that, as latency increased, so did response variability. The SD of latency was smallest for the early onset response waves V and A (0.23 and 0.35 ms, respectively), and increased in the middle with SD latency (up to 0.69 ms) for wave D, and decreased again towards the offset of the response (0.39 ms). Moreover, and as would be suggested by the within-norm latencies exhibited by the Arabic and Hebrew speakers, the BioMARK scores which are computed by the software for clinical use and requires visual identification of waves V and A only, of all listeners in this group were normal and fell within the range of 0–5.

The relationships between click and speech-evoked measures were also examined. In our entire test sample, the latency of click wave V correlated moderately, but significantly, with the latencies of speech onset response V and A ($r = 0.65$, $r = 0.60$, respectively, $p < 0.0001$; see Figure 4). These findings suggest that while there may be some shared processing reflected in the click and speech onset latency measures as shown in the figure, each provides a separate type of information and therefore, a delayed speech evoked measure does not necessarily predict a delayed click. Thus, across both groups of speakers, variability of the major components of the speech evoked response and the correlations between the speech- and click indices were similar to those reported previously for English speakers (Song et al, 2006).

Discussion

The present study was the first (to our knowledge) to compare brainstem encoding of a /CV/ syllable either between Arabic and Hebrew speakers, or between speakers of these Semitic languages and English speakers. The major finding of this comparison has been that the main characteristics of the speech-evoked brainstem response to /da/, that were previously described among English speakers (Johnson et al, 2005; Russo et al, 2004), are similar across those

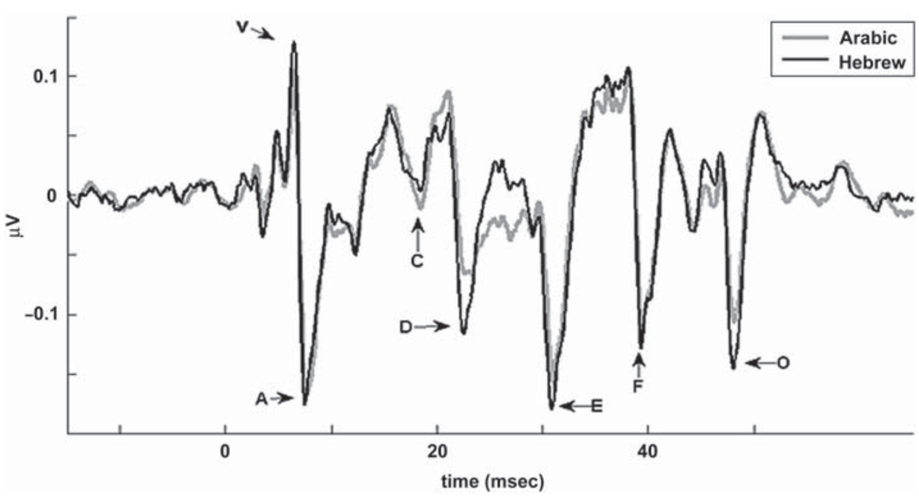

Figure 2. Grand average speech-ABR waveform for both Arabic and Hebrew speaking groups No significant differences were found in the latencies or amplitudes of either of the seven response peaks between Arabic (grey) and Hebrew (black) speaking groups.

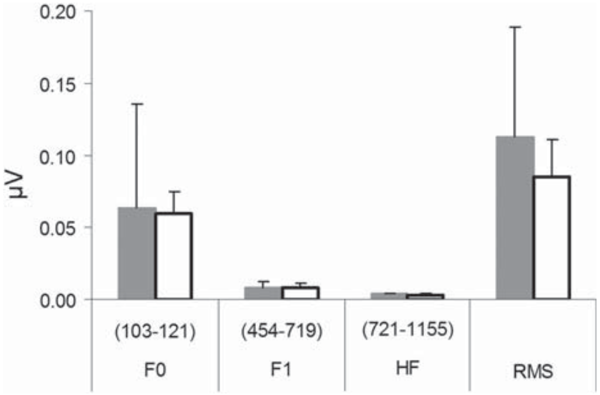

Figure 3. Comparisons of the sustained measures in Arabic and Hebrew speaking groups. Average spectral amplitude was calculated for three frequency ranges and compared between the Arabic groups (left, filled grey bars) and the Hebrew groups (right, empty bars). (1) F0: 103–121 Hz, amplitude of the spectral component corresponding to the stimulus fundamental frequency (F0 amp); (2) F1: low harmonics: 180–410 Hz (not included), and middle harmonics: 454–719 Hz amplitude of the spectral component corresponding to first formant frequencies of the stimulus (F1 amp); (3) HF: high harmonics: 721–1155 Hz.

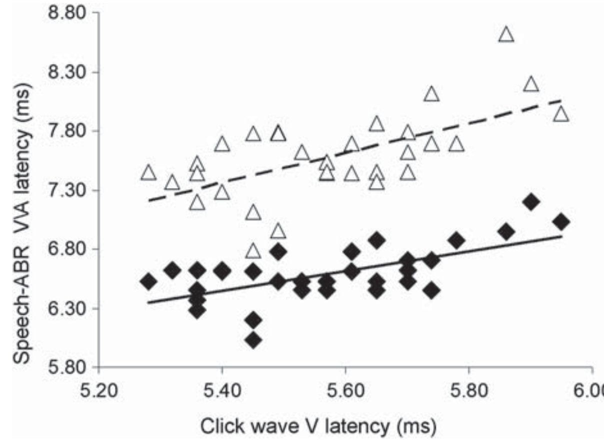

Figure 4. Click-evoked vs. speech-evoked response timing. Speech ABR measures: wave V latency (bottom, diamonds) and wave A latency (top, triangles) as a function of click wave V latency. Lines depict the linear fit of the click and speech measures in the whole group.

three language groups. Furthermore, the pattern of the correlations between the latency of click-evoked response and the onset latency of the speech-evoked response suggests that even when tested with a stimulus that is different in its acoustic properties from the /da/ syllable naturally produced by Hebrew and Arabic speakers, the speech-evoked response still provides information that cannot be fully accounted for by the click response (Song et al, 2006).

The Israeli data presented here was statistically indistinguishable from the US norms for this age group. Therefore the brainstems of English speakers for whom the study /da/ stimulus closely mimics the acoustic-phonetics of their first language, do not appear to encode it differently than the brains of Hebrew and Arabic speakers for whom the acoustics of the test stimulus were foreign. This finding seems in contrast to the findings that the brainstems of Mandarin speakers encode pitch patterns that convey linguistic meaning in Mandarin, but not in English, more robustly than English speakers (Krishnan et al, 2005, 2009). However, it is also possible that the lack of difference stems from the fact that while our Arabic and Hebrew speaking participants have relatively little experience with the English VOT, the +5 ms VOT used in the current study is closer to the VOT of the consonant /t/ in these languages. Thus, the stimulus we used still carried some meaning to our subjects. Indeed, when asked in the end of the experiment what they heard, more than half of the participants reported they heard a /ta/ sound.

Table 2. Comparisons between the Israeli test sample and US norms on timing measures. Mean and (SD) of the test sample are shown in the middle of the left section, among the lower and upper bounds of the sample 95% confidence interval (CI). Right-most column shows the US 95% CI. The overlap between the Israeli and US CIs is equivalent to an insignificant result of a t-test with $\alpha = 0.05$.

	Latency (ms)			
	Israeli 95% CI			*US 95% CI*
Discrete peaks	*lower*	*Mean (SD)*	*upper*	
V	6.50	6.59 (0.23)	6.66	6.63–6.74
A	7.43	7.57 (0.35)	7.69	7.51–7.68
C	18.35	18.43 (0.23)	18.52	18.35–18.67
D	22.37	22.62 (0.69)	22.88	22.62–23.00
E	30.88	31.01 (0.34)	31.13	30.90–31.15
F	39.36	39.43 (0.21)	39.50	39.45–39.69
O	48.06	48.19 (0.39)	48.35	48.14–48.36
VA slope (μV/ms)	−0.42	−0.37 (0.14)	−0.31	−0.43–0.36

It is also interesting to note that even though, as a group, our Israeli sample did not differ from the US normative sample, 22 of our participants (nearly 65%) had wave V latencies shorter than 6.63 ms (the lower bound of the US confidence interval). We believe the most likely reason for this observation is chance variation and our relatively small sample size. However, it could also be speculated that it results from the fact that our participants may have perceived an ambiguous stimulus (something in between /da/ and /ta/) and were trying to force it into a native phonetic category. Indeed, when asked about the identity of the stimulus, more than a third of the current sample indicated they were hesitant (four participants said they did not know what the stimulus was; the others said they were not confident in their response). More research with a wider range of stimuli and perceptual tasks is required to address this point.

Conclusions

The major characteristics of the brainstem response to the speech sound /da/ are similar across speakers of English, Hebrew, and Arabic. Taken together, our findings thus suggest that the US based procedure and most importantly the scoring procedure that is based on normative data obtained in US English speakers can be used in speakers of other languages without having to replace the synthetic syllable used with a rendition more similar to the natural acoustic form of the specific language (Arabic and Hebrew in our case). As a growing body of literature has revealed that brainstem encoding of specific elements of speech sounds are impaired in clinical populations such as language based learning disabilities (see Banai et al, 2007 for review), the current study suggests that the US based procedure can be used to supplement diagnosis in these non English speaking groups.

The current study also suggests that the early processing of stop consonants may be similar among speakers of different languages in which this consonant is present and therefore, the US norms can probably be used with other non English speaking groups in addition to Arabic and Hebrew speakers.

Acknowledgements

We thank all the study participants. Special thanks to Nina Kraus for providing us with the United States BioMARK norms for this age group. This study was supported by the Israel Science Foundation (LHSI 1842/07 to K.B.). Parts of this work have been presented at the annual meeting of the Israeli Society of Auditory Research, Tel-Aviv, October 13, 2009.

Declaration of interest: The authors report no conflict of interest.

References

Abrams D.A., Nicol T., Zecker S.G. & Kraus N. 2006. Auditory brainstem timing predicts cerebral asymmetry for speech. *J Neurosci*, 26(43), 11131–11137.

ANSI. 1989. *Specifications for Audiometers (ANSI S3.6-1989)*. New York: American National Standards Institute.

Banai K., Abrams D. & Kraus N. 2007. Sensory-based learning disability: Insights from brainstem processing of speech sounds. *Int J Audiol*, 46(9), 524–532.

Banai K., Hornickel J., Skoe E., Nicol T., Zecker S. et al. 2009. Reading and subcortical auditory function. *Cereb Cortex*, 19(11), 2699–2707.

Banai K., Nicol T., Zecker S.G. & Kraus N. 2005. Brainstem timing: Implications for cortical processing and literacy. *J Neurosci*, 25(43), 9850–9857.

Cacace A.T. & McFarland D.J. 2005. The importance of modality specificity in diagnosing central auditory processing disorder. *Am J Audiol*, 14(2), 112–123.

Chandrasekaran B. & Kraus N. 2009. The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*.

Galbraith G.C., Arbagey P.W., Branski R., Comerci N. & Rector P.M. 1995. Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport*, 6(17), 2363–2367.

Hall J.W. 1992. *Handbook of Auditory Evoked Responses*. Needham Heights, USA: Allyn & Bacon.

Hood L.J. 1998. *Clinical Applications of the Auditory Brainstem Response*. San Diego, USA: Singular Publishing Group, Inc.

Jacobsen J. 1985. *The Auditory Brainstem Response*. San Diego, USA: College-Hill Press.

Johnson K.L., Nicol T., Zecker S.G., Bradlow A.R., Skoe E. et al. 2008. Brainstem encoding of voiced consonant: Vowel stop syllables. *Clin Neurophysiol*, 119(11), 2623–2635.

Johnson K.L., Nicol T.G. & Kraus N. 2005. Brain stem response to speech: A biological marker of auditory processing. *Ear Hear*, 26(5), 424–434.

Josey A. 1985. Auditory brainstem response in site of lesion testing. In: J. Katz (ed.), *Handbook of Clinical Audiology*. Baltimore, USA: Williams and Wilkins, pp. 534–548.

Khateeb G. 2000. VOT production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics*, 8, 95–122.

King C., Warrier C.M., Hayes E. & Kraus N. 2002. Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neurosci Lett*, 319(2), 111–115.

Kraus N. & Nicol T. 2005. Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. *Trends Neurosci*, 28(4), 176–181.

Krishnan A., Swaminathan J. & Gandour J.T. 2009. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *J Cogn Neurosci*, 21(6), 1092–1105.

Krishnan A., Xu Y., Gandour J. & Cariani P. 2005. Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res*, 25(1), 161–168.

Maddieson I. 1984. *Patterns of Sound*. Cambridge, UK: Cambridge University Press.

Møller A.R. & Jannetta P. 1985. Neural generators of the auditory brainstem response. In: J.T. Jacobson (ed.), *The Auditory Brainstem Response*, San Diego: College-Hill Press, pp. 13–32.

Musiek F.E. 1991. Auditory evoked responses in site-of lesion assessment. In: W.F. Rintelmann (ed.), *Hearing Assessment*. Austin: PRO-ED, pp. 383–427.

Polley D.B., Steinberg E.E. & Merzenich M.M. 2006. Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci*, 26(18), 4970–4982.

Raphael L.J., Tobin Y., Faber A., Most T., Kollia H.B. et al. 1995. Intermediate values of voice onset time. In: F. Bell-Berti & L.J. Raphael (eds.), *Producing Speech: Contemporary Issues*. New York: AIP Press.

Russo N.M., Nicol T., Musacchia G. & Kraus N. 2004. Brainstem responses to speech syllables. *Clin Neurophysiol*, 115(9), 2021–2030.

Russo N.M., Nicol T.G., Zecker S.G., Hayes E.A. & Kraus N. 2005. Auditory training improves neural timing in the human brainstem. *Behav Brain Res*, 156(1), 95–103.

Song J.H., Banai K., Russo N.M. & Kraus N. 2006. On the relationship between speech- and non-speech-evoked auditory brainstem responses. *Audiol Neurootol* (4), 233–241.

Starr A. & Don M. 1988. Brainstem potentials evoked by acoustic stimuli. In: T.W. Picton (ed.), *Handbook of Electroencephalography and Clinical Neurophysiology*. Amsterdam: Elsevier, pp. 97–150.

Tremblay K.L., Piskosz M. & Souza P. 2003. Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol*, 114(7), 1332–1343.

Wible B., Nicol T. & Kraus N. 2005. Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain*, 128 (part 2), 417–423.

Wong P.C., Skoe E., Russo N.M., Dees T. & Kraus, N. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci*, 10(4), 420–422.