

Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages

Jennifer Krizman^{a,b,c}, Viorica Marian^{b,c}, Anthony Shook^{b,c}, Erika Skoe^{a,c}, and Nina Kraus^{a,c,d,e,f,1}

^aAuditory Neuroscience Laboratory, Department of Communication Sciences; ^bBilingualism and Psycholinguistics Laboratory, Department of Communication Sciences; Departments of ^cNeurobiology and Physiology and ^eOtolaryngology; and ^dNorthwestern University Institute of Neuroscience, ^fNorthwestern University, Evanston, IL 60208

Edited by Michael Merzenich, W. M. Keck Center for Integrative Neuroscience, San Francisco, CA, and approved March 28, 2012 (received for review February 6, 2012)

Bilingualism profoundly affects the brain, yielding functional and structural changes in cortical regions dedicated to language processing and executive function [Crinion J, et al. (2006) *Science* 312:1537–1540; Kim KHS, et al. (1997) *Nature* 388:171–174]. Comparatively, musical training, another type of sensory enrichment, translates to expertise in cognitive processing and refined biological processing of sound in both cortical and subcortical structures. Therefore, we asked whether bilingualism can also promote experience-dependent plasticity in subcortical auditory processing. We found that adolescent bilinguals, listening to the speech syllable [da], encoded the stimulus more robustly than age-matched monolinguals. Specifically, bilinguals showed enhanced encoding of the fundamental frequency, a feature known to underlie pitch perception and grouping of auditory objects. This enhancement was associated with executive function advantages. Thus, through experience-related tuning of attention, the bilingual auditory system becomes highly efficient in automatically processing sound. This study provides biological evidence for system-wide neural plasticity in auditory experts that facilitates a tight coupling of sensory and cognitive functions.

brainstem | electrophysiology | multilingualism

Experience shapes how the nervous system responds to sensory input, such that the history of inputs fine-tunes the response to subsequent stimulation (1–4). Through this experience-dependent plasticity, learning-induced changes in the neural processing of behaviorally-relevant stimuli can be seen. For example, with improved juggling ability, novice jugglers demonstrate structural enhancements in a cortical region associated with processing and storage of complex visual motion (5). Similarly, the bilingual, a mental juggler of two languages (6), shows structural and functional enhancements in cortical regions involved in language use and executive control (7, 8), likely resulting from a lifetime of communicating in two languages.

By virtue of interacting in multiple languages, bilinguals, relative to monolinguals, experience an enriched linguistic environment. Although only one language (i.e., the target language) is overtly engaged during communication, the bilingual's non-target language is also coactivated and available, meaning that at any given time, one language is being suppressed (9–11). The need to constantly control two languages confers advantages in the executive system (12, 13), the system that directs cognitive processing. These effects have been demonstrated primarily using visual stimuli and are heightened in children and older adults (e.g., ref. 14). Specifically, bilinguals, relative to monolinguals, are better able to monitor conflicting sensory information and tune into a relevant stimulus or stimulus features amid irrelevant information, via a process known as inhibitory control (12).

Inhibitory control abilities are likely tightly coupled to focusing and maintaining attention (15), and these cognitive processes may still be developing during adolescence (16–18), a major biological transition during the second decade of human life. During this period, synaptic pruning and neural restructuring is occurring in areas of the brain known to be involved in executive function

(19), suggesting that this cognitive ability is not fully crystallized by the onset of this developmental period. On the other hand, the maturation of this process may be malleable with experience as evidenced by the fact that knowing another language impacts the development of executive function in bilingual children (13, 20, 21). Given that the neural infrastructure of executive function is still in flux during adolescence and that bilingual children outperform monolingual peers on tasks of executive function (i.e., inhibitory control and selective attention), we predicted that bilingual adolescents would also demonstrate neural enhancements and behavioral gains in tasks that engage these cognitive abilities. To test this prediction, performance on a task of integrated visual and auditory sustained selective attention was compared between highly-proficient Spanish–English bilinguals and English monolinguals. The neural underpinnings of this bilingual advantage were assessed in the auditory domain.

The auditory system is an ideal model for studying the effects of top-down mechanisms, such as attention or inhibitory control, on sensory processing. Focusing on a target voice in a noisy, complicated soundscape is typical of communication in today's bustling world (e.g., a noisy school cafeteria) and doing this well requires selective attention and active inhibitory control (22). The anatomical substrates of this top-down control, which acts to modulate auditory processing within the subcortical pathway (23), are efferent projections linking the cortex to thalamic, brainstem, and peripheral structures (i.e., cochlea). The auditory brainstem response to complex sounds (cABR), a measure of auditory encoding strength and fidelity, provides a biological snapshot of the putative role of bottom-up and top-down processes that shape experience-dependent plasticity (24, 25). Indeed, the cABR is malleable with training, whether relatively short in duration (26, 27) or more protracted, such as lifelong experience with tonal languages or musicianship (28, 29, 30). Interestingly, musicians also demonstrate advantages in cognitive processing (i.e., attention and memory) that mirror those seen in bilinguals (31, 32); and, in lifelong musicians, these advantages correlate with enhanced neural processing of sound, as measured by the cABR (28, 29, 33). The strongest relationship between musicians' cognitive abilities and neural encoding of sound is seen when the target sound is presented within the context of a competing stream, and these relationships are presumed to develop from the sensory enrichment provided through musical experience.

Given that bilingualism, like musical training, is a form of sensory enrichment that translates to gains in cognitive abilities and that these cognitive gains in attention and memory are known to modulate subcortical processing of auditory stimuli, we predicted that bilinguals would show enhanced cABRs to the speech syllable [da]. Furthermore, because attention and memory facilitate

Author contributions: J.K., E.S., and N.K. designed research; J.K. performed research; J.K. and A.S. analyzed data; and J.K., V.M., A.S., E.S., and N.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: nkraus@northwestern.edu.

the process of focusing on a target speaker in noise, we predicted that the greatest enhancements would be seen when the stimulus is presented in the context of multitalker babble relative to a quiet acoustic background. Specifically, we predicted that the fundamental frequency (F0) of the syllable, a feature that underlies pitch perception, facilitates grouping of auditory objects, is robustly represented in the cABR, and is sensitive to experience and perceptual abilities (34–36), would demonstrate the strongest bilingual advantage.

Results

Bilinguals, relative to monolinguals, showed enhanced subcortical representation of the fundamental frequency of the speech sound as well as improved sustained selective attention. Attention abilities correlated with strength of the F0 when the stimulus was presented in multitalker babble, but not when it was presented in quiet.

Electrophysiology. For both the formant transition and steady-state frequency-following responses of the cABR (20–60 and 60–180 ms, respectively), group differences in the brainstem response to F0 were assessed using a 2 (within subject condition: quiet, babble) \times 2 (between language group: monolingual, bilingual) repeated-measures ANOVA. Within the steady-state frequency-following response, bilinguals demonstrated more robust subcortical encoding of the F0 in both quiet and multitalker babble ($F = 16.866$, $P < 0.0005$); bilinguals were also less affected by the addition of the competing stream, as evidenced by the significant interaction between condition and language group ($F = 6.32$, $P = 0.015$). Post hoc t tests for the steady-state region confirmed this bilingual advantage was evident when the stimulus was presented in quiet ($t_{46} = -3.137$, $P = 0.003$) and in the context of multitalker babble ($t_{46} = -4.946$, $P < 0.00005$) (Fig. 1). During the formant transition, in both quiet and multitalker babble conditions, there was a trend for bilinguals to have more robust F0 encoding ($F = 4.371$, $P = 0.052$), although there was no group by condition interaction ($F = 1.888$, $P = 0.176$).

Sustained Selective Attention. A multivariate ANOVA was used to identify group differences in performance on the three output measures of the attention paradigm (auditory attention, visual attention, full-scale attention). There was a main effect of group ($F = 3.246$, $P = 0.031$) with bilinguals outperforming monolinguals on the auditory ($F = 9.234$, $P = 0.004$), visual ($F = 5.401$, $P = 0.025$) and full-scale (i.e., collapsed across sensory modalities $F = 9.53$, $P = 0.003$) measures of sustained selective attention (Fig. 2).

Correlations Between Sustained Selective Attention and Electrophysiology. To further investigate the effects of bilingualism in the auditory domain, correlations were run on performance on the auditory attention task and neural encoding of the F0 in quiet and in multitalker babble. Although auditory attention did not correlate with F0 amplitude in quiet ($r = 0.254$, $P = 0.081$), neural encoding of the F0 in multitalker babble was positively correlated with sustained selective auditory attention ($r = 0.442$, $P = 0.002$) as can be seen in Fig. 2. When performing separate correlations for the two language groups, the correlation held for the bilingual group ($r = 0.483$, $P = 0.02$), but not the monolingual group ($r = 0.09$, $P = 0.668$), further supporting the notion that bilingual experience is what drives the relationship between F0 encoding and auditory attention.

Additionally, in the visual domain, the same pattern was seen, such that visual attention did not correlate with F0 amplitude in quiet ($r = 0.263$, $P = 0.071$), but did correlate with F0 amplitude in multitalker babble ($r = 0.393$, $P = 0.006$). Again, this relationship was driven by bilingual ($r = 0.482$, $P = 0.020$), but not monolingual ($r = -0.136$, $P = 0.581$), performance.

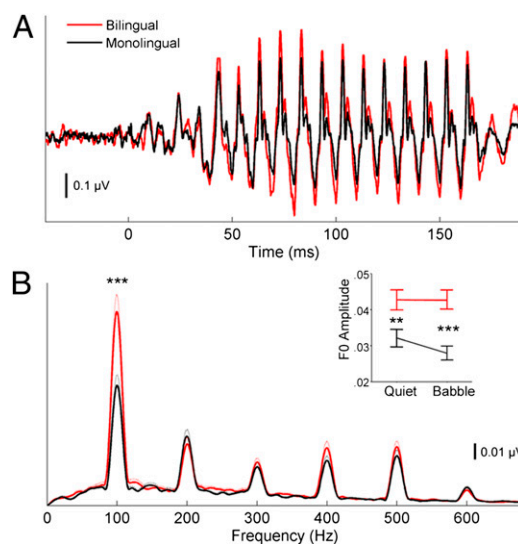
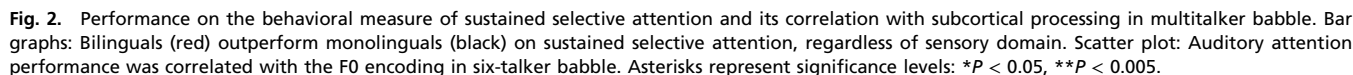


Fig. 1. Subcortical response of bilinguals (red) and monolinguals (black) to the speech sound [da] presented in multitalker babble. (A) Bilinguals show a larger auditory brainstem response relative to monolinguals. (B) Amplitudes of the individual component frequencies in the steady-state (60–180 ms) region of the response to [da] in multitalker babble. Thin lines represent ± 1 SEM. Inset in B displays the mean amplitude (± 1 SE) of the fundamental frequency in quiet and in multitalker babble for bilinguals and monolinguals. For monolinguals, there is a decrease in the amplitude of the fundamental frequency (F0, 100 Hz) when the stimulus is presented in multitalker babble relative to when it is presented in quiet. In contrast, bilinguals show virtually no change in F0 amplitude between the two conditions. Asterisks represent significance levels: $**P < 0.005$, $***P < 0.0001$.

Discussion

This study establishes a neural signature of bilingual experience, whereby bilinguals, relative to monolinguals, have enhanced subcortical representation of the F0 to a target sound presented in a noisy background coupled with heightened sustained selective attention. We argue that enhanced processing of sound and its relationship to attention reveal a biological basis for enhanced flexibility and efficiency of auditory processing in bilinguals. Our findings, combined with previous work, suggest that the bilingual brain undergoes widespread neural specialization that encompasses subcortical and cortical structures relating to language and cognitive processing (7, 8, 37–39). This neural specialization likely results from the bilingual's complex linguistic environment, including a rich diversity in phonetic, phonological, and grammatical structure both within and between talkers. The relationship between sustained attention and subcortical encoding of F0 is a compelling demonstration of experience-dependent plasticity, especially given that there is no correlation between these two measures within the monolingual group. Indeed, these sensory and cognitive changes may be driven synergistically within the bilingual neural system. This is because both robust representation of the F0 and sustained attention are required for accurate perception of auditory events in our acoustically dynamic world; however, the need for well honed sensory and cognitive processes may be higher in a bilingual's phonologically rich and cognitively demanding soundscape.

In bilinguals, immersion in an enriched environment may strengthen attention directed to all linguistic stimuli. With continued exposure, their heightened attention becomes increasingly focused on the behaviorally-relevant stimulus features, such as the F0. Subsequently, the auditory system becomes tuned to automatically process sound more efficiently, as seen here in the cABR. This tuning is likely driven by an interplay of bottom-up and top-down influences, where top-down attentional processes target the most behaviorally relevant features of the stimulus,



In conclusion, we provide evidence that continuously manipulating sounds across two languages leads to an expertise in how sound is encoded in the bilingual brain. The neural enhancements observed in multitalker babble intersect with bilinguals' known advantages in cognitive control and are similar to advantages seen in musicians. In both groups of auditory experts (i.e., musicians and bilinguals), enhanced experience with sound results in an auditory system that is highly efficient, flexible and focused in its automatic sound processing, especially in challenging or novel listening conditions. Thus, converging evidence from both musicians and bilinguals points to subcortical plasticity as providing a biological basis for advantages in real-world experiences with sound.

Materials and Methods

Subjects. Subjects were 48 incoming freshmen attending three public high schools in Chicago, IL. Inclusionary criteria included, normal IQ (Wechsler Abbreviated Scale of Intelligence, WASI; bilinguals: 98.95 ± 8.12 ; monolinguals: 97.7 ± 11.1 ; $F = 0.200$, $P = 0.657$), normal hearing defined as air conduction thresholds < 20 dB normal hearing level (nHL) for octaves from 125 to 8,000 Hz, with no apparent air-bone conduction gap, click-evoked brainstem response latencies within normal limits [the 100- μ s click stimulus was presented at 80 dB sound pressure level (SPL) at a rate of 31 per s], and no external diagnosis of an attention disorder (ADHD or ADD). Monolinguals ($n = 25$; 52% female) and Spanish–English bilinguals ($n = 23$; 56.5% female) were matched on age (bilinguals: 14.8 ± 0.54 y; monolinguals: 14.6 ± 0.46 y; $F = 2.801$, $P = 0.101$) and socioeconomic status (SES) based on maternal education (48) such that 52% of monolingual families and 65% of bilingual families reported between middle school and senior year of high school as the highest maternal education.

Language proficiency was measured by the Language Experience and Proficiency Questionnaire (LEAP-Q, ref. 49). To be included in the study, all subjects had to report high English proficiency (≥ 8 out of 10 for the average self-report of English speaking and understanding proficiency, monolinguals: 9.32 ± 0.8 ; bilinguals 9.36 ± 0.6). Spanish–English bilinguals additionally had to report high Spanish proficiency (≥ 8 out of 10 for the average self-report of Spanish speaking and understanding proficiency; 8.36 ± 0.8). The bilingual subjects also reported speaking and learning Spanish at home (100% of bilinguals spoke Spanish at home; 78% also reported speaking English at home), and their parents/guardians reported that the child spoke two languages. The bilingual subjects were all early bilinguals; their age of first exposure was about 3 y of age (50) for both languages (Spanish: 3.04 ± 2 y; English 3.47 ± 2 y). Of the bilingual subjects, 61% identified Spanish as their native language, whereas 39% considered English their native language.

Monolingual subjects reported no exposure to a second language and their parents also reported that the child only knew English. In the monolingual group, 12 (6 female, 7 low socioeconomic standing) of the subjects were of Hispanic descent. The remaining 13 subjects self-identified as African American ($n = 11$; 6 female, 5 low socioeconomic standing) or Caucasian ($n = 2$; 1 female, 1 low socioeconomic standing). Within the monolingual group, the Hispanic and non-Hispanic subgroups did not differ on any of the measures that were analyzed in the current study (WASI $P > 0.250$; full-scale sustained selective attention $P = 0.161$; auditory sustained selective attention $P > 0.250$; visual sustained selective attention $P = 0.081$; F0 amplitude in quiet $P > 0.250$; F0 amplitude in six-talker babble $P > 0.250$). Given that the groups did not differ on these measures, the results seen in the current study cannot be driven by ethnic or cultural differences between the monolingual and bilingual groups. Because all subjects in the bilingual group were of Hispanic descent, this subanalysis was not performed in this language group.

Sustained Selective Attention. Sustained selective attention was assessed by the Integrated Visual and Auditory Continuous Performance Test (IVA+Plus, www.braintrain.com), a 20-min test with 500 trials of 1s and 2s presented in a pseudorandom order to the visual and/or auditory modalities. For this test, the subject clicks the mouse only when a 1 (but not a 2) is seen or heard. The subject's responses during the test capture abilities of attention, control, and focus, both collectively and individually within the auditory and visual domains. Responses were converted to age-normed standard scores. To assess sustained selective attention in an ecologically valid setting, subjects

were administered this test over headphones at their high school using a laptop computer that was placed 60 cm from the participant.

Electrophysiological Recording. Stimulus and recording. Stimulus and recording parameters followed those described in ref. 23. The complex stimulus [da] is a dynamic, six-formant, 170-ms sound synthesized at a 20-kHz sampling rate using a Klatt synthesizer (51). Except for the initial 5-ms stop burst, this syllable is voiced throughout with a steady fundamental frequency ($F_0 = 100$ Hz). This stimulus is characterized by a 50-ms formant transition (transition between [d] and [a]) followed by a 120-ms steady-state [a] portion during which the formants are unchanging. The [da] stimulus was presented 6,300 times with an 81-ms interstimulus interval presented in alternating stimulus polarities to the right ear at 80-dB SPL through an insert earphone (ER-3; Etymotic Research) using the stimulus presentation software NeuroScan Stim2 (Sound module; Compumedics). Responses, which originate primarily from the inferior colliculus (52), were differentially recorded in a sound-attenuated, electrically-shielded chamber using NeuroScan Acquire4 at a sampling rate of 20 kHz. Ag/Ag-Cl electrodes were applied in a vertical montage from Cz-to-earlobe with forehead as ground. During electrophysiological testing, the participant watched a movie of his or her choice in a comfortable reclining chair. The left ear was unoccluded enabling the participant to hear the movie soundtrack played at < 40 dB SPL, an insufficient intensity to mask the stimulus. The [da] was presented alone and in the context of multitalker babble. The multitalker babble (four female and two male voices) was created by mixing six tracks of English nonsense sentences in Cool Edit Pro, version 2.1 (Syntrillium Software, 2003), into a 45-s duration babble track that was presented at a signal-to-noise (SNR) of +10 dB relative to the [da] based on the root mean square (RMS) amplitude of the entire track.

Data averaging. For both the quiet and babble conditions, electrophysiological responses were off-line bandpass filtered in Neuroscan Edit from 70 to 2,000 Hz (12 dB/octave, zero phase-shift) to include energy within the phase-locking limits of the inferior colliculus (52, 53) and to minimize low-frequency myogenic noise and cortical activity. Responses were then averaged over a -40 - to 190-ms window, with stimulus onset occurring at time 0. An artifact reject criterion of ± 35 μ V was applied, resulting in prestimulus baseline corrected final averages comprising $\sim 6,000$ sweeps.

Analysis of the F0 magnitude of the auditory brainstem response. In MATLAB (Mathworks) a fast Fourier transform was performed separately for the formant transition (20–60 ms) and the steady-state response (60–180 ms). From the resultant spectrum, average amplitudes of specific frequency bins were calculated. Each bin was 40 Hz wide, centered on the stimulus F0 (100 Hz).

ACKNOWLEDGMENTS. We thank the students and their families as well as the teachers and staff (especially Kate Ulett for initiating this neuro-education liaison) of the high schools participating in this study. We also acknowledge the members of the Auditory Neuroscience Laboratory, especially Travis White-Schwoch and Trent Nicol, and the Bilingualism and Psycholinguistics Laboratory for their helpful comments on earlier versions of the manuscript. We thank the members of the Auditory Neuroscience Laboratory that helped with testing, especially Rafael Escobedo. This research is funded in part by National Science Foundation Grant 1015614 (to N.K.), the Mathers Foundation (to N.K.), the Hugh Knowles Center of Northwestern University (to N.K.), National Institute of Child Health and Human Development Grant R01HD059858 (to V.M.), and National Research Service Award Institutional Research Training Grant T32 DC009399-01A10 (to J.K.).

- Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci* 13:87–103.
- Kilgard MP, Merzenich MM (1998) Plasticity of temporal information processing in the primary auditory cortex. *Nat Neurosci* 1:727–731.
- Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223.
- Buonomano DV, Merzenich MM (1998) Cortical plasticity: From synapses to maps. *Annu Rev Neurosci* 21:149–186.
- Draganski B, et al. (2004) Neuroplasticity: Changes in grey matter induced by training. *Nature* 427:311–312.
- Kroll JF (2009) The consequences of bilingualism for the mind and the brain. *Psychol Sci Public Interest* 10:i–ii.
- Abutalebi J, et al. (2011) Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb Cortex*, 10.1093/cercor/bhr287.
- Mechelli A, et al. (2004) Neurolinguistics: Structural plasticity in the bilingual brain. *Nature* 431:757.
- Marian V, Spivey M (2003) Competing activation in bilingual language processing: Within-and between-language competition. *Biling Lang Cogn* 6:97–115.
- Marian V, Spivey M (2003) Bilingual and monolingual processing of competing lexical items. *Appl Psycholinguist* 24:173–193.
- Thierry G, Wu YJ (2007) Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc Natl Acad Sci USA* 104:12530–12535.
- Bialystok E (2011) Reshaping the mind: The benefits of bilingualism. *Can J Exp Psychol* 65:229–235.
- Carlson SM, Meltzoff AN (2008) Bilingual experience and executive functioning in young children. *Dev Sci* 11:282–298.
- Bialystok E, Martin MM, Viswanathan M (2005) Bilingualism across the lifespan: The rise and fall of inhibitory control. *Int J Biling* 9:103–119.
- Reck SG, Hund AM (2011) Sustained attention and age predict inhibitory control during early childhood. *J Exp Child Psychol* 108:504–512.
- van der Molen MW (2000) Developmental changes in inhibitory processing: Evidence from psychophysiological measures. *Biol Psychol* 54:207–239.
- van der Stelt O, Kok A, Smulders FT, Snel J, Boudewijn Gunning W (1998) Cerebral event-related potentials associated with selective attention to color: developmental changes from childhood to adulthood. *Psychophysiology* 35:227–239.
- Luna B, Garver KE, Urban TA, Lazar NA, Sweeney JA (2004) Maturation of cognitive processes from late childhood to adulthood. *Child Dev* 75:1357–1372.

