Supporting Online Material
for
"Human voice recognition depends on language ability"
T.K. Perrachione, S.N. Del Tufo, J.D.E. Gabrieli
Published 29 July 2011, Science 333, 595 (2011)
DOI: 10.1126/science.1207327

Materials & Methods

Participants

Native English-speaking controls (N = 16; 9 female) 18-30 years of age (M = 21.2; SD = 2.98) with a self-reported history free from neurological, psychiatric, speech, language, or reading impairments were matched with individuals with dyslexia (N = 16; 8 female) between 16-38 years of age (M = 24; SD = 6.8). Inclusionary criteria for dyslexia consisted of a prior clinical diagnosis or lifelong history of reading disability and scoring below the 16th percentile (one standard deviation below the age-normed mean) on any two subtests from the following standard clinical reading and language assessments: *Woodcock Reading Mastery Test-Revised* (WRMT-R/NU) (7), *Test of Word Reading Efficiency* (TOWRE) (8), and *Comprehensive Test of Phonological Processing* (CTOPP) (9). Groups were matched based on cognitive performance ("Matrices" and "Block Design" from the *Wechsler Abbreviated Scale of Intelligence*, WASI; (10)), working memory (*Wechsler Adult Intelligence Scale* WAIS-IV; (11)), age, and education. All participants were right-handed based on questionnaire responses (adapted from the Edinburgh Handedness Inventory; (12)). All participants indicated no prior experience with Mandarin Chinese. Informed written consent, approved by the MIT Committee on the Use of Humans as Experimental Subjects, was obtained from all participants prior to participation.

Stimuli

Two sets of ten sentences designed for acoustic assessment were recorded for this experiment: one spoken in English (13), the other in Mandarin (14). The English sentences were read by five male native speakers of American English (aged 19-26 years, M = 21.6). The Mandarin sentences were read by five male native speakers of Mandarin Chinese (aged 21-26 years, M = 22.6). No talker read sentences in both languages, and none of the individuals recorded as talkers participated in the listening experiment. Recordings were made in a sound-attenuated chamber via a SHURE SM58 microphone using a Creative USB Sound Blaster Audigy 2 NX sound card, sampled at 22.05 kHz and normalized for RMS amplitude to 70 dB SPL. Recordings of sentences were 1.46sec to 4.09sec in duration (M = 2.43, SD = 0.54). In each language, five sentences were used during the familiarization and practice phases, and all ten were used during the final voice recognition test. These stimuli have been used in prior experiments of voice recognition by native speakers of English and Chinese (2,15).

Procedure

Participants learned to identify five talkers in each of two language conditions (English and Mandarin) from the sound of their voice. Each talker was associated with a distinct cartoon avatar (Fig. S1). Training and testing on voice recognition were completed in each language condition separately, and the order was counterbalanced across listeners. During an initial familiarization phase, participants heard each of the voices in succession while the corresponding avatars were displayed on a computer screen. Participants then actively practiced identifying the talkers with corrective feedback: The five avatars appeared on the screen while a recording from one talker was played, and participants selected the avatar matching the voice they heard. If participants selected incorrectly, the computer indicated the correct response. During the task, all instructions were presented both as text on the screen and as auditory prompts recorded by an additional female talker. The familiarization and active practice

phases were repeated over five training sentences, and each sentence was practiced ten times. Following training, participants undertook a 50-item talker identification test, in which they identified the voices without feedback. Participants completed the self-paced experiment in a quiet room. Stimuli were presented binaurally at a comfortable level over Sennheiser HD-250 linear II circumaural headphones using an Edirol UA-25EX sound card.

Supporting Online References

8. Woodcock, R.W. (1998). *Woodcock Reading Mastery Tests – Revised/Normative Update (WRMT-R/NU)*. Circle Pines, MN: American Guidance Service.
9. Torgesen, J.K., Wagner, R., & Rashotte, C. (1999). *Test of Word Reading Efficiency (TOWRE)*. Austin, TX: Pro-Ed.
10. Wagner, R., Torgesen, J.K., & Rashotte, C. (1999). *Comprehensive Test of Phonological Processing (CTOPP)*. Austin, TX: Pro-Ed.
11. Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence* (3rd Ed.). San Antonio, TX: The Psychological Corporation.
12. Wechsler, D. (2008). *Wechsler memory scale.* (4th Ed.). San Antonio, TX: Pearson
13. Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97-113.
14. Institute of Electrical and Electronics Engineers (1969). IEEE recommended practices for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics,* 17, 225-246.
15. Open Speech Repository. (2005). *Open Speech Repository*. Retrieved August 2005 from http://www.voiptroubleshooter.com/open_speech/index.html
16. Perrachione, T.K., Pierrehumbert, J.B., & Wong, P.C.M. (2009). Differential neural contributions to native- and foreign-language talker identification. *Journal of Experimental Psychology - Human Perception and Performance*, 35, 1950-1960.

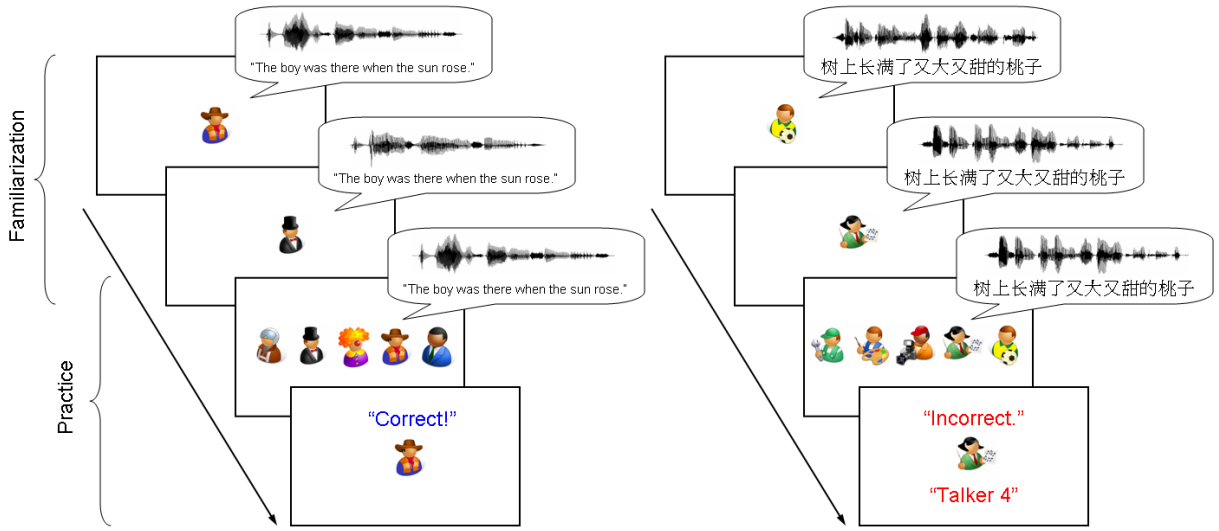

Figure S1: Graphical depiction of training paradigm. Listeners learned to recognize 5 English and 5 Chinese talkers from the sound of their voice. Talkers were paired with distinct icons, and acoustic waveforms illustrate variability in talker characteristics. Participants were familiarized with individual talkers and practiced recognizing them with feedback. Talker identification accuracy in each condition was assessed with a post-training test.

		Control		Dyslexia		
Test	Subtest	Raw Score	Standard Score	Raw Score	Standard Score	Cohen's *d*
WASI						
	Block Design	60.6 ± 7.4	61.4 ± 5.4	53.9 ± 13.4	57.1 ± 8.1	0.644
	Matrix Reasoning	29.8 ± 2.3	58.6 ± 4.7	29.8 ± 3.8	58.4 ± 67.5	0.022
	Performance IQ	119.9 ± 8.5	116.8 ± 8.8	115.5 ± 11.0	112.2 ± 11.0	0.487
CTOPP						
	Elision	18.9 ± 1.5	10.9 ± 1.5	15.7 ± 3.4	8.3 ± 22.7	1.401
	Blending Words	18.5 ± 1.4	12.5 ± 1.4	14.1 ± 3.5	8.9 ± 2.5	1.833
	Non-word Repetition	15.7 ± 2.1	11.8 ± 1.9	9.7 ± 1.9	6.8 ± 1.3	3.134
	Rapid Digit Naming	22.9 ± 5.1	10.4 ± 2.6	31.2 ± 8.6	6.9 ± 2.8	1.315
	Rapid Letter Naming	22.8 ± 4.5	11.1 ± 2.7	35.4 ± 8.9	5.5 ± 2.9	2.072
	Rapid Object Naming	39.4 ± 5.6	10.6 ± 3.1	51.7 ± 8.5	6.3 ± 2.1	1.721
WRMT-R/NU						
	Word ID	101.7 ± 2.8	112.4 ± 8.1	91.3 ± 6.9	94.2 ± 7.8	2.368
	Word Attack	42.4 ± 1.7	121.3 ± 13.3	31.6 ± 4.4	92.0 ± 8.0	2.762
TOWRE						
	Sight Word Reading	99.0 ± 8.3	106.1 ± 11.3	79.9 ± 15.4	85.3 ± 13.0	1.765
	Decoding	54.7 ± 6.1	104.8 ± 11.4	35.4 ± 10.7	76.7 ± 18.3	1.899
	Total	200.0 ± 40.9	106.5 ± 11.6	166.2 ± 21.9	75.4 ± 19.8	1.975
WAIS-IV						
	Digit Span Total	8.6 ± 1.5	9.6 ± 1.9	7.9 ± 1.9	8.8 ± 2.5	0.403
Age (years)		21.3 ± 2.7		23.9 ± 6.8		0.536
Education (years)		15.3 ± 1.5		15.1 ± 2.4		0.286

Table S1: Cognitive and behavioral assessment profile of participants. Dyslexic and control participants were matched on performance IQ, working memory, age, and education, but differed on measures of reading ability and phonological processing. Values indicate mean ± standard deviation. Abbreviations: WASI: *Wechsler Abbreviated Scale of Intelligence, 3rd Ed.* (10); CTOPP: *Comprehensive Test of Phonological Processing* (8); WRMT-R/NU: *Woodcock Reading Mastery Test-Revised* (6), TOWRE: *Test of Word Reading Efficiency* (8); WAIS-IV: *Wechsler Adult Intelligence Scale* (9). Cohen's *d* shows the effect size of the group difference in standard scores.