Human Voice Recognition Depends on Language Ability

Tyler K. Perrachione,[1]* Stephanie N. Del Tufo,[1] John D. E. Gabrieli[1,2]*

The ability to recognize individual conspecifics from their communicative vocalizations is an adaptive trait evinced widely among social and territorial animals, including humans. Studies of human voice recognition compare this ability to nonverbal processes, such as human perception of faces or nonhuman animals' perception of vocalizations (1). However, the human voice is also the principal medium for the human capacity of language, as conveyed through speech. Human listeners are more accurate at identifying voices when they can understand the language being spoken (2), an advantage thought to depend on listeners' knowledge of phonology—the rules governing sound structure in their language. Leading theories of dyslexia propose that impoverished phonological processing often underlies impaired reading ability in this disorder (3, 4). We therefore hypothesized that, if voice recognition by human listeners relies on linguistic (phonological) representations, listeners with dyslexia would be impaired compared with control participants when identifying voices speaking their native language (because of impaired phonological processing) but unimpaired in voice recognition for an unfamiliar, foreign language (where both individuals with and without dyslexia lack relevant language-specific phonological representations).

We assessed participants with and without dyslexia for their ability to learn to recognize voices speaking either the listener's native language (English) or an unfamiliar, foreign language (Mandarin Chinese). In each language, participants learned to associate five talkers' voices with unique cartoon avatars and were subsequently tested on their ability to correctly identify those voices. The participants' task was to indicate who of the five talkers spoke in each trial [five-alternative forced choice; chance = 20% accuracy (5)]. Despite using the same vocabulary, all speakers of a language differ in their pronunciations of words (6), and listeners can use their phonological abilities to perceive these differences as part of a speaker's vocal identity. A repeated-measures analysis of variance revealed that, compared with controls, dyslexic participants were significantly impaired at recognizing the voices speaking English but unimpaired for those speaking Chinese (group × condition interaction, $P < 0.0006$) (Fig. 1).

English-speaking listeners with normal reading ability were significantly more accurate identifying voices speaking English than Chinese (paired t test, $P < 0.0005$), performing on average 42% better in their native language (7). English-speaking listeners with dyslexia were no better able to identify English-speaking voices than Chinese-speaking ones (paired t test, $P = 0.65$), with an average performance gain of only 2% in their native language. Correspondingly, dyslexic listeners were significantly impaired compared with controls in their ability to recognize English-speaking voices (independent-sample t test, $P < 0.0021$). Dyslexic listeners were as accurate as controls when identifying the Chinese-speaking voices (independent-sample t test, $P = 0.83$), demonstrating that their voice-recognition deficit was not due to generalized auditory or memory impairments. Moreover, for the dyslexic partic-

ipants, greater impairments on clinical assessments of phonological processing were correlated with worse accuracy for identifying English-speaking voices (both Pearson's $r > 0.6$, $P < 0.015$). Although the diagnostic criterion for dyslexia is impairment in developing typical reading abilities, these data show that reading difficulties are accompanied by impaired voice recognition. This inability to learn speaker-specific representations of phonetic consistency may reflect a weakness in language learning that contributes to impoverished long-term phonological representations in dyslexia.

For humans, the ability to recognize one another by voice relies on the ability to compute the differences between the incidental phonetics of a specific vocalization and the abstract phonological representations of the words that vocalization contains. When the abstract linguistic representations of words are unavailable (because the stimulus is unfamiliar, as in foreign-language speech) or impoverished (because native-language phonological representations are compromised, as in dyslexia), the human capacity for voice recognition is significantly impaired. This reliance on our faculty for language distinguishes human voice recognition from the recognition of conspecific vocalizations by other nonhuman animals.

References and Notes

1. P. Belin, S. Fecteau, C. Bédard, *Trends Cogn. Sci.* **8**, 129 (2004).
2. T. K. Perrachione, P. C. M. Wong, *Neuropsychologia* **45**, 1899 (2007).
3. L. Bradley, P. E. Bryant, *Nature* **301**, 419 (1983).
4. J. D. E. Gabrieli, *Science* **325**, 280 (2009).
5. Materials and methods are available as supporting material on *Science* Online.
6. J. Hillenbrand, L. A. Getty, M. J. Clark, K. Wheeler, *J. Acoust. Soc. Am.* **97**, 3099 (1995).
7. Native Chinese-speaking controls exhibit the opposite pattern, recognizing Chinese-speaking voices more accurately than English-speaking ones (2), revealing the critical factor to be listeners' language familiarity, not properties inherent to the voice stimuli or languages themselves.

Supporting Online Material

www.sciencemag.org/cgi/content/full/333/6042/595/DC1
Materials and Methods
Fig. S1
Table S1
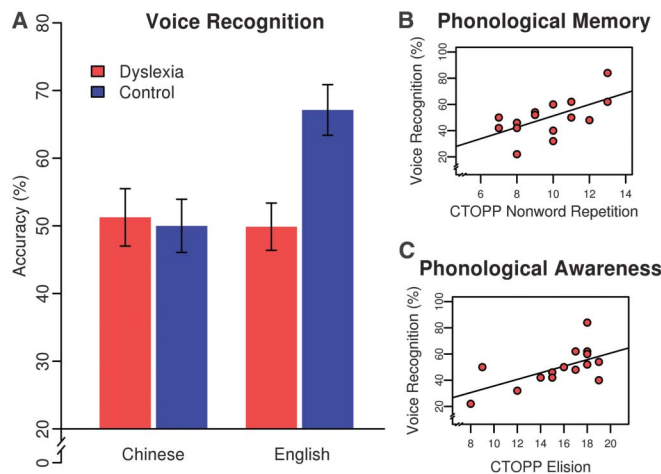References (8–16)

[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]McGovern Institute for Brain Research and Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: tkp@mit.edu (T.K.P.); gabrieli@mit.edu (J.D.E.G.)

Fig. 1. (A) Mean voice-recognition performance of dyslexic and control listeners (error bars indicate SEM). All individuals scored above chance (20%), shown as baseline. (**B** and **C**) Relationships between clinical measures of language (phonological) ability in dyslexia and voice-recognition ability. CTOPP, Comprehensive Test of Phonological Processing.