Sex Differences in the Discrimination of Non-Native Speech Sounds

Vera Kempe (v.kempe@abertay.ac.uk)
Division of Psychology, University of Abertay Dundee
Dundee, DD1 1HG, United Kingdom

John C. Thoresen (john.thoresen@epfl.ch)
Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne,
1015 Lausanne, Switzerland

Patricia J. Brooks (patricia.brooks@csi.cuny.edu)
Department of Psychology, College of Staten Island and the Graduate Center, City University of New York
Staten Island, NY 10314, USA

Abstract

This study examined sex differences in the discrimination of minimal pairs of foreign language (non-native) tonemes. Adult native speakers of English (237 women and 177 men), with no prior exposure to a tonal language, performed an AX-task, which required them to discriminate between rising and falling-rising Norwegian tonemes. When controlling for nonverbal intelligence, prior exposure to foreign languages, and age, sensitivity measures (*A'*) showed a clear male advantage. Thus, the sex differences previously observed in non-linguistic temporal processing tasks appear to extend to the discrimination of unfamiliar non-native speech sounds. These sex differences in auditory processing may be due to anatomical differences between men and women in the ratio of white to grey matter in the left hemisphere, which, in turn, might affect speed of neural transmission. These findings contribute to the ongoing debate on cognitive effects of putative sex differences in intra- and inter-hemispheric connectivity.

Keywords: non-native speech perception; tonal contrast; sex differences; adult L2 learning; auditory processing.

Introduction

Auditory processing of temporal sequences underlies the neural representation of speech and has been implicated in impairments in language development; e.g., dyslexia and Specific Language Impairment (Goswami et al., 2002; Talcott et al., 2000; Tallal, 1980). However, little is known about individual differences in auditory processing in the non-clinical adult population, and, specifically, individual differences in the ability of adults to discriminate the speech sounds of foreign (non-native) languages. Illuminating the basis of individual differences in non-native speech processing may help to explain some of the considerable variance in outcomes observed among adult foreign language (L2) learners (Johnson & Newport, 1989). So far, only a few studies have explored individual differences in the processing of non-native speech sounds (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Golestani & Zatorre, 2009). Thus, we know very little about which factors, besides age of first exposure (Flege, MacKay & Meador, 1999), make an adult more or less successful in processing non-native speech sounds.

Research on temporal processing as a predictor of psychometric intelligence (Rammsayer & Brandler, 2007) and working memory capacity (Troche & Rammsayer, 2009) has revealed a sex difference, with men outperforming women in temporal order judgments (Szelag et al., 2011; Wittman & Szelag, 2003) and temporal discrimination tasks (Rammsayer & Troche, 2010). Factor-analytical approaches have traced the male advantage to a latent variable – temporal resolution power, which has been linked to neural oscillation rate determining speed and accuracy of neural transmission (Jensen, 1982). The male advantage is not confined to the auditory modality, however, but has also been observed for tactile temporal processing (Rostad, Mayer, Fung & Brown, 2007), suggesting that it affects general temporal processing in the sub-second range.

In addition to sex differences in pure temporal information processing tasks, there is evidence for a male advantage in the discrimination of pitch contours of computer-generated waveforms, comprising a fundamental frequency and two formants, which were presented binaurally (McRoberts & Sanders, 1992). Pitch contour discrimination requires sensitivity to changes in pitch over time and therefore relies on temporal processing. Rapidly changing values of one or several acoustic parameters (e.g., formant transitions) play a crucial role in distinguishing different speech sounds—for example, notoriously difficult phonological contrasts like the dental-retroflex contrast (for English speakers) or the r/l contrast (for Japanese speakers) require sensitivity to rapid spectral changes. The present study therefore aims to examine whether a male advantage can also be found in the ability to discriminate natural non-native speech contrasts. Natural speech sounds differ from synthetic stimuli in their greater variability within speech sound categories and in the complexity of their acoustic characteristics.

We chose to examine sensitivity to lexical tones as one example of such a non-native speech contrast. We used Norwegian tonemes as many dialects of Norwegian have a simple tonal system with pitch accents that distinguish otherwise homophonous bisyllabic words. Detecting these tonal contrasts requires tracking temporal changes in pitch contours of bi-syllabic words. We tested adult native English speakers' sensitivity to the tonal contrast between rising and falling-rising tonemes, which are illustrated in Figure 1.

If sex differences in non-linguistic temporal processing extend to linguistic stimuli we would expect to see a male advantage in the processing of an unfamiliar Norwegian tonal contrast by native English speakers.

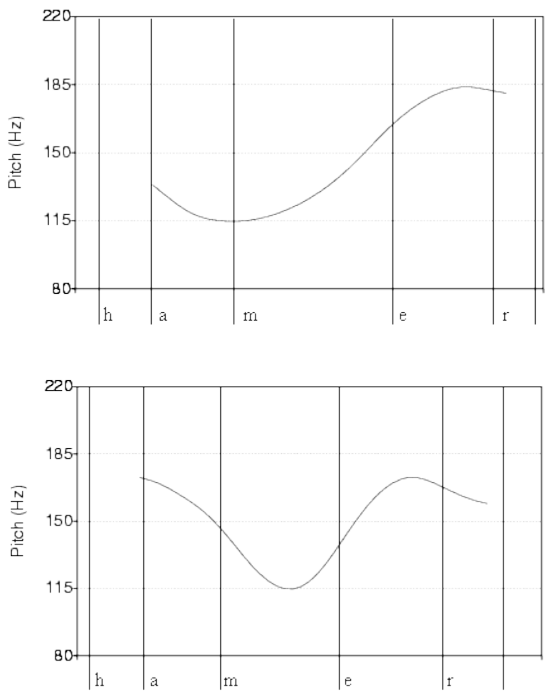

Figure 1: Illustrations of the different pitch contours of minimal pairs of Norwegian tonemes. The upper panel shows the rising tone for the word 'Hammer' [a proper noun]. The lower panel shows the falling-rising tone for the word 'hammer' [the tool][1].

Method

We pooled data from six experiments on non-native discrimination of Norwegian tonal contrasts that were conducted over a period of five years (Kempe, Thoresen & Brooks, 2007, 2008; Kempe, Brooks, Marronaro & Thoresen, 2010; Kempe, Thoresen, Kirk, Brooks & Schaeffler, 2011). These experiments tested native speakers of English differing in dialectal background (American, English and Scottish), and varied with respect to other cognitive abilities tested (e.g., verbal working memory capacity) and other speech sound contrasts presented in addition to the tonal contrast. All six experiments controlled for nonverbal intelligence and prior exposure to other languages. It is necessary to control for nonverbal intelligence because of the well-established link between processing speed and temporal processing on the one hand and psychometric intelligence on the other hand (e.g., van Raavenzwaaj, Brown & Wagenmakers, 2011; Rammsayer & Brandler, 2007; Sheppard & Vernon, 2008). It is also important to make sure that any observed differences cannot be accounted for by prior exposure to tonal contrasts.

Participants: A total of 458 participants (197 men) were tested in various locations in the United Kingdom and in the United States (New York City). Participants' mean age was 22.8 years (range 17-61 years). Of the participants, 282 (117 men) were native speakers of American English and 176 (80 men) were native speakers of Scottish English.

An additional ten native speakers of Norwegian (four men), aged 20 to 22 years, were tested to confirm that the chosen tonal contrast can reliably be discriminated by native speakers of the language.

Materials: To capture the within-category variability characteristic of natural speech sounds, a male native speaker of Norwegian recorded two different instances of each of 16 bisyllabic Norwegian words comprising 8 minimal tonal contrast pairs. In half of the pairs, the first stressed syllables contained short vowels (mean length 80 ms); in the other half they contained long vowels (mean length 144 ms). These eight minimal pairs are listed in Table 1. Stimuli were recorded at a sampling rate of 44.1 kHz and presented to participants through Sennheiser headphones.

Table 1: Minimal pairs of Norwegian words used for tone discrimination. Note that the members of a pair are homophones despite differences in spelling.

rising tone	falling-rising tone
short vowel	
bønder [farmers]	*bønner* [beans]
lammet [lamb]	*lamme* [to paralyze]
sulten [hunger]	*sulten* [hungry]
verket [creation]	*verke* [to ache]
long vowel	
bøter [fines]	*bøter* [to repent]
laget [team]	*lage* [to make]
suget [suction]	*suge* [to suck]
været [weather]	*være* [to be]

To establish whether pitch contours were indeed sufficiently distinct between the two tonemes, we measured pitch of the steady-state part of the vowels in the first and second syllable. As Figure 1 indicates, there should be a larger difference between pitch on the first and the second syllable for a rising pitch contour than for a falling-rising pitch contour. Pitch measurements confirmed that the median pitch difference between syllables for the rising pitch contour (38 Hz) was significantly larger than the median pitch difference for the falling-rising pitch contour (4 Hz), Mann-Whitney $U = 2.0$, $p < .001$, $r = .84$, indicating

[1]Figure reprinted from the project Lingo resource at the Norwegian University of Science and Technology, Trondheim, http://www.ling.hf.ntnu.no/ipa/no/tema_008.html

a sufficiently large measurable difference in acoustic characteristics between the two tonemes.

Procedure: The Norwegian tonal contrasts were presented in an AX discrimination procedure requiring participants to make judgments about whether pairs of words sounded the 'same' or 'different'. The 32 'same' trials consisted of different instances of the same word spoken with the same pitch accent. The 32 'different' trials consisted of minimal pairs of words spoken with different pitch accents. The two words in each pair were presented with an inter-stimulus interval of 200 ms.

Participants also completed the Cattell Culture-Fair Test of Nonverbal Intelligence, Scale 3, Form A (Cattell & Cattell, 1973), and a language background questionnaire, used to confirm participant status as a native English speaker and to inquire about prior exposure to languages other than English. Participants were asked to rate their reading, writing, speaking and comprehension abilities in each of their languages on a scale from 1 (rudimentary) to 6 (native-like).

Results

Norwegian native speakers: Each participant's performance was converted to an *A'* score, a measure of sensitivity that corrects for individual differences in bias. *A'* is a non-parametric analogue to *d'* and has values ranging from 0 and 1, with 0.5 corresponding to chance. The mean *A'* score for the native speakers of Norwegian was .93 (*SD* = 0.04), which supports the validity of our stimuli and confirms that discrimination of the native tonal contrast did not pose any problems for native speakers.

English native speakers: Seventeen men and 18 women reported some familiarity with a tonal language (e.g., Chinese). A further three men and six women failed to provide proficiency ratings in the language background questionnaire. All these participants were excluded from the analyses leaving a total of 177 men and 237 women.

The mean *A'* score for the entire sample was 0.71 (*SD* = 0.14). Table 2 presents means and standard deviations for men and women, along with results of Bonferroni-corrected t-tests comparing men and women on age, Culture Fair non-verbal intelligence test scores (CF IQ), number of learned foreign languages (L2s), and mean proficiency self-ratings for first and second L2s. If participants had not studied any L2, the corresponding rating scores were set to 0. These comparisons showed that while women had higher proficiency self-ratings in their first L2, women's discrimination of Norwegian tonemes was significantly lower than men's.

Table 2: Means and standard deviations (in parenthesis) of the various measurements for men and women. The last column shows the results of a t-test (** indicates significance after Bonferroni correction at $p < .01$).

	men	women	$t(412); p$
age	22.5 (7.1)	23.0 (7.3)	-0.56; .574
CF IQ score	25.0 (5.1)	24.0 (5.1)	1.96; .050
# L2s	1.4 (0.8)	1.6 (0.8)	-2.50; .013
self-rating 1st L2	2.3 (1.3	2.7 (1.5)	-2.90; .004**
self-rating 2nd L2	0.7 (1.1)	1.0 (1.2)	-2.40; .018
A'	0.74 (0.13)	0.70 (0.14)	3.18; .002**

To account for potential uncontrolled effects of the different testing conditions and contexts in the six experiments, we computed standardized *A'* scores for each experiment separately. These standardized *A'* scores served as the dependent variable in a multiple regression analysis with age, Culture Fair nonverbal intelligence test scores, the various language background variables, and sex (coded as a dummy variable) as predictors (see Table 3).

Table 3: Results of a multiple regression analysis of all predictors and sex, coded as dummy variable, on standardized *A'* scores for toneme discrimination ability.

	β	*t*	*p*
age	-.073	-1.50	.134
CF IQ score	.161	3.24	.001
# L2s	.049	0.69	.492
self-rating 1st L2	-.071	-1.28	.201
self-rating 2nd L2	.076	1.04	.300
sex	-.152	-3.11	.002

The model accounted for a total of 5.8% of the variance, $F(6,405) = 5.2$, $p < .001$, and showed men outperforming women, over and above a facilitative effect of non-verbal intelligence. A further stepwise regression analysis with all predictors entered at the first step, and sex entered at the second step, showed that sex accounted for a unique 2% of variance, cumulative $F(1,405) = 9.7$, $p < .01$.

To check whether the effect of non-verbal intelligence was present in both sexes, we performed separate multiple regression analyses for men and women. In both analyses, the only significant effect was that of the Culture Fair test (men: $\beta = .16$, $p < .05$; women: $\beta = .17$, $p < .05$). The relationship between non-verbal intelligence and sensitivity to non-native tonal contrasts in men and women is depicted in Figure 2.

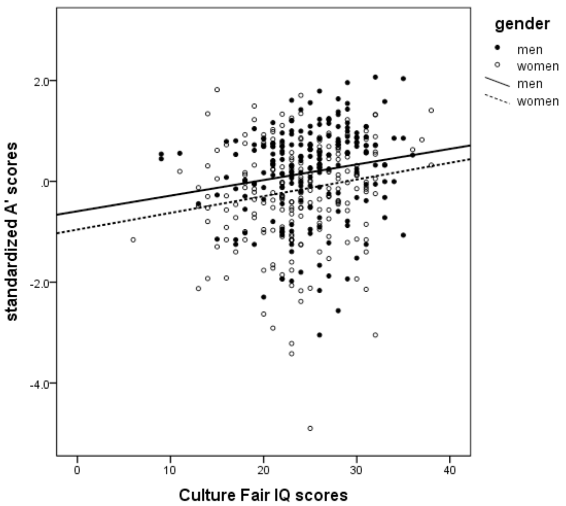

Figure 2: Correlation between Culture Fair scores and sensitivity to Norwegian tonal contrasts for men (black circles, solid line) and women (white circles, dashed line).

Discussion

Our findings demonstrate a small but significant male advantage in non-native toneme discrimination, which cannot be attributed to sex differences in prior exposure to foreign languages or non-verbal intelligence. The lack of a link between prior language exposure and processing of non-native speech sounds is consistent with similar findings from other studies (Golestani & Zatorre, 2009). Note that the effect of non-verbal intelligence on toneme discrimination ability was independent of the effect of sex, confirming the general link between temporal information processing—which is one component of non-native speech sound discrimination—and psychometric intelligence.

The sex effect may seem unexpected because studies that employ non-native speech sound discrimination tasks or temporal auditory processing tasks typically do not compare the performance of men and women. Moreover, given that the sex effect is very small, the sample sizes in such studies are often not large enough for it to be detectable. However, recently, Bowles, Silbert, Jackson & Doughy (2011) reported a male advantage for discrimination of two Hindi consonant contrasts involving differences in Voice Onset Time in a large sample of 1,185 male and 395 female native speakers of American English. This suggests that the male advantage in temporal information processing clearly extends to the processing of difficult non-native speech sounds containing rapid spectral changes.

What mechanisms may be responsible for this sex effect? Golestani et al. (2007) have shown in a perceptual training study that faster learning of non-native speech sounds, involving rapid spectral changes, was associated with differences in brain anatomy. Specifically, faster learning was linked to larger overall white matter volumes in left Heschl's gyrus and increased degree of left > right asymmetry in white matter density in auditory cortex. Increased white matter volume may indicate greater myelination, which would result in more rapid neural transmission crucial for perception of rapid spectral changes. It may also be due to a greater number of white matter fibers connecting language regions within and between cortical hemispheres, for example, connecting the auditory cortex with anterior and posterior language regions. Given that for these perceptual learning tasks performance at the outset was highly correlated with speed of learning (Golestani & Zatorre, 2009) it is reasonable to speculate that similar anatomical changes may also distinguish individuals who perform better in perceptual discrimination tasks when presented with a non-native speech sound for the first time. In addition, white matter volume has been shown to be negatively correlated with variability in isochronous tapping in the sub-second range (Ullén, Forsman, Blom, Karabanov & Madison, 2008) suggesting that white matter volume can be implicated in rapid temporal processing in other domains as well.

Comparisons of male and female brain anatomy and cytoarchitecture have revealed larger white matter to grey matter ratios in men than women, and less white matter asymmetry between hemispheres in women (Gur et al., 1999). Gur et al. (1999) suggested that maintaining grey matter volume consisting of somatodendritic tissue—responsible for computation—at the relative expense of myelinated connective tissue—responsible for information transmission—may be a reasonable evolutionary strategy for dealing with the smaller cranial volumes of females, where transmission occurs over relatively shorter distances than in males.

How can these conjectures about the anatomical substrate responsible for a male advantage in temporal processing be reconciled with findings of (a) somewhat higher verbal abilities in women and (b) the absence of mean sex differences in psychometric intelligence? Meta-analyses (Hyde & Linn, 1988; Lynn & Mikk, 2009) have shown a reliable albeit very small female advantage in verbal abilities, mainly related to speech production and reading and writing abilities. This seems to be at odds with the present finding of a male advantage in temporal processing which extends to the processing of non-native speech sounds. However, there is much more to verbal abilities than the processing of non-native speech sounds, making it unlikely that a sex difference in one capacity will dominate the complex interaction of skills required for the various aspects involved in language learning and processing.

Mean sex differences in general intelligence have generally proven to be elusive (Johnson, Carothers & Deary, 2009) despite sex differences in reaction times (Der & Deary, 2006) and temporal processing (Rammsayer & Troche, 2010), – parameters that have been shown to be predictive of general intelligence (Sheppard & Vernon, 2008). The present study is in agreement with these findings as the trend towards slightly higher Culture Fair non-verbal intelligence scores in men was not significant after Bonferroni correction, despite the fact that Culture Fair scores correlated significantly with sensitivity to the Norwegian tonal contrast. A review of research on sex differences in various timed tests revealed that men are faster on reaction time and finger tapping tests while women are faster in naming and symbol copying and neither sex outperforms the other in general intelligence (Roivainen, 2011). Thus, while reaction time and temporal information processing appear to explain some of the variance in general intelligence, other performance components are also bound to play a role and these components do not necessarily favor men.

It is important to keep in mind that the observed sex effect in non-native speech sound processing was very small. Future research will have to explore to what extent the male advantage in non-native speech sound processing generalizes to other tasks; e.g., identification tasks or AXB-tasks which may be more taxing on working memory or on the ability to form long-term representations of novel speech sounds. It is even less clear whether the male advantage in non-native speech sound processing benefits other aspects of adult foreign language learning, such as morphosyntax or vocabulary acquisition. To clarify these issues, studies of individual differences in various aspects of language learning should include sex as a variable into their analyses.

Despite these limitations, the reported findings underscore the importance of studying sex differences in cognitive tasks as one of the domains that allow researchers to explore potential cognitive repercussions of neuro-anatomical differences.

Acknowledgments

The authors would like to thank Felix Schaeffler for advice in constructing the stimuli and Neil W. Kirk and James Munro for their help in running the experiments. Parts of this study were funded by a Small Grant from the journal *Language Learning*.

References

Bowles, A. R., Silbert, N. H., Jackson, S. R., & Doughy, C. J. (2011). Individual differences in working memory predict second language learning success. Poster presented at the 52nd Annual Meeting of *The Psychonomic Society*, Seattle, WA.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299–2310.

Cattell, R. B., & Cattell, H. E. P. (1973). *Measuring Intelligence with the Culture-Fair Tests*. Champaign, IL: Institute for Personality and Ability Testing.

Der, G., & Deary, I. J. (2006). Reaction time age changes and sex differences in adulthood. Results from a large, population based study: The UK Health and Lifestyle survey. *Psychology and Aging, 21,* 62–73.

Flege, J. E., MacKay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America, 106*, 2973–2987.

Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain & Language, 109*, 55–67.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex, 17*, 575–582.

Goswami, U., Thompson, J., Richardson, U., Stainthorp, R., Highes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proceedings of the National Academy of Sciences USA, 99*, 10911–10916.

Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W., Hughett, P., & Gur, R. E. (1999). Sex differences in brain gray and white matter in healthy young adults: Correlations with cognitive performance, *Journal of Neuroscience, 19,* 4065-4072.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 107,* 139-155.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence*. New York: Springer.

Johnson, W., Carothers, A., & Deary, I. J. (2009). A role for the X chromosome in sex differences in general intelligence? *Perspectives in Psychological Science, 4,* 598-611.

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60-99.

Kempe, V., Brooks, P. J., Marronaro, R., & Thoresen, J. (2010) Individual differences in the perception of a non-native tonal contrast. Poster presented at the Workshop on *Psycholinguistic Approaches to Speech Recognition in Adverse Conditions*. Bristol, UK.

Kempe, V., Thoresen, J. C., & Brooks, P. J. (2007). Differences in foreign language phoneme perception and production. Poster presented at the 48th Annual Meeting of *The Psychonomic Society*, Long Beach, USA.

Kempe, V., Thoresen, J. C., & Brooks, P. J. (2008). Norwegian toneme perception by non-native speakers. Poster presented at the 49th Annual Meeting of *The Psychonomic Society*, Chicago, IL.

Kempe,V., Thoresen, J. C., Kirk, N. W., Brooks, P. J., & Schaeffler, F. (2011) Perceptual and cognitive predictors of non-native phoneme discrimination. Poster presented at the 52nd Annual Meeting of *The Psychonomic Society*, Seattle, WA.

Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Trames, 13,* 3–13.

McRoberts, G. & Sanders, B. (1992). Sex differences in performance and hemispheric organization for a nonverbal auditory task. *Perception and Psychophysics, 51,* 118-122.

Rammsayer, T., & Troche, S. (2010). Sex differences in the processing of temporal information in the sub-second range. *Personality and Individual Differences, 49,* 923-927.

Rammsayer, T. H., & Brandler, S. (2007). Performance on temporal information processing as an index of general intelligence. *Intelligence, 35,* 123–139.

Roivainen, E. (2011). Gender differences in processing speed: A review of recent research. *Learning and Individual differences, 21*, 145-149.

Rostad, K., Mayer, A., Fung, T. S., & Brown, L. N. (2007). Sex-related differences in the correlations for tactile temporal thresholds, interhemispheric transfer times, and nonverbal intelligence. *Personality and Individual Differences, 43,* 1733–1743.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information processing: A review of 50 years of research. *Personality and Individual Differences, 44,* 535-551.

Szelag, E., Szymaszek, A., Aksamit-Ramotowska, A., Fink, M., Ulbrich, P., Wittmann, M., & Pöppel, E. (2011). Temporal processing as a base of language universals: Cross-linguistic comparisons on sequencing ability with some implications for language therapy. *Restorative Neurology and Neuroscience, 29,* 35-45.

Talcott, J. B., Witton, C., McClean, M., Hansen, P. C., Rees, A., Green, G. G. A., & Stein, J. F. (2000). Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Sciences USA, 97,* 2952–2962.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*, 182-198.

Troche, S. J., & Rammsayer, T. H. (2009). The influence of temporal resolution power and working memory capacity on psychometric intelligence. *Intelligence, 37,* 479–486.

Ullén, F., Forsman, L., Blom, Ö., Karabanov, A., & Madison, G. (2008). Intelligence and variability in a simple timing task share neural substrates in the prefrontal white matter. *Journal of Neuroscience, 28,* 4238–4243.

van Raavenzwaaj, D., Brown, S., & Wagenmakers, E.-J. (2011). An integrated perspective on the relation between response speed and intelligence. *Cognition, 119,* 381-393.

Wittmann, M., & Szelag, E. (2003). Sex differences in perception of temporal order. *Perceptual and Motor Skills, 96,* 105–112.