# Stimulus, Task, and Learning Effects on Measures of Temporal Resolution: Implications for Predictors of Language Outcome

**Nicholas A. Smith**
**Laurel J. Trainor**
**Kellie Gray**
**Judy A. Plantinga**
**David I. Shore**
McMaster University, Hamilton, Ontario

**Purpose:** Some studies find that temporal processing ability predicts language outcome whereas other studies do not. Resolution of this debate is hindered by the variety of temporal measures used, nonsensory loading of the tasks, and differential amounts of practice across studies. The goal of this study was to examine the effects of stimulus properties, experimental task, and perceptual learning on listeners' gap detection performance.
**Method:** Gap detection thresholds were obtained from adults with normal hearing and language ability. The effects of marker frequency similarity and marker duration on thresholds were examined in yes-no, two-interval forced-choice (2IFC), and dual-pair comparison tasks (which vary in nonsensory loading) over 4 days of testing.
**Results:** Thresholds were highest for gaps defined by markers with disparate frequencies (1000 and 4000 Hz; i.e., between-channel gap detection), and with longer (300 ms) trailing markers, obtained using yes-no and 2IFC tasks. However, these effects were attenuated with training or the initial use of the dual-pair comparison task.
**Conclusions:** These results suggest that gap detection thresholds reflect a variety of sensory and nonsensory factors. Understanding these underlying factors is critical to any evaluation of the relation between temporal processing and language outcome.

**KEY WORDS: temporal resolution, language, perceptual learning, backward masking, gap detection**

Speech and other complex auditory patterns contain many fine temporal details that listeners must resolve for accurate encoding and eventual understanding. *Gap detection thresholds*—the shortest detectable interruption between two sounds—represent one common measure of temporal resolution. We examined how listeners' gap detection thresholds were affected by (a) stimulus factors, such as marker frequency and duration; (b) nonsensory demands of the task; and (c) perceptual learning.

Beyond addressing important basic questions of auditory temporal processing, the present study is relevant to applied research that examines the potential role of basic sensory processes in deficits of reading and language. The *temporal processing hypothesis* (also referred to as the *rate-processing constraint hypothesis*) states that language outcome can be predicted by measures of auditory temporal processing (e.g., Benasich & Tallal, 2002; Tallal & Gaab, 2006; Tallal et al. 1996; Wright, Lombardino, et al., 1997). This hypothesis has considerable prima facie plausibility given that the relative timing and sequence of phonetic information are essential aspects of speech. Gap detection in particular may bear some relation to those processes involved in the processing of speech (Phillips, 1999). For example, a temporal gap in the onset of voicing provides a critical distinction between voiced and unvoiced stop consonants (e.g., /b/ versus /p/). Gaps between words can also be important. For example, increasing the duration of the silent interval between the words *gray* and *chip* can cause listeners to perceive *great chip* (Repp, Liberman, Eccardt, & Pesetsky, 1978). Temporal processing has also been linked to reading ability in that dyslexics, who show profound difficulties in reading, also show deficits in tasks requiring the processing of rapid visual and auditory stimuli.

However, the relation between temporal processing and language deficits is not clear-cut. For example, a number of studies have failed to find correlations between reading measures and thresholds for the detection of gaps in noise (Ahissar, Protopapas, Reid, & Merzenich, 2000; Breier, Fletcher, Foorman, Klaas, & Gray, 2003; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1998). On the other hand, links have been found between language ability and backward masking (Ahissar et al, 2000; Marler & Champlin, 2005; Wright, Lombardino, et al., 1997), temporal order judgments (Au & Lovegrove, 2001), and frequency discrimination (McAnally & Stein, 1996).

One contributing factor to the complexity of the debate concerns the methodological variability across studies. Reviews of the evidence acknowledge conflicting conclusions (Habib, 2000; McArthur & Bishop, 2001). Nevertheless, Farmer and Klein (1995, p. 485) argue that the existing data "constitute sufficient evidence to warrant further investigation of the link between temporal processing deficits and dyslexia." In their review, McArthur and Bishop (2001) identified the need to assess the reliability and validity of the psychophysical tasks used in individual subjects and to assess the variability in measures of an individual's performance in the task across test sessions. It is an essential prerequisite that any study of the predictive relation between temporal processing and language outcome make use of psychophysical measures that reliably predict individual subjects' performance across test sessions as well as predict performance on other tasks that aim to measure the same underlying sensory and perceptual abilities (the principle of converging operations; see, e.g., Benasich, Thomas, Choudhury, & Leppänen, 2002). A measure (i.e., of temporal processing) cannot be expected to predict something else (i.e., language outcome) if it cannot first predict itself.

Despite many studies supporting and refuting the temporal processing hypothesis, very few studies have begun with a rigorous examination of the reliability and validity of their psychophysical measures. Typically listeners are tested once in a temporal processing task—usually with stimuli that are novel to the listener—and the correlation between performance on this task and performance on language tasks is evaluated. In the present study, we systematically examined the effects of spectral and temporal stimulus characteristics, nonsensory task demands, and perceptual learning. This approach permits a better understanding of the role that different stimulus and psychological factors play in gap detection and of how these factors change with perceptual learning across test sessions.

## Stimulus Factors

*Marker frequency.* The auditory system performs a spectral analysis on incoming sounds and encodes them in terms of their component frequencies. This is achieved by means of a tonotopically organized basilar membrane which projects to the auditory cortex via a collection of channels in the auditory periphery that are tuned to different frequencies. The way in which the auditory system detects gaps in a stimulus is dependent on the spectral properties of the gap markers themselves. If the leading and trailing markers have the same or similar frequency content, the gap (specifically, the offset of the leading marker and the subsequent onset of the trailing marker) can be encoded within a common set of frequency channels (within-channel gap detection). However, if the leading and trailing markers have different frequency content then each marker is encoded by a different, nonoverlapping set of frequency channels (between-channel gap detection). As a consequence, the gap must be detected using more central auditory mechanisms that integrate information across multiple frequency channels. A number of studies have shown dramatically increased thresholds for between-channel, relative to within-channel, thresholds (Fitzgibbons, Pollatsek, & Thomas, 1974; Formby, Gerber, Sherlock, & Magder, 1998; Formby, Sherlock, & Li, 1998; Grose, Hall, Buss, & Hatch, 2001; Phillips & Hall, 2000; Phillips, Taylor, Hall, Carr, & Mossop, 1997; Taylor, Hall, Boehnke, & Phillips, 1999). Furthermore, Phillips and Smith (2004) found that listeners' between- and within-channel gap detection thresholds are only weakly correlated, suggesting that different auditory processing mechanisms underlie performance in the two tasks. Besides being a putative measure of more central auditory processing, Phillips (1999) has argued that between-channel gap detection tasks may provide a better model of voice onset time (VOT) perception than within-channel tasks because the sounds before and after the delay are spectrally dissimilar and activate different perceptual channels. In the present study, marker frequency was manipulated with the goal of examining the differential effects of perceptual learning and task demands on central (between-channel gap detection) and peripheral (within-channel gap detection) processes.

### Marker Duration

Gap detection performance depends both on the duration of the silent gap and also on the properties of the sound markers used to generate the gap. The presence of gaps may be obscured by leading markers due to the effects of forward-masking and adaptation (Penner, 1977; Wiegrebe & Krumbholz, 1999), and by trailing markers due to the effects of backward masking. One important reason for examining the role of backward masking in gap detection is an observed relation between excessive backward masking and language impairments (Hartley & Moore, 2002; McArthur & Hogben, 2000; Wright, Lombardino, et al., 1997).

We tested short (10 ms) and long (300 ms) trailing markers to address the following questions. First, do longer trailing markers increase thresholds through increased backward masking? Previous research provides mixed reports of the effects of duration. For example, whereas Elliott (1964) reported no effects of noise masker duration on tone detection, the results of others (Penner, 1974; Penner, Cudahy, & Jenkins, 1974) do suggest some effects of masker duration on backward masking of clicks, particularly when the masker follows the target within a critical masking interval of 10 ms or less. In any case, the potential masking of gaps by longer duration trailing markers is particularly relevant to speech processing given that vowels following VOT gaps that define consonants in speech are typically closer to 300 ms than 10 ms.

The second question examined here is whether backward masking is affected by frequency differences between leading and trailing markers. As is the case with duration effects in backward masking, reports of frequency effects are mixed. Backward masking of tones by narrow-band noise is greatest when the probe tone frequency corresponds to the frequency of the noise band (Elliott, 1967; Wright, 1964). However, using tone maskers, Miyazaki and Sasaki (1984) failed to find effects of masker frequency on backward masking of tones.

Finally, what role do nonsensory factors (task demands and perceptual learning) play in backward masking of gaps? In this vein, previous studies have found that the effects of backward masking can be reduced with training (Bishop, Carlyon, Deeks, & Bishop, 1999; Buss, Hall, Grose, & Dev, 1999; Kallman & Brown, 1986; Marler, Champlin, & Gillam, 2001; Roth, Kishon-Rabin, & Hildesheimer, 2001) and that backward-masking thresholds are related to cognitive factors, such as IQ (Hartley, Wright, Hogan, & Moore, 2000).

## *Experimental Task: Nonsensory Factors*

The ability to discriminate gap and no-gap stimuli was tested using three different experimental tasks. In the yes-no task, listeners were presented with a single stimulus, or interval, and were asked to decide whether the stimulus contained or did not contain a gap. In the two-interval forced-choice (2IFC) task, both the gap and the no-gap stimuli were presented on a given trial (in random order), after which listeners decided which of the two intervals contained the gap. In the dual-pair comparison task (also referred to as a *4IAX task*), stimuli were arranged into two pairs, one of which had two identical no-gap stimuli and the other of which contained one gap stimulus and one no-gap stimulus. In this case, the listeners' task was to decide which of the pairs was different. It is important to note that the dual-pair comparison task is not a four-alternative forced-choice (4AFC) task in which chance performance would be equal to 25%, but rather it was a task in which each of the two intervals contained a pair of stimuli.

The present study compares thresholds (at a fixed percent correct) across tasks. This measure was adopted primarily because it was straightforward to implement in an adaptive staircase procedure, which permitted the efficient testing of multiple experimental factors across multiple test sessions. In terms of percent correct, comparisons across tasks are both straightforward and problematic. Although the yes-no, 2IFC, and dual-pair comparison tasks are equated in terms of chance performance (50%), they each have different underlying $d'$s. In terms of detection-theoretic representations of the decision space underlying each task, percent correct for a given $d'$ in a 2IFC task exceeds that observed in a yes-no task, and a yes-no task is, in turn, higher than that observed in a dual-pair comparison task (Macmillan, Kaplan, & Creelman, 1977; Micheyl & Messing, 2006). However, comparison across tasks is made more complicated by the fact that numerous empirical tests have failed to find the relations predicted by theory (Creelman & Macmillan, 1979; Jesteadt & Bilger, 1974), suggesting the additional role of other factors such as interstimulus timing, stimulus range, and perceptual memory that are not adequately captured by the models (Macmillan & Creelman, 2005). Although a detection-theoretic comparison of these tasks would be desirable in many ways, these analyses typically involve large numbers of trials performed by highly practiced listeners and would therefore preclude the present examination of early perceptual learning occurring over a few relatively short test sessions.

The present comparison of performance in yes-no, 2IFC, and dual-pair comparison tasks is interesting because in all cases, the basic stimulus processing is equated. That is, the stimulus differences to be detected between gap and no-gap stimuli are the same. What differs across tasks is the cognitive processing required for correct performance. In the yes-no task, listeners must identify the single stimulus, using their memory of the gap and no-gap stimuli. In the 2IFC task, listeners need only to recognize the gap stimulus relative to the no-gap stimulus presented within that trial. Finally, in the dual-pair comparison task, listeners need not even identify or recognize the gap but simply determine whether the stimuli in each pair are the same or different. It is therefore reasonable to expect that as the task progresses from a yes-no to a 2IFC to a dual-pair comparison task, the cognitive demands placed on the listener decrease. Lending credence to this assumption, Bishop and colleagues (1999) found that children's backward-masking thresholds were higher when tested with a two-alternative forced-choice task, in which children had to decide which of the two intervals contained the tone, than with a three-interval forced-choice task, in which children were required to decide which of the three intervals was different than the other two (though this study did not control for differences in chance performance between tasks). Alternatively, on the basis of tests of reminder paradigms (reviewed by Macmillan & Creelman, 2005, pp. 180–183), one might predict lower performance in the dual-pair comparison task.

The primary value of the task manipulation is evident when one considers the potential predictive relation between gap-detection performance and reading ability. Do the previously reported correlations between temporal processing and reading reflect an assessment of auditory temporal resolution, or do cognitive—task-related—factors mediate this relation? Examining the joint contribution of sensory processes, which are more or less constant across tasks, and cognitive processes, which vary across tasks, will start to unravel this interaction.

## Perceptual Learning

Experience and learning improve the ability to respond to stimuli of all types (Gibson, 1969; Goldstone, 1998). Although many studies seek to avoid or control the effects of learning or practice (i.e., within-subjects perceptual studies), perceptual learning itself has been the focus of many auditory studies. With training, listeners show improved frequency discrimination (Irvine, Martin, Klimkeit, & Smith, 2000), auditory pattern processing (Leek & Watson, 1984), and sensitivity to nonnative speech contrasts (Logan, Lively, & Pisoni, 1991; Wang, Spence, Jongman, & Sereno, 1999). In fact, the potential for perceptual learning to occur is germane to attempts (that stem from the temporal processing hypothesis) to remediate language deficits through training in temporal processing tasks (Friel-Patti, Loeb, & Gillam, 2001; Gillam, Crofford, Gale, & Hoffman, 2001; Tallal et al., 1996). In support of this, auditory training has been shown to improve children's phonological processing (Merzenich et al., 1996; Moore, Rosenberg, & Coleman, 2005).

Perceptual learning has implications for tests of the relation between perceptual measures and language outcome. Initial performance levels may be quite different than performance levels on a practiced task, and individuals may differ in the rate at which performance improves (Karmarkar & Buonomano, 2003; Watson, Kelly, & Wroton, 1976; Wright, Buonomano, Mahncke, & Merzenich, 1997). Some kinds of tasks may be more amenable to perceptual learning than others.

The present study measured thresholds on a variety of gap detection tasks over 4 days of testing. Previous studies have examined perceptual learning in temporal processing tasks. Wright, Buonomano, et al. (1997) demonstrated improvement in temporal-interval discrimination over 10 days of testing, in a task in which listeners were asked to discriminate comparison intervals from a standard interval of 100 ms defined by short 1-kHz tone pips. Overall difference thresholds improved with training, though for some listeners they improved more than others. Furthermore, Wright, Buonomano, et al. found that learning generalized to 100-ms standard intervals marked by a different frequency (4 kHz), but not to intervals of different lengths defined by the same frequency as that used in training. Some degree of temporal specificity of learning has also been found for the discrimination of intervals defined by vibratory pulses on the skin, despite generalization to untrained skin locations and to auditorily defined intervals (Nagarajan, Blake, Wright, Byl, & Merzenich, 1998).

## Scope of the Present Study

The present study tested performance in three related gap detection tasks (yes-no, 2IFC, and dual-pair comparison) across 4 days of testing with two primary stimulus manipulations—marker duration (10 ms or 300 ms) and marker frequency content (1000–1150 Hz or 1000–4000 Hz). A number of predictions were made.

1. Thresholds would decrease with practice, both over the course of 4 days of testing and within a test session, the second threshold obtained being lower than the first.
2. Thresholds would be higher for 1000–4000 Hz markers, than for 1000–1150 Hz markers, which would be consistent with previous tests of between-channel gap detection.
3. Thresholds would be higher for 300 ms trailing markers than for the 10 ms trailing markers.
4. Thresholds would be higher in the yes-no tasks than in the 2IFC and dual-pair comparison tasks, due to the increased cognitive load.

By testing gap detection under a variety of conditions, in this study we aimed to assess both the sensory and cognitive aspects of temporal resolution. Any such interaction would underscore the importance of considering nonsensory factors when examining the relation between auditory temporal processing and language ability.

# Methods

## Participants

Eighteen adult undergraduate students between 18 and 23 years of age were tested, all of whom had normal audiometric thresholds and reported no language or reading problems. Half of the listeners ($n$ = 9) were assigned to the 10–10 ms marker duration condition, and the other half ($n$ = 9) to the 10–300 ms marker duration condition (described subsequently). Participants were paid $50 for completion of all four test sessions, within the span of 1 week.

## Stimuli

Listeners' gap detection thresholds were tested using gap and no-gap stimuli, both of which consisted of leading and trailing tone markers. In the gap stimuli, the duration of the gap between the leading and trailing markers varied adaptively between 1 ms and 120 ms from trial to trial. In the no-gap stimuli, the trailing marker immediately followed the leading marker.

The use of sine tones in gap detection tasks poses problems not present in other kinds of gap detection tasks due to the introduction of a number of confounding factors (see Formby & Forrest, 1991). These include the creation of *spectral splatter*—energy at a range of frequencies other than the nominal frequency—caused by the abrupt onsets and offsets of sine tones. Furthermore, the use of any kind of discrete marker (as opposed to gaps in continuous noise) introduces the potential for confounds between the gap duration and the total duration of the stimulus when the markers have fixed durations. These factors are problematic because they provide alternative potential cues (i.e., other than gap duration) to the presence of a gap between tone markers.

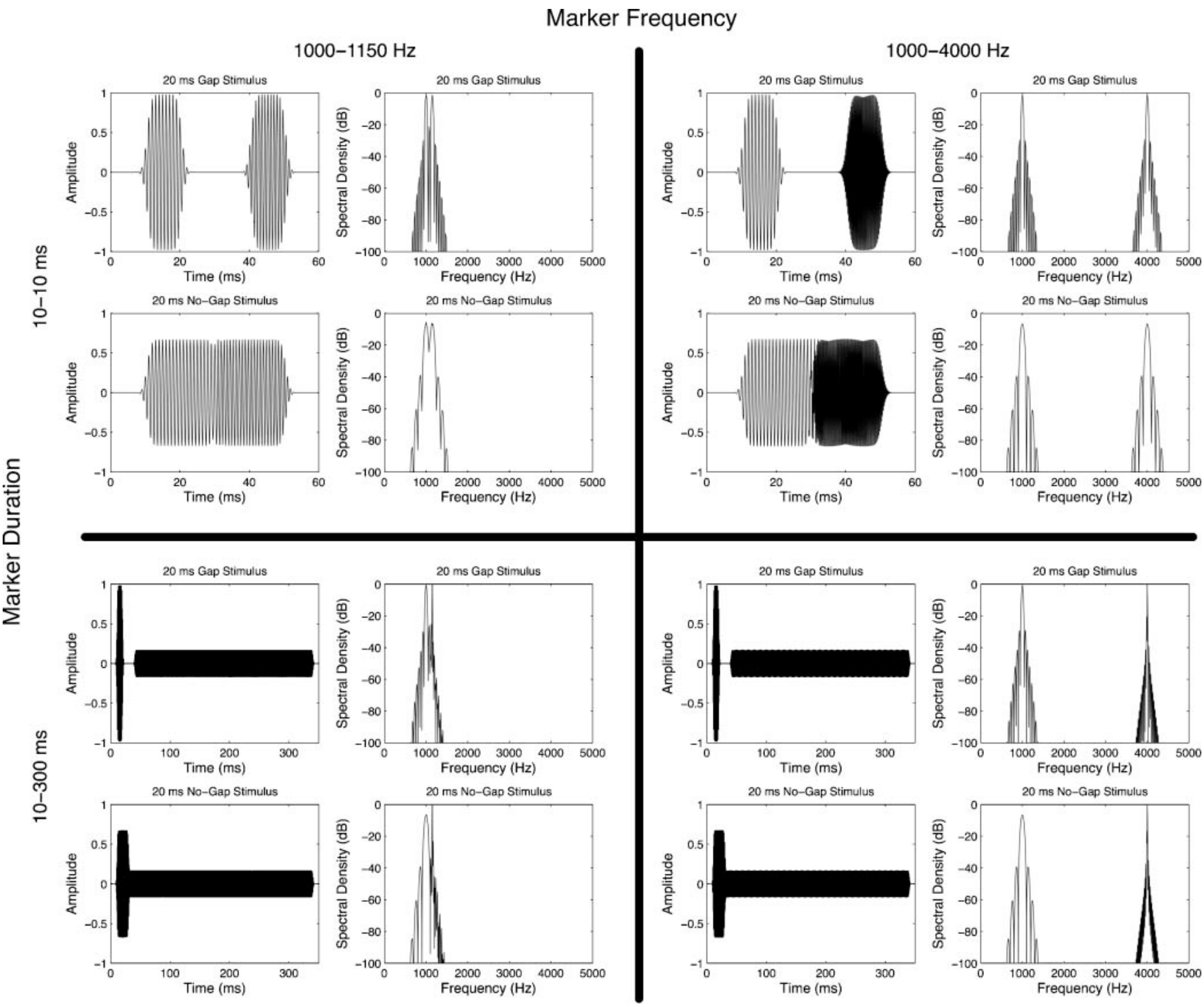
To control for these confounds, gap and no-gap stimuli were created as follows. Each marker consisted of a single pure tone, temporally enveloped by a series of overlapping Gaussian windows using a method adapted from Schneider, Pichora-Fuller, Kowalchuk, and Lamb (1994). Each Gaussian window had a standard deviation of 1 ms, and the temporal offset between peaks of successive windows was 1 ms. Consequently, the duration of each marker was defined as the number of Gaussian windows in the series. Gap size was defined in terms of the number of 1-ms Gaussian windows between the last window of the first marker and the first window of the second marker. Schneider et al. demonstrated that when tones are enveloped in this way, the spectral density functions of gap and no-gap stimuli are similar in terms of the range and profile, thereby reducing the potentially confounding role of spectral factors in gap detection. Furthermore, when Schneider et al. reduced the stimulus intensity, bringing remote frequency regions under the threshold of audibility, they failed to find decreases in performance that would be expected if listeners' gap detection performance were driven by the spectral cues. As in Schneider et al., spectral density functions for gap and no-gap stimuli (shown in Figure 1) show similar ranges and profiles.

A number of different kinds of stimuli were created using the general method just described. Two different trailing marker durations were tested. In the 10-10 ms condition, the leading and trailing markers were each 10 ms in duration. In the 10–300 ms condition, the leading marker was 10 ms long, and the trailing marker was 300 ms long. The 10- and 300-ms-long markers were created by using series of either 10 or 300 Gaussian windows.

The effect of marker frequency separation was also tested. In the 1000–1150 Hz condition, the leading and trailing markers were 1000 and 1150 Hz, respectively. In the 1000–4000 Hz condition, the leading and trailing markers were 1000 and 4000 Hz, respectively. The 1000-Hz marker was always in the leading position.

One unavoidable consequence of adding gaps of various durations between two markers of fixed durations is that gap duration becomes perfectly correlated with the total duration of the stimulus. As a result, listeners could potentially perform the task by identifying the longer stimulus as the one containing the gap without resolving the temporal gap between markers. To control for this confound, each gap stimulus had an associated no-gap stimulus of equal total duration. In effect, this was done by lengthening the durations of the markers of the leading and trailing markers in the no-gap stimuli. However, this alone would introduce an additional complication because lengthening the markers in the no-gap stimuli would increase their total power, or perceived intensity, relative to the their associated gap stimuli. Listeners could perform the task by identifying the less intense stimulus as the gap stimulus without having resolved the actual gap between markers. To control for this, gap and no-gap stimuli were equated for total power. One additional potential confound remains: that of the relative intensity of the leading and trailing markers in the 10–300 ms condition. There is a greater intensity difference between the leading and trailing markers

**Figure 1.** Waveforms and spectral analyses for gap and no-gap stimuli (20 ms given here as an example) shown with marker frequencies of 1000–1150 Hz and 1000–4000 Hz and marker durations of 10–10 ms and 10–300 ms. Gap and no-gap versions of each stimulus pair are equated for total power and duration and have similar spectra.




in the gap stimuli than there is between leading and trailing markers in the no-gap stimuli.

## Procedure

Listeners were seated in a single-walled, sound-attenuated booth. Stimuli were presented through Sennheiser HDA 200 headphones at 62 dB. Listeners' gap detection thresholds were obtained using an adaptive 2-down, 1-up staircase procedure (Levitt, 1971). This procedure adjusts the gap duration from trial to trial, on the basis of listeners' correct/incorrect responses, in order to find the gap duration corresponding to threshold performance level of 71%. The staircase terminated after eight reversals, and the geometric average gap duration over the last six reversals was taken as the threshold. Gap detection thresholds were obtained for both within- and between-channel stimuli using three different experimental tasks: yes-no, 2IFC, and dual-pair comparison.

In the yes-no task, listeners were presented with a single gap or no-gap stimulus and were asked to identify the stimulus as a gap or no-gap stimulus, presented with equal probabilities. In the 2IFC task, listeners were presented with a pair of stimuli (one gap stimulus and one no-gap stimulus in a random order) and were asked to decide which of the two stimuli contained the gap. The interstimulus interval (ISI) was 500 ms. In the dual-pair comparison task, listeners were presented with two pairs of stimuli. One pair contained two no-gap stimuli, and the other pair contained one gap and one no-gap stimulus. The ISI between stimuli within each pair was 500 ms, and the ISI between pairs was 1,000 ms. Listeners were asked to judge which of the two stimulus pairs contained different stimuli. No feedback was given.

Listeners were assigned to either the 10–10 ms or 10–300 ms marker duration condition, and tested on 4 different days over the course of a week. On each day, two thresholds were obtained for each marker frequency condition (1000–1150 Hz and 1000–4000 Hz) and each of the three experimental tasks (yes-no, 2IFC, and dual-pair comparison). The order of conditions tested on each day was counterbalanced across listeners, and for each listener, the task was counterbalanced across days. The decision to test the effect of marker duration between subjects was made for practical reasons: to limit the duration of a testing session to 1 hr and to avoid excessively complex counterbalancing.

## Results

Two gap detection thresholds were obtained for each listener in each experimental condition on each day. The averages of these two thresholds, averaged across listeners for marker duration (10–10 ms and 10–300 ms), marker frequency (1000–1150 Hz and 1000–4000 Hz), and task (yes-no, 2IFC, and dual-pair comparison), are plotted as a function of testing day (1, 2, 3 and 4) in Figure 2. A preliminary analysis of the data revealed that the variances were not homogenous across levels of the experimental factors. In particular, thresholds that were higher on average tended to be more variable. In order to correct for this, for the purpose of adhering to the assumptions of analysis of variance (ANOVA), the data were log (base 10) transformed. Furthermore, a number of violations of the sphericity assumption were detected and remedied using Greenhouse-Geisser corrections. In all cases, the corrected analyses replicated those in which sphericity was assumed, and so, for the sake of simplicity of presentation, the uncorrected analyses are presented in the following paragraphs.

In comparing thresholds across tasks, response bias is a potential concern, particularly for the yes-no task in which in the gap duration can vary over a sequence of trials in which only one of the stimulus type (i.e., gap or no gap) is presented. An analysis revealed a neutral bias overall: 48.9% “gap” responses in the yes-no task; 50.8% “second interval contains gap” responses in the 2IFC task; and 45.7% “second pair of stimuli are the same” responses in the dual-pair comparison task.

The transformed thresholds were submitted to a five-way repeated measures ANOVA examining the between-subjects factor of marker duration (10–10 ms vs. 10–300 ms) and the within-subjects factors of marker frequency (1000–1150 Hz and 1000–4000 Hz), experimental task (yes-no, 2IFC, and dual-pair comparison), testing day (1 to 4), and replication (first or second threshold obtained).

Significant main effects were found for all factors tested. As expected, there was a significant main effect of testing day, $F(3, 48) = 21.03$, $p < .001$, with thresholds showing a significant linear trend, $F(1, 16) = 31.69$, $p < .001$, across days. No significant higher order trends were observed.

As expected, a significant main effect of marker frequency was found, $F(1, 16) = 93.70$, $p < .001$, with higher thresholds for 1000–4000 Hz markers than for 1000–1150 Hz markers. A significant main effect of task was found, $F(2, 32) = 22.28$, $p < .001$, with higher thresholds in the yes-no and 2IFC tasks than in the dual-pair comparison task. A significant main effect of trailing marker duration was found, $F(1, 16) = 11.04$, $p < .01$, with higher thresholds for the 10–300 ms group than for the 10–10 ms group. Finally, a significant effect of repetition was found, $F(1, 16) = 17.58$, $p < .001$, with the second threshold obtained on a given day being lower than the first.

A number of significant interaction effects were also found. The two-way interaction of Marker Duration × Task, $F(2, 32) = 9.36$, $p < .01$, illustrates that the negative effect of longer (300 ms) trailing markers on gap thresholds was strongest in the yes-no task ($p < .001$; Bonferroni adjustments for all multiple comparisons) and 2IFC task ($p < .005$) and could be alleviated by using a less cognitively demanding dual-pair comparison task (*ns*).

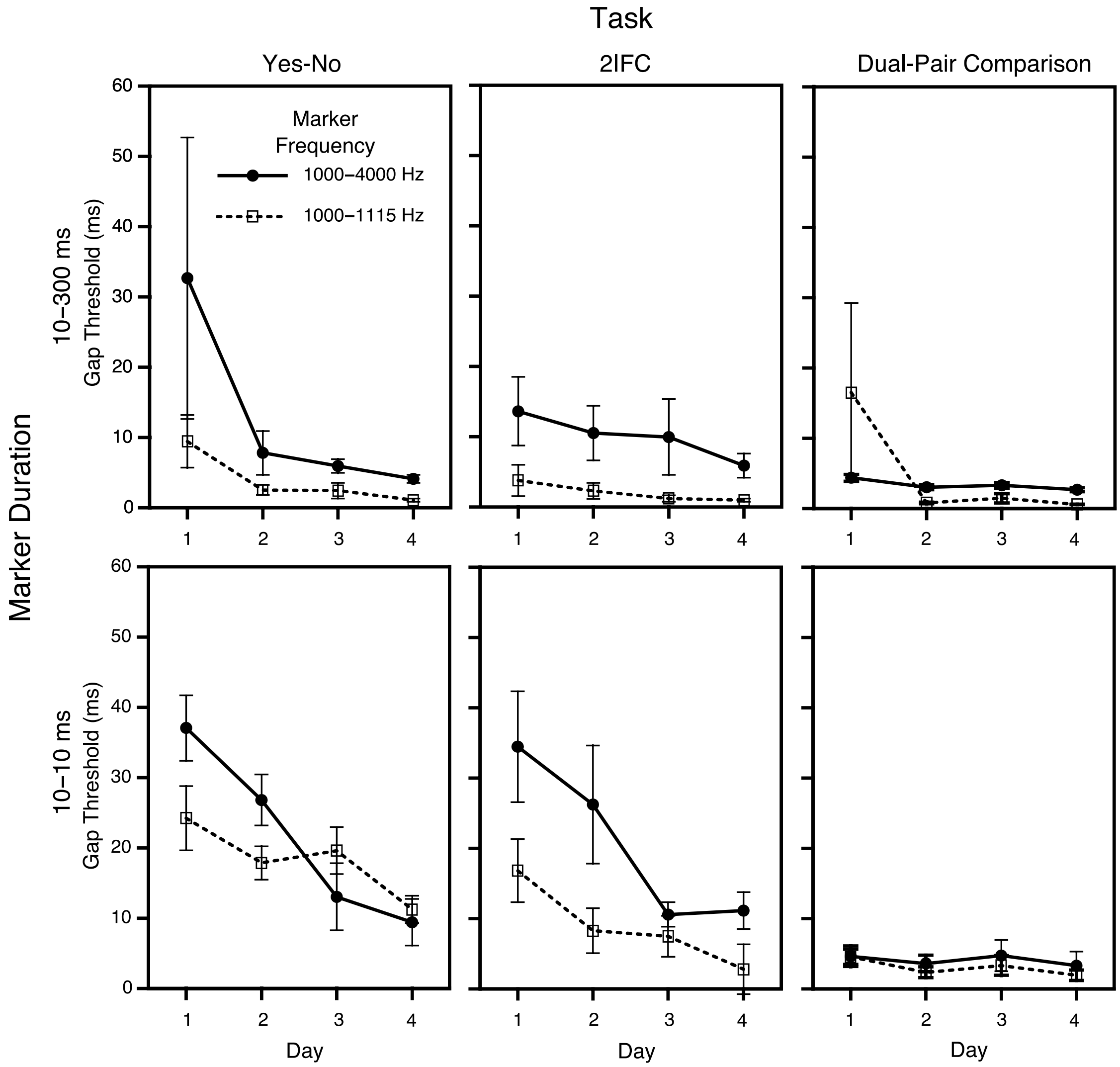
A significant two-way interaction of Marker Frequency × Task, $F(2, 32) = 5.48$, $p < .01$, was found. Thresholds were significantly higher for the 1000–4000 Hz markers than for the 1000–1150 Hz markers in all tasks (all *p*s < .001); the magnitude of the difference was slightly greater for the 2IFC task.

A significant two-way interaction of Marker Frequency × Marker Duration was found, $F(1, 16) = 20.96$, $p < .001$. This interaction reflects an increased effect of marker frequency separation for the 10–10 ms group ($p < .001$), than for the 10–300 ms group ($p < .01$).

A significant two-way interaction of Task × Marker Duration was found, $F(2, 32) = 9.36$, $p < .001$. The increase in thresholds associated with longer (i.e., 300 ms) trailing markers was greatest for the yes-no task ($p < .001$), followed by the 2IFC task ($p < .005$), and was smallest for the dual-pair comparison task (*ns*). All other interaction effects were nonsignificant.

Although the repetition factor did not interact significantly with other factors, which would provide evidence for varying degrees of reliability across conditions,

**Figure 2.** Mean gap detection thresholds as a function of marker frequency condition (1000–1150 Hz and 1000–4000 Hz), marker duration condition (10–10 ms and 10–300 ms), experimental task (yes-no, 2IFC, and dual-pair comparison), and test day. Standard error bars (within subjects) are shown. 2IFC = two-interval forced-choice.



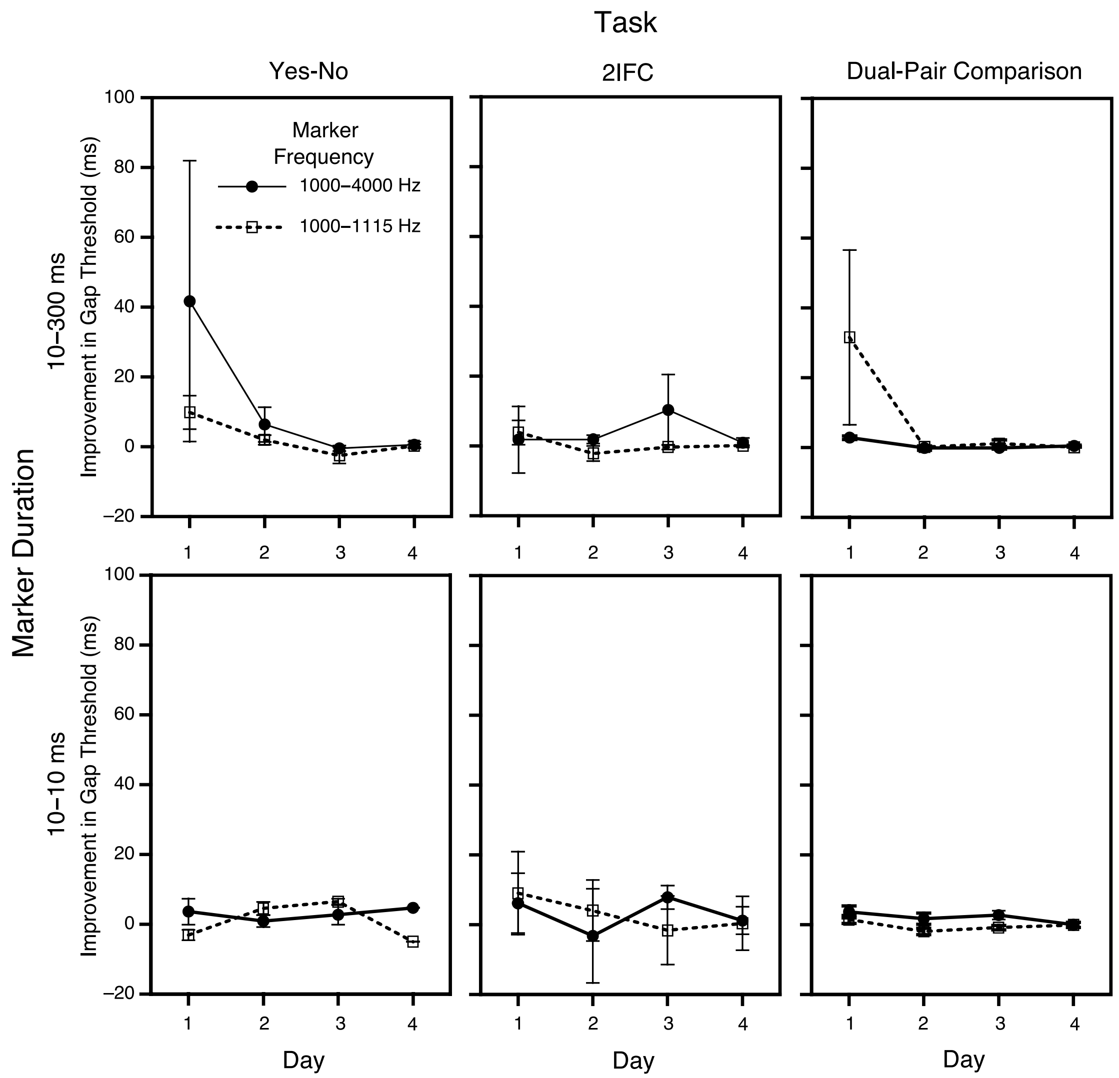

it is nevertheless worthwhile to examine differences between the first and second threshold obtained, as a function of other factors. As is shown in Figure 3, change in threshold is most often in the direction of improvement (reflecting the main effect of repetition discussed previously). Also, the few examples of large within-session improvement generally occurred on Day 1, but the degree of this improvement varied widely across listeners. However, in most cases thresholds obtained on each repetition were reliable to within a few milliseconds of each other.

## Discussion

Listeners' ability to detect gaps varies considerably, both between and within listeners as a function of the spectral and temporal properties of the markers, the experimental task, and practice. These effects demonstrate the different roles of sensory and cognitive processes engaged in temporal processing in different contexts. Many of these factors have been shown individually in previous studies, of which the present results provide a consistent extension. However, the present study shows that these factors interact within individual listeners. These interactions have important implications for the use of gap detection thresholds as a potential predictor of language outcome.

Listeners demonstrated significant perceptual learning. Overall, their gap thresholds decreased over four days of testing. This is not surprising, given that perceptual learning is a ubiquitous phenomenon found

**Figure 3.** Mean within-session improvement in gap detection thresholds as a function of marker frequency condition (1000–1150 Hz and 1000–4000 Hz), marker duration condition (10–10 ms and 10–300 ms), experimental task (yes-no, 2IFC, and dual-pair comparison), and test day. Standard error bars are shown.




across domains. However, it is more interesting that the amount of perceptual learning was greater for some conditions than for others. In particular, the greatest practice-related improvement was found for the yes-no and 2IFC tasks, with spectrally dissimilar (1000–4000 Hz) markers and with longer (300 ms) trailing marker durations. In other words, perceptual learning was greatest in the conditions in which initial performance was worse.

Unpracticed performance on the dual-pair comparison task was near ceiling (or at least approached the smallest gap possible with our temporal enveloping techniques). This highlights the fact that the perceptual/cognitive mechanisms (i.e., those beyond basic temporal resolution) involved in the yes-no and 2IFC tasks are amenable to learning. The important implication is that unpracticed performance on such tasks may not provide an accurate index of temporal resolution per se.

In terms of individual factors tested, we found that listeners' thresholds were significantly higher for gaps defined by the 1000–4000 Hz markers than the 1000–1150 Hz markers. This is consistent with previous research that has shown that gap thresholds increase as a function of frequency difference between markers (Fitzgibbons et al., 1974; Formby, Gerber, et al., 1998; Formby, Sherlock, & Li, 1998; Grose et al., 2001; Phillips & Hall, 2000; Taylor et al., 1999). Furthermore, our observed marker frequency effects are also consistent with the view that more central auditory processing mechanisms are engaged when listeners must integrate information across disparate frequency channels (Phillips et al., 1997). However, beyond confirming the results of previous studies, by testing gap detection over 4 days of testing and with different tasks, we show that these putative central processes are capable of perceptual learning and benefit from the additional contextual information provided by 2IFC and dual-pair comparison tasks. This contextual information, namely examples of gap and no-gap stimuli presented within a given trial, eliminates the need for listeners to identify stimuli on the basis of long-term memory traces and allows them to make relative judgments. This is shown in the dramatic reduction in the effect of large frequency separation (in the 1000–4000 Hz condition) over successive days of testing as well as through the use of 2IFC and dual-pair comparison tasks, even on the first day of testing.

The effects of learning and task also interacted with those of trailing marker duration in similar ways. Thresholds were significantly higher with 300 ms trailing markers than with 10 ms trailing markers, suggesting a duration effect in backward masking. Alternatively, an effect of marker duration similarity is also possible—that is, when the two markers have the same duration (10–10 ms), some processes affected by similarity may increase gap thresholds. In addition, even during the first test session, the effect is reduced considerably by using a dual-pair comparison task. Altogether this supports the view that this effect is influenced by nonsensory cognitive factors (learning, memory, context) that vary with experimental task.

One important implication of the learning and task effects observed here concerns the relative contribution of the ear's ability to encode the stimulus and the listeners' ability to use this information to generate correct decisions in the various experimental tasks. Given the same basic gap and no-gap stimuli used across conditions, variance in performance appears to be more related to the latter. This raises the obvious question of what aspect of temporal processing might relate to language outcome. Does language outcome depend on basic sensory encoding of temporal information, or does it depend more on the complex perceptual abilities to process and integrate this information with respect to memory and other contextual information? The present results clearly show that gap detection thresholds (and presumably other measures of temporal processing) reflect a variety of sensory, perceptual, and cognitive limitations, combined in varying degrees as a function of stimulus properties, task, and learning. By choosing a particular set of stimuli and experimental procedures, any test of the temporal processing hypothesis makes an assumption about what aspect of temporal processing plays a putative role in the language outcome.

In the context of the relation between temporal processing and language outcome, there are a number of important considerations. If we assume, for the moment, that auditory temporal processing does play a role in language processing, then the present findings raise the possibility that this relation will only be observed with some tasks and some stimuli. Alternatively, it may be that previously reported correlations between temporal processing and language outcome are largely a result of nontemporal processing factors, such as perceptual or working memory. That is, a significant relation between auditory temporal processing and language ability may have emerged in the literature because the specific tasks and stimuli chosen in different designs required a cognitive component that is related to language outcome even though auditory temporal processing itself may not be. It is also possible that both types of factors are related to language outcome, but the use of different experimental designs or stimuli have lent support to one conclusion or another.

One theme emerges from studies that support the temporal processing hypothesis. These studies typically employ tasks and stimuli that involve the perception of the relation between different elements or features, rather than the mere detection of a single feature or element. For example, the ability to detect a gap in noise is a poor predictor of language outcome (Ahissar, Protopapas, Reid, & Merzenich, 2000; Breier, Fletcher, Foorman, Klaas, & Gray, 2003; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1998), but the ability to represent spectrotemporal relations between elements, as in temporal order judgments (Au & Lovegrove, 2001) or the judgments of the directionality of frequency changes (rather than the detection of a frequency change), do appear to have predictive power (Banai & Ahissar, 2006).

Not all putative temporal processing measures tap the same underlying processing mechanisms. The present results clearly show that listeners' temporal resolution cannot be adequately described in terms of a single performance measure, such as a gap detection threshold. Different aspects of temporal processing are engaged, even within a family of related gap detection tasks. Acknowledgment of the multidimensional nature of temporal processing is also critical for making sense of conflicting reports of the relation between temporal processing measures and language outcome.

The present results illustrate that the various gap detection tasks offer different sets of advantages. For example, the dual-pair comparison task provides highly reliable thresholds in the sense that listeners' performance did not vary much within and between test sessions. This may be considered a strength. However, there is very little variability between listeners, which may limit the use of dual-pair comparison tasks as predictors of language outcome. The present study did not test listeners with language problems, and it remains to be seen whether these listeners might also fall within this tightly constrained range. This of course depends on the underlying basis of the language deficits.

In contrast, thresholds obtained in yes-no and 2IFC tasks do demonstrate greater variability across listeners, which may afford a greater sensitivity to individual differences in temporal processing. This quality is advantageous for attempts to relate these differences to language outcome. However, these individual differences are drastically reduced with training as thresholds improve, suggesting that the individual differences are related more to cognitive processing factors than to temporal resolution itself.

Clearly a listener's temporal resolution cannot be adequately described in terms of a single performance measure, such as a gap detection threshold, as many stimulus characteristics and processing factors play a role. Some aspects of temporal processing may be related to language outcome, along the lines of the temporal processing hypothesis, and others may not. It is only by better understanding the perceptual processes underlying particular temporal processing tasks that conflicting results in the literature can be reconciled to provide a coherent account of the possible role, if any, of temporal processing ability in language deficits.

## Acknowledgments

This research was supported by the Canadian Institutes of Health Research (awarded to the second author), the Natural Sciences and Engineering Research Council of Canada (awarded to the second and fifth authors), the Canadian Language and Literacy Research Network (awarded to the second and fifth authors), and the Ontario Ministry of Development and Trade (a Premier's Research Excellence Award to the fifth author) on infrastructure provided through a new opportunities grant to the fifth author from the Canadian Foundation for Innovation and the Ontario Innovation Trust. We thank Walt Jesteadt for comments on this article.

## References

**Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. M.** (2000). Auditory processing parallels reading abilities in auditory. *Proceedings of the National Academy of Sciences, 97,* 6832–6837.

**Au, A., & Lovegrove, B.** (2001). Temporal processing ability in above average and average readers. *Perception & Psychophysics, 63,* 148–155.

**Banai, K., & Ahissar, M.** (2006). Auditory processing deficits in dyslexia: Task or stimulus related? *Cerebral Cortex, 16,* 1718–1728.

**Benasich, A. A., & Tallal, P.** (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural Brain Research, 136,* 31–49.

**Benasich, A. A., Thomas, J. L., Choudhury, N., & Leppänen, P. H. T.** (2002). The importance of rapid processing abilities to early language development: Evidence from converging methodologies. *Developmental Psychobiology, 40,* 278–292.

**Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J.** (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language and Hearing Research, 42,* 1295–1310.

**Breier, J. I., Fletcher, J. M., Foorman, B. R., Klaas, P., & Gray, L. C.** (2003). Auditory temporal processing in children with specific reading disability with and without attention deficit hyperactivity disorder. *Journal of Speech, Language, and Hearing Research, 46,* 31–42.

**Buss, E., Hall, J. W., Grose, J. H., & Dev, M. D.** (1999). Development of adult-like performance in backward, simultaneous, and forward masking. *Journal of Speech, Language, and Hearing Research, 42,* 844–849.

**Creelman, C. D., & Macmillan, N. A.** (1979). Auditory phase and frequency discrimination: A comparison of nine procedures. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 146–156.

**Elliott, L. L.** (1964). Backward masking: Different durations of the masking stimulus. *Journal of the Acoustical Society of America, 36,* 393.

**Elliott, L. L.** (1967). Development of auditory narrow-band frequency contours. *Journal of the Acoustical Society of America, 42,* 143–153.

**Farmer, M. E., & Klein, R. M.** (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review, 2,* 460–493.

**Fitzgibbons, P., Pollatsek, A., & Thomas, I.** (1974). Detection of temporal gaps within and between perceptual tonal groups. *Perception & Psychophysics, 16,* 522–528.

**Formby, C., & Forrest, T. G.** (1991). Detection of silent temporal gaps in sinusoidal markers. *Journal of the Acoustical Society of America, 89,* 830–837.

**Formby, C., Gerber, M. J., Sherlock, L. P., & Magder, L. S.** (1998). Evidence for an across-frequency, between-channel process in asymptotic monaural temporal gap detection. *Journal of the Acoustical Society of America, 103,* 3554–3560.

**Formby, C., Sherlock, L. P., & Li, S.** (1998). Temporal gap detection measured with multiple sinusoidal markers: Effects of marker number, frequency, and temporal position. *Journal of the Acoustical Society of America, 104,* 984–998.

**Friel-Patti, S., Loeb, D. F., & Gillam, R. B.** (2001). Looking ahead: An introduction to five exploratory studies of Fast ForWord. *American Journal of Speech-Language Pathology, 10,* 195–202.

**Gibson, E. J.** (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

**Gillam, R. B., Crofford, J. A., Gale, M. A., & Hoffman, L. M.** (2001). Language change following computer-assisted language instruction with Fast ForWord or Laureate Learning Systems software. *American Journal of Speech-Language Pathology, 10,* 231–247.

**Goldstone, R. L.** (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612.

**Grose, J. H., Hall, J. W., Buss, E., & Hatch, D.** (2001). Gap detection for similar and dissimilar gap markers. *Journal of the Acoustical Society of America, 109,* 1587–1595.

**Habib, M.** (2000). The neurological basis of developmental dyslexia: An overview and working hypothesis. *Brain, 123,* 2373–2399.

**Hartley, D. E. H., & Moore, D. R.** (2002). Auditory processing efficiency deficits in children with developmental language impairments. *Journal of the Acoustical Society of America, 112,* 2962–2966.

**Hartley, D. E. H., Wright, B. A., Hogan, S. C., & Moore, D. R.** (2000). Age–related improvements in auditory backward and simultaneous masking in 6- and 10-year-old children. *Journal of Speech, Language, and Hearing Research, 43,* 1402–1415.

**Irvine, D. R. F., Martin, R. L., Klimkeit, E., & Smith, R.** (2000). Specificity of perceptual learning in a frequency discrimination task. *Journal of the Acoustical Society of America, 108,* 2964–2968.

**Jesteadt, W., & Bilger, R. C.** (1974). Intensity and frequency discrimination in one- and two-interval paradigms. *Journal of the Acoustical Society of America, 55,* 1266–1276.

**Kallman, H. J., & Brown, S. C.** (1986). Auditory backward recognition masking with well-practiced listeners. *The Journal of Auditory Research, 26,* 239–246.

**Karmarkar, U. R., & Buonomano, D. V.** (2003). Temporal specificity of perceptual learning in an auditory discrimination task. *Learning & Memory, 10,* 141–147.

**Leek, M. R., & Watson, C. S.** (1984). Learning to detect auditory pattern components. *Journal of the Acoustical Society of America, 76,* 1037–1044.

**Levitt, H.** (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49,* 467–477.

**Logan, J. S., Lively, S. E., & Pisoni, D. B.** (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America, 89,* 874–886.

**Macmillan, N. A., & Creelman, C. D.** (2005). *Detection theory: A user's guide*. Mahwah, NJ: Erlbaum.

**MacMillan, N. A., Kaplan, H. L., & Creelman, C. D.** (1977). The psychophysics of categorical perception. *Psychological Review, 84,* 452–471.

**Marler, J. A., & Champlin, C. A.** (2005). Sensory processing of backward-masking signals in children with language-learning impairment as assessed with the auditory brainstem response. *Journal of Speech, Language, and Hearing Research, 48,* 189–203.

**Marler, J., Champlin, C. A., & Gillam, R. B.** (2001). Backward and simultaneous masking measured in children with language-learning impairments who received intervention with Fast ForWord or Laureate Learning Systems software. *American Journal of Speech-Language Pathology, 10,* 258–268.

**McAnally, K. I., & Stein, J. F.** (1996). Auditory temporal coding in dyslexia. *Proceedings of the Royal Society of London, 263,* 961–965.

**McArthur, G. M., & Bishop, D. V. M.** (2001). Auditory perceptual processing in people with reading and oral language impairments: Current issues and recommendations. *Dyslexia, 7,* 150–170.

**McArthur, G. M., & Hogben, J. H.** (2000). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *Journal of the Acoustical Society of America, 109,* 1092–1100.

**Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P.** (1996, January 5). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271,* 77–80.

**Micheyl, C., & Messing, D. P.** (2006). Likelihood ratio, optimal decision rules, and correct response probabilities in a signal detection theoretic, equal-variance Gaussian model of the observer in the 4IAX paradigm. *Perception & Psychophysics, 68,* 725–735.

**Miyazaki, K., & Sasaki, T.** (1984). Pure-tone masking patterns in nonsimultaneous masking conditions. *Japanese Psychological Research, 26,* 110–119.

**Moore, D. R., Rosenberg, J. F., & Coleman, J. S.** (2005). Discrimination training of phonemic contrasts enhances processing in mainstream school children. *Brain and Language, 94,* 72–85.

**Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., & Merzenich, M. M.** (1998). Practice related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere and modality. *The Journal of Neuroscience, 18,* 1559–1570.

**Penner, M. J.** (1974). Effect of masker duration and masker level on forward and backward masking. *Journal of the Acoustical Society of America, 56,* 179–182.

**Penner, M. J.** (1977). Detection of temporal gaps in noise as a measure of the decay of auditory sensations. *Journal of the Acoustical Society of America, 61,* 552–557.

**Penner, M. J., Cudahy, E., & Jenkins, G. W.** (1974). The effect of masker duration on forward and backward masking. *Perception & Psychophysics, 15,* 405–410.

**Phillips, D. P.** (1999). Auditory gap detection, perceptual channels, and temporal resolution in speech perception. *Journal of the American Academy of Audiology, 10,* 343–354.

**Phillips, D. P., & Hall, S. E.** (2000). Independence of frequency channels in auditory temporal gap detection. *Journal of the Acoustical Society of America, 108,* 2957–2963.

**Phillips, D. P., & Smith, J. C.** (2004). Correlations among within-channel and between-channel auditory gap detection thresholds in normal listeners. *Perception, 33,* 371–378.

**Phillips, D. P., Taylor, T. L., Hall, S. E., Carr, M. M., & Mossop, J. E.** (1997). Detection of silent intervals between noises activating different perceptual channels: Some properties of "central" auditory gap detection. *Journal of the Acoustical Society of America, 101,* 3694–3704.

**Repp, B. H., Liberman, A. M., Eccardt, T., & Pesetsky, D.** (1978). Perceptual integration of acoustic cues for stop, fricative, and affricative manner. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 621–637.

**Roth, D. A., Kishon-Rabin, L., & Hildesheimer, M.** (2001). Auditory backward masking and the effect of training in normal hearing adults. *Journal of Basic & Clinical Physiology & Pharmacology, 12,* 145–159.

**Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., & Lamb, M.** (1994). Gap detection and the precedence effect in young and old adults. *Journal of the Acoustical Society of America, 95,* 980–991.

**Shulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H.** (1998). Role of auditory temporal processing for reading and spelling disability. *Perceptual and Motor Skills, 86,* 1043–1047.

**Tallal, P., & Gaab, N.** (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neuroscience, 29,* 382–390.

**Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M., & Merzenich, M. M.** (1996, January 5). Language comprehension in language learning impaired children improved with acoustically modified speech. *Science, 271,* 81–84.

**Taylor, T. L., Hall, S. E., Boehnke, S. E., & Phillips, D. P.** (1999). Additivity of perceptual channel-crossing effects in auditory gap detection. *Journal of the Acoustical Society of America, 105,* 563–566.

**Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A.** (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106,* 3649–3658.

**Watson, C. S., Kelly, W. J., & Wroton, H. W.** (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America, 60,* 1176–1186.

**Wiegrebe, L., & Krumbholz, K.** (1999). Temporal resolution and temporal masking properties of transient stimuli: Data and an auditory model. *Journal of the Acoustical Society of America, 105,* 2746–2756.

**Wright, B. A., Buonomano, D. V., Mahncke, H. W., & Merzenich, M. M.** (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *The Journal of Neuroscience, 17,* 3956–3963.

**Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M.** (1997, May 8). Deficits in auditory temporal and spectral resolution in language impaired children. *Nature, 387,* 176–178.

**Wright, H. N.** (1964). Backward masking for tones in narrowband noise. *Journal of the Acoustical Society of America, 36,* 2217–2221.

Received March 7, 2007

Revision received October 24, 2007

Accepted April 1, 2008

DOI: 10.1044/1092-4388(2008/07-0058)

Contact author: David I. Shore, Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada. E-mail: dshore@mcmaster.ca.

Nicholas A. Smith is now at Boys Town National Research Hospital, Omaha, Nebraska.