Different Response Patterns between Auditory Spectral and Spatial Temporal Order Judgment (TOJ)

Leah Fostick

Department of Communication Disorders, Ariel University, Ariel, Israel

Harvey Babkoff

Department of Psychology, Ashkelon Academic College, Ashkelon, Israel

Author note

Leah Fostick is the corresponding author for this manuscript. Mail to Leah Fostick, Department of Communication Disorders, Ariel University, Ariel 40700, Israel, Leah.fostick@ariel.ac.il, Tel. +972-54-4925354, Fax. +972-3-9765743
Abstract

Temporal order judgment (TOJ) thresholds have been widely reported as valid estimates of the temporal disparity necessary for correctly identifying the order of two stimuli. Data for two auditory TOJ paradigms are often reported in the literature: (1) spatially-based TOJ in which the order of presentation of the same stimulus to the right and left ear differ; and (2) spectrally-based TOJ in which the order of two stimuli differing in frequency, are presented to one ear or to both ears simultaneously. Since the thresholds reported using the two paradigms differ, the aim of the current study was to compare their response patterns. The results from three different experiments showed that: (1) while almost none of the participants were able to perform the spatial TOJ task when ISI=5msec, with the spectral task, 50% reached an accuracy level of 75% when ISI=5msec; (2) temporal separation was only a partial predictor for performance in the spectral task, while it fully predicted performance in the spatial task; and (3) training improved performance markedly in the spectral TOJ task, but had no effect on spatial TOJ. These results suggest that the two paradigms may reflect different perceptual mechanisms.

Keywords: auditory temporal order judgment (TOJ); Spectral TOJ; Spatial TOJ; perceptual cues
Introduction

Temporal order judgment (TOJ) refers to the ability to correctly perceive the temporal order of at least two stimuli. TOJ has been studied in different modalities, in the current paper, we focus on auditory TOJ. The study of the human capacity to judge the temporal order of two stimuli was described by Hirsh (1959) and Hirsh and Sherrick (1961). In these early studies, the authors reported that the stimulus-onset-asynchrony (SOA) required to identify the order of two successive stimuli was approximately 20 msec, regardless of stimulus modality or presentation mode. However, more recent studies of auditory temporal order judgment usually reported TOJ inter-stimulus-interval (ISI) thresholds ranging from 30-60 msec (e.g., Babkoff, Zukerman, Fostick, & Ben-Artzi, 2005; Ben-Artzi, Fostick, & Babkoff, 2005; Broadbent & Ladefoged, 1959; Efron, 1973; Fay, 1966; Fink, Churan, & Wittmann, 2005; Fink, Ulbrich, Churan, & Wittmann, 2006a; Green, 1971; Kanabus, Szelag, Rojek, & Pöppel, 2002; Kinney, 1961; Szelag, Kanabus, Kolodziejeczyk, Kowalska, & Szuchnik, 2004; Szymaszek, Sereda, Pöppel, & Szelag, 2009; Szymaszek, Szelag, & Sliwowska, 2006).

Elevated TOJ thresholds have been associated with pathologies in a variety of sensori-motor and cognitive functions (Lewandowska, Piatkowska-Janko, Bogorodzki, Wolak, & Szela, 2010; Mauk & Buonomano, 2004; Pöppel, 1994, 1997; Tallal, Merzenich, Miller, & Jenkins, 1998). Furthermore, there is a growing body of evidence that associates elevated TOJ thresholds in a number of different sub-populations with language difficulties. These sub-populations include aphasic patients (von Steinbuchel, 1998; von Steinbuchel, Wittmann, Strasburger, Szelag, 1999; Fink et al., 2006a); dyslexic
readers (Ben-Artzi et al., 2005; Fostick, Bar-El, Ram-Tsur, 2012; Fostick, Ben-Artzi, Babkoff, 2008; Reed, 1989; Tallal, 1980); aging adults (Ben-Artzi, Babkoff, & Fostick, 2011; Fink et al., 2005; Fostick, Ben-Artzi, & Babkoff, 2007; Szymaszek et al., 2006, 2009); children with autistic spectrum disorders (ASD) (Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011); children and young adults with attention deficit/hyperactivity disorder (ADHD) (Breier, Gray, Fletcher, Foorman, Klaas, 2002; Kleiner, Negbi, Or, Zuaretz, & Fostick, 2011); and sleep deprived young adults (Babkoff et al., 2005; Fostick and Zukerman, 2010). Moreover, several studies have reported an association between elevated TOJ thresholds and phonological awareness among young and aging adults (Fink et al., 2005), aphasic patients (Fink et al., 2006a), dyslexic readers (Ben-Artzi et al., 2005; Fostick and Zukerman, 2010; Tallal, 1980) and sleep deprived young adults (Babkoff et al., 2005; Fostick and Zukerman, 2010).

A number of researchers have suggested that the perceptual process of order judgment involves two phases: (1) direct perception of the elements in a pair of tones separated temporally, dependent on the inter-stimulus interval, i.e., the perception of the order of their occurrence; and (2) a holistic perception of the tonal patterns created by the order of the elements without separate perception of the elements and the judgment of difference between two tones whose temporal order is inverted, based on the perceived differences between these patterns (Ben-Artzi et al., 2005; Thomas & Fitzgibbons, 1971; Warren, 1974a; Warren and Ackroff, 1976). However, in a number of studies, Warren and his associates (Warren, 1974a, 1974b; Warren, 1993; Warren and Ackroff, 1976; Warren and Byrnes, 1975; Warren and Obusek, 1972) have shown that sequence perception tasks might involve two types of sequence perception: Direct identification of
the elements and their order, and Holistic pattern recognition, such as in speech and music recognition (Warren, 1993). Efron (1973) also made a distinction between identifying the order of elements in a sequence and discriminating between sequences with different orders. He defined pairs of two-tone stimuli, experienced as a unitary perceptual event, as Micropatterns.

The brain mechanisms involved in judgments of temporal order have been investigated in a number of studies. Several imaging studies have reported that the left posterior sylvian region and the left temporal parietal junction are involved in judgments of temporal order (Bernasconi, Grivel, Murraya, & Spierer, 2010; Bernasconi, Manuel, Murray, & Spierer, 2011; Lewandowska et al., 2010); while others have also indicated the presence of prefrontal activation to rapid auditory stimuli (Lewis and Miall, 2003; Pastor, Macaluso, Day, & Frackowiak, 2006; von Steinbuchel et al., 1999). Brechmann & Scheich (2005) and Lewandowska et al. (2010) concluded that the locus of maximum cortical activity also depends on the characteristics of the stimuli and task.

A variety of different experimental paradigms have been used to study TOJ. In general, these experimental paradigms can be divided into two classes: 1) those that are spectrally based (Ben-Artzi et al., 2005; Brechmann & Scheich, 2005; Bregman, Ahad, Crum, & O’reilly, 2000; Fink et al., 2005, 2006a; Fink, Ulbrich, Churan, & Wittmann 2006b; Kanabus et al., 2002; Rahne et al., 2008; Reed, 1989; Szymaszek et al., 2006, 2009; Tallal, 1980); and 2) those that are spatially based (Bakoff et al., 2005; Ben-Artzi et al., 2005; Bernasconi et al., 2010, 2011; Brechmann & Scheich, 2005; Fink et al., 2005, 2006a, 2006b; Szymaszek et al., 2006, 2009). In spectrally based TOJ paradigms, the participant is presented with two brief duration sinusoidal tones (usually ranging
between 15-75 msec) separated by an inter-stimulus-interval (ISI) and is required to judge the order of presentation of the tones (high-low or low-high). The tones can be presented either monaurally (both tones presented to one ear) or diotically (both tones presented to both ears). In the spatially based TOJ paradigm, the stimuli can be either sinusoidal tones (of the same frequency) or clicks, one of which is presented to one ear while the other is presented to the other ear. The stimuli are separated by an ISI and the participant is required to judge the order of the stimuli based on the order of their presentation (left-right or right-left). The ISI required for correct identification of the order (according to the method in use) is the TOJ threshold.

Both classes of TOJ paradigm have been used and reported in the literature for almost 60 years. The general understanding of the meaning of temporal order is that the judgment represents a measure of the perception of the temporal relationships between two or more stimuli, and is therefore, a fairly basic indicator, or marker, of "temporality" (Hirsh and Sherrick, 1961). One difficulty with this explanation is that TOJ can be performed not only by a perception of the temporal relation between the stimuli presented, but also by perceiving the global pattern of the sounds (e.g., Warren 1974a, 1974b). Another difficulty is the fact that the threshold values of ISIs necessary for correct judgment of order, that are reported for the spatial vs. the spectral paradigm, are quite different. Table 1 summarizes studies reporting spectral and spatial TOJ thresholds. Note that in general, the ISIs necessary for successful judgment of spectral TOJ are smaller than those reported for the spatial TOJ paradigm (mean threshold = 21.5±13.5msec VS 49.14±14.72msec, respectively).
Therefore, the aim of the current study was to compare the response patterns of spatially-based TOJ and spectrally-based TOJ paradigms. We first replicated the finding in the literature, that spectral TOJ can be performed for shorter ISIs than spatial TOJ. Second, we compared the predicted value of the time separation necessary between the onset of the leading tone to the onset of the lagging tone (stimulus-onset-asynchrony, SOA) for accurate performance of the spectral TOJ task with the spatial TOJ task. Finally, we compared the learning curve while performing spectral TOJ with the learning curve while performing spatial TOJ.

**Experiment 1**

As noted above, the spectral TOJ thresholds reported in the literature are shorter than those reported for spatial TOJ (see Table 1). This means that participants can successfully perform spectral TOJ with much shorter ISIs than spatial TOJ. In order to replicate the findings that spectral TOJ can be performed at shorter ISIs than spatial TOJ, the aim of Experiment 1 was to compare the performance level of spectral and spatial TOJ, and especially when the two 15 msec duration stimuli are separated by only 5 msec (SOA = 20 msec).

**Method**

**Participants.** Sixty-eight undergraduate students (age 21-39, 43% women) participated in the experiment. All participants had audiometric thresholds ≤ 20 dB for 500 – 4,000 Hz.
**Tasks and stimuli. Spectral TOJ.** Pairs of 15 msec pure tones, one, 1kHz and the other, 1.8kHz, were presented diotically at 40 dB SL, and with 1 msec rise/fall times on each trial. In half of the trials the order of the tones was the high frequency (1.8 kHz) tone followed by the low frequency tone, and in the reverse order on the other half of the trials. Participants were required to reproduce the order in which they heard the tones (high tone first then low tone, or the reverse) by pressing the relevant keys on the keyboard in the order they heard. Tone pairs were presented with an ISI of 5, 10, 15, 30, 60, 90, 120 or 240 msec. Each ISI value was repeated 16 times, resulting in a total of 32 trials per participant for the current analysis.

Experimentation followed four training phases (see also Ben Artzi et al., 2005). In the first training phase, participants were familiarized with the stimuli used in the study by listening to five low frequency tones followed by five high frequency tones. In the second phase, the participants were trained to associate each tone with the proper response key in 32 trials in each they were required to press the correct key for each tone they heard. Feedback was given following each response (“right” or “wrong”). In the third phase, the participants were required to press the correct key for each tone, but without any feedback. All participants passed this phase successfully with at least 20 correct trials out of 24. In the fourth phase, participants were trained to reproduce the order of two tones. This phase was similar to the conditions of the experiment, but with an ISI = 240 msec and with only 32 trials. Performance in this phase was accompanied by appropriate feedback after each response (“right” or “wrong”).
In order to present the tones at 40 dB SL, hearing threshold was measured by a two alternative forced choice 2-down-1-up adaptive staircase procedure. The threshold was calculated as the average of the last eight out of 10 reversals.

Spatial TOJ. Pairs of 15 msec pure 1 kHz tones were presented monaurally at 40 dB SL, with 1 msec rise/fall time on each trial. The tones were always presented with an inter-stimulus interval (ISI) of either: 5, 10, 15, 30, 60, 90, 120 or 240 msec separating them. On half of the trials, the order of the presentation of the tones was left ear first followed by the right ear, and in the reverse order on the other half of the trials. Participants were required to reproduce the order in which they heard the tones (left ear first then right ear, or the reverse) by pressing the relevant keyboard keys in the order corresponding to what they heard. Each ISI value was repeated 16 times, resulting in a total of 32 trials for the current analysis.

The four phases of training were the same as for the spectral TOJ paradigm. The participants’ absolute hearing threshold was measured by a two alternative forced choice 2-down-1-up adaptive staircase procedure. The threshold was the average of the last eight out of 10 reversals.

Apparatus. A Pentium1 personal computer controlled the stimulus presentation and recorded responses for both paradigms. The auditory stimuli were generated by a sound-generator device (TDT-system II: Tucker-Davis Technologies, Gainesville, FL), and were presented binaurally through TDH- 49 headphones. Tasks were programmed using Matlab™ software version 6.5. Screening for hearing level was performed using Danplex DA64 or Maico Hearing Instruments Ltd MA32 audiometers.
**Procedure.** Prior to experimentation, participants received full explanations about the study, followed by written informed consent. Since the study was a within-subjects design, the spectral and spatial TOJ paradigms were tested in a random order, each one beginning with the training phase.

**Results**

A three-way repeated measures analysis was performed with leading ear, ISI, and paradigm as within subjects variables. No effect was found for leading ear ($F_{(1,67)}=1.23, p>.05$), or any leading ear * ISI ($F_{(7,469)}=952, p>.05$) or leading ear * paradigm ($F_{(1,67)}=1.529, p<.01$) interactions. As no differences were found between leading ears, the data of both leading ears were averaged. Figure 1(a) shows percent correct by ISI for each paradigm. Significant main effects were found for paradigm ($F_{(1,67)}=15.90, p<.001$) and ISI ($F_{(7,469)}=178.90, p<.001$). Overall, higher accuracy was found for the spectral task (80.4%), as compared with the spatial (72.3%), and significant differences between all ISI values, except from between 10 and 15 msec (66.6% and 67.5%, respectively, $p>.05$). A significant ISI*paradigm interaction was also found ($F_{(7,469)}=89.15, p<.001$).

Better performance in spectral TOJ was found at ISIs of 5 to 60 msec (5 msec: $t_{(67)}=7.667, p<.001$; 10 msec: $t_{(67)}=8.341, p<.001$; 15 msec: $t_{(67)}=7.792, p<.001$; 30 msec: $t_{(67)}=6.450, p<.001$; 60 msec: $t_{(67)}=2.351, p<.05$), no difference was found at 90 msec ($t_{(67)}=-.424, p>.05$), and better performance for spatial task performance was found at 120 and 240 msec ($t_{(67)}=-2.677, p<.01$ and $t_{(67)}=-5.173, p<.001$, respectively).

Pearson correlations revealed no association between the paradigms at most of the ISIs (5 msec: $r=.164, p>.05$; 10 msec: $r=.063, p>.05$; 15 msec: $r=.189, p>.05$; 30 msec:
r=.115, p>.05; 60 msec: r=.160, p>.05), but a significant correlation was found at the longer ISIs (90 msec: r=.357, p<.01; 120 msec: r=.335, p<.01; 240 msec: r=.348, p<.01).

We compared the number of participants who identified the correct order of the tones at 75% or better for the two TOJ paradigms at each ISI (Fig. 1). The results indicated that at ISIs of 60 and 90 msec, similar number of participants were able to correctly identify the order of the tones at 75% accuracy or better with both paradigms ($\chi^2(1)=3.18$, $p>.05$ and $\chi^2(1)=1.98$, $p>.05$, respectively). At longer ISIs (120 and 240 msec), more participants were able to correctly identify the order of the tones at the spatial paradigm ($\chi^2(1)=12.91$, $p<.001$ and $\chi^2(1)=45.61$, $p<.001$, respectively). However at shorter ISIs (10, 15, and 30 msec) more participants were able to correctly perform the spectral than the spatial paradigm ($\chi^2(1)=11.28$, $p<.01$; $\chi^2(1)=95.67$, $p<.001$; and $\chi^2(1)=65.96$, $p<.001$, respectively).

At the shortest ISI (5 msec), 31 participants (46%) succeeded in identifying the order of the tones at the spectral paradigm, but only two participants (3%) were successful in performing the spatial TOJ paradigm (Fig. 1). This difference was significant ($\chi^2(1)=18.19$, $p<.001$). As reported earlier, at this short ISI there were significant differences in accuracy and no correlation between the paradigms.

---

1 Recently we tested the effect of the frequency difference between the stimuli on performance in spectral TOJ. We found that when the frequencies were 1 and 1.1kHz, 50% of participants succeeded in reproducing the order of the tones when ISI was 0 msec, and 32% when the frequencies were 1 and 3.5kHz ($\chi^2(1)=1.19$, $p>.05$). These results suggest that increasing the difference in frequency between the tones does not positively increase the performance of spectral TOJ. The findings were presented at the Auditory Perception, Cognition, and Action Meeting, Minneapolis, MN, November 15, 2012.
Discussion

The results show significant differences between the two paradigms in the ability to correctly identify the order of the tones. There is a gradual increase in the accuracy rate of the spatial TOJ task and in the number of participants performing correctly as the ISI increases. However, it seems that the accuracy rate and the number of participants performing the spectral TOJ task correctly, does not change that much with ISI. Interestingly, this finding implies that most of the participants, who could not perform the spectral task at short ISIs, still could not correctly identify the order of the tones with longer ISIs, which is reflected in higher accuracy rate in long ISIs in the spatial task. Therefore, it seems that increasing ISI does not influence accuracy in the same way for spectral TOJ as it does for spatial TOJ. In addition, the lack of correlation between the paradigms at most of the ISIs suggests that the participants use different judgment criteria, at least in the short-to-mid range ISIs.

The aim of the current analysis was to test participants’ ability to reproduce the order of two tones at very short ISIs, namely 5 msec. The data indicate that most if not all participants are unable to accurately judge spatial TOJ with ISI as short as 5 msec (SOA = 20 msec). In contrast, 46% of the participants were able to correctly judge spectral TOJ at the same ISI = 5 msec. These results complement other reports in the literature of very short spectral TOJ thresholds (Table 1a) and of participants being able to accurately judge spectral TOJ even when stimulus intervals were close to zero. For example, Fink et al. (2005) found that when performing spectral TOJ, 43% of the young adults they tested reached an accuracy of nearly 100% when SOA was 11 msec (ISI = 1 msec). Ben-Artzi et al. (2005) showed that young control participants reached an asymptote in the
performance of spectral TOJ and achieved an average of 90% correct when ISI = 8 msec, while they averaged only about 55% correct on the spatial TOJ paradigm at ISI = 8 msec.

The difference in the pattern of performance between spectral and spatial TOJ, although performed by the same participants, indicates that a very small separation between stimuli, if any, is required to perform the spectral TOJ paradigm, while a substantially longer temporal separation is required for performing the spatial paradigm. In fact, temporal separation between stimuli may not be the most crucial parameter for successful performance of the spectral TOJ paradigm, while it certainly seems the most crucial parameter for spectral TOJ. Experiment 2 was designed to test the hypothesis that the temporal separation between stimuli may play different roles in the performance of spectral and spatial TOJ.

Experiment 2

The aim of Experiment 2 was to test the role of temporal separation between the first and the second stimuli in the performance of spectral and spatial TOJ. The total separation time between the stimuli is composed of the duration of the first tone presented plus the ISI, namely, the stimulus-onset asynchrony (SOA) (Babkoff and Fostick, in press; Fostick, Ben-Artzi, & Babkoff, 2011). In order to compare the role of the temporal separation between the first and the second stimuli in each task, we compared the performance accuracy and response time of TOJ when plotted as a function of SOA for both paradigms.
Methods

Participants. Fifty-six undergraduate students (mean age = 22 years, 74% females) participated in the experiment. Half of them were randomly assigned to the spectral TOJ task, and half to the spatial TOJ task.

Tasks and stimuli. Spectral TOJ. Pairs of pure tones, one at 1kHz and the other at 1.8kHz, were presented diotically at 60 dB SPL, and with 1 msec rise/fall time. Tone durations were 5, 10, 20, 30 or 40 msec and were presented with an ISI of 5, 10, 15, 30, 60, 90, 120 or 240 msec, yielding SOA values of 10-280 msec. On each trial, participants were presented with a pair of tones, consisting of one high frequency and one low frequency tone. In half of the trials the order of the tones was high frequency tone first then low frequency tone, and in the reverse order in the other half of the trials. Each ISI-duration combination was repeated 16 times, resulting in a total of 1,280 trials per participant (2 tone orders x 8 ISIs x 5 stimulus durations x 16). All stimuli were presented in a random order. After every 32 trials, participants received a short recess. Participants were required to reproduce the order in which they heard the tones (high tone first then low tone, or the reverse) by pressing the relevant keyboard keys in the order corresponding to what they heard. Experimentation followed the four-phases of training as described in Experiment 1.

Spatial TOJ. Pairs of 1kHz pure tones were presented monaurally at 60 dB SPL, and with 1 msec rise/fall time. Tone durations and ISIs were exactly the same as for the spectral TOJ paradigm. As in the spectral TOJ paradigm, on each trial participants were presented with a pair of tones, consisting of one tone to the left ear and one tone to the right ear. In half of the trials the order of the tones was: left ear first then right ear, and
the reverse order in the other half of the trials. Each ISI-duration combination was repeated 16 times, resulting in a total of 1,280 trials (2 tone orders x 8 ISIs x 5 stimulus durations x 16). All stimuli were presented in a random order. After every 32 trials participants received a short recess. Participants were required to reproduce the order in which they heard the tones (left ear first then right ear, or the reverse) by pressing the relevant keyboard keys in the order corresponding to what they heard. Experimentation followed the four-phases of training as described in Experiment 1.

**Apparatus.** A 486 personal computer controlled the stimulus presentation and recorded responses and response times for both paradigms. Response time represents the time elapsed from the offset of the second stimulus to the participant’s pressing the first response key. Auditory stimuli were generated by a sound-generator device (TDT-system II: Tucker-Davis Technologies, Gainesville, FL), and then presented binaurally through TDH- 49 headphones. Tasks were programmed in C.

**Procedure.** The study was done in a between-subjects design. Each task began with a training phase, followed by the experiment itself. Prior to experimentation the participants received full explanation about the study and signed an informed consent.

**Results**

The combination of stimulus duration values and ISIs resulted in 29 different values of SOA ranging from 10- 280 msec. Three-way repeated measures ANOVA with leading ear and ISI as within subjects variables and TOJ paradigm as between subjects variable revealed no effect for leading ear (F(1,55)=.180, p>.05), or any leading ear * ISI (F(7,385)=.733, p>.05) and leading ear * paradigm (F(1,55)=.123, p>.05) interaction. As no
difference between leading ears was found, the data of both leading ears was averaged. In order to test the predictive value of SOA for the spectral and spatial TOJ paradigms, accuracy and response time for both tasks were plotted against SOA. Figure 2 shows response accuracy and response time plotted as a function of SOA for the spatially-based TOJ paradigm, while comparable data for the spectrally-based TOJ paradigm are shown in Fig. 3.

Table 2 presents the estimated predicting value of SOA using linear, 2\textsuperscript{nd} order polynomial, Gaussian, and logistic curves. Plotting accuracy or response time as a function of SOA predicted 86 to 97\% of the variance in accuracy for the spatial TOJ paradigm and 94 to 97\% for response time. However, the same equations account for zero to 54\% of the variance for response accuracy and for only 36 to 62\% of the response time variance for the spectral TOJ paradigm. A Fisher r-to-z transformation revealed significant differences between the predicting value of the spatial TOJ and spectral TOJ accuracy and response time for all fitting functions.

Discussion

Experimental paradigms designed to test temporal order judgments assume that the physical dimension for response accuracy is the temporal disparity between the presentations of two stimuli whose order is judged. Consequently, accuracy should be a monotonic function of the temporal disparity between the two stimuli and increase systematically from chance level at a short SOA (5-to-10 msec) to an (almost) error-less judgment as a function of increasing temporal intervals. While the previous description certainly fits the spatial TOJ paradigm, the spectral TOJ accuracy data, however, do not
always show systematic increases from chance level at around 5-10 msec to highly accurate as a function of increases in SOA, but in fact show a jump from low to high accuracy even at very short SOAs. The data of Experiment 2 thus complement the data of Experiment 1.

As shown in Figure 2, the changes both in accuracy and response time with the spatial TOJ paradigm follow the classical form of the psychometric function, beginning with very low accuracy levels (52%-57%) and fairly long reaction times (1128-1111 msec) associated with the shorter SOAs (10-35 msec), followed by a slow increase in accuracy (59%-83%) and decrease in reaction time (1136-946 msec) in the mid-range of SOA (40-125 msec), and finally reaching an asymptote of average high accuracy levels (81%-93%) and fairly stable reaction times (939-764 msec) at the longer ranges of SOA (130-280 msec). In contrast, the curves describing the relationship of accuracy and reaction time to SOA in the spectral TOJ paradigm appear very different. First, the overall amount of variance explained by SOA in the spatial TOJ paradigm (86-97% for accuracy and 94-97% for response time) is significantly greater than that explained by SOA in the spectral TOJ paradigms (0-33% for accuracy and 36-62% for response time). Second, the range of changes in accuracy and reaction time as a function of SOA in the spectral TOJ paradigm appear to be totally different from that of the spatial TOJ paradigm. In contrast to the latter, there appears to be a very sharp change both in accuracy (67%-to-86%) and in response time (1318-to-901 msec) in a very short range of SOA values (10-30 msec) that brings both sets of variables to asymptotic levels (73%-89% and 1134-764 msec) by the mid-range of the SOA values (35-280 msec) (note the actual plotted values in Figs. 1 and 2).
These findings suggest that while performance of the spatial paradigm is explained almost fully by the temporal onset disparity between the stimuli, the same conclusion cannot be applied to performance of the spectral TOJ paradigm. For performance of the spatial paradigm, the stimulus duration and ISI are almost completely interchangeable. We recently published additional data supporting this conclusion (see Babkoff and Fostick, in press). However, this is not the case with the spectral TOJ paradigm, even though the shortest tone duration is long enough so that it does not alter tonality (Turnbull, 1944). Experiment 3 was designed to test whether participants can be trained as easily to perform both the spectral and the spatial TOJ paradigms within the same number of training sessions.

Experiment 3

The aim of Experiment 3 was to test the effect of training on performance while performing the spectral and spatial TOJ tasks. We contend that if the temporal manipulation produces the same improvement in judgment in both tasks then the effect of learning in both tasks should be fairly similar, if not the same. In Experiment 3 we compare the TOJ thresholds obtained by the same individuals from two consecutive sessions while performing each of the two TOJ paradigms.

Method

Participants. Thirty undergraduate students (mean age = 24 years, 42% females) performed the spatial TOJ task and 19 (mean age = 22 years, 89% females) performed the
spectral TOJ task. All were screened for normal hearing (at least 20 dB SPL for .5, 1, 2, and 4 kHz).

**Tasks and stimuli. Spectral TOJ.** Pairs of 15 msec pure tones, one at 1kHz and the other at 1.8kHz were presented diotically at 40 dB SL, and with 1 msec rise/fall time. On each trial participants were presented with a pair of tones, consisting of one high frequency and one low frequency tone. Participants were required to reproduce the order in which they heard the tones (high frequency tone first then low frequency tone, or the reverse) by pressing the relevant keyboard keys in the order corresponding to what they heard. Tone pairs were presented in random order using a two alternative forced choice, 2-down-1-up, adaptive staircase procedure. The initial ISI was 150 msec which was reduced by 20 msec at each step until it reached ISI= 90 msec. From 90 to 40 msec, ISI was reduced by 10 msec on each step, from 40 to 15 msec, ISI was reduced by 5 msec, and from 15 to 0 msec ISI was reduced by 2.5 msec on each step. The experiment was terminated after 10 reversals, and the threshold was calculated as the average of the last eight reversals. The task was performed in two consecutive sessions. TOJ threshold was recorded for each session separately.

Experimentation followed the four-phases of training as described in Experiment 1. Absolute hearing threshold for the pure tone stimuli were measured prior to performing the TOJ tasks, as described in Experiment 1.

**Spatial TOJ.** Pairs of 15 msec 1 kHz pure tones were presented monaurally at 40 dB SL, and with 1 msec rise/fall time. On each trial participants were presented with a pair of tones, consisting of one tone to the left ear and one tone to the right ear. On half of the trials the order of the tones was: left ear first then right ear, and the reverse order in
the other half of the trials. Participants were required to reproduce the order in which they heard the tones (left ear first then right ear, or the reverse) by pressing the relevant keyboard keys in the order corresponding to what they heard. Tone combinations were presented in a random order using a two alternative forced choice 2-down-1-up adaptive staircase procedure with exactly the same steps as in the spectral TOJ paradigm. The experiment was terminated after 10 reversals, and the threshold was calculated as the average of the last eight reversals. The task was performed in two consecutive sessions. TOJ threshold was recorded for each session separately.

Experimentation followed the four-phases of training as described in Experiment 1. Absolute hearing threshold for the pure tone stimuli were measured prior to performing the TOJ tasks, as described in Experiment 1.

**Apparatus.** A Pentium1 personal computer controlled the stimulus presentation and recorded responses for both paradigms. The auditory stimuli were presented through TDH-49 headphones. Tasks were programmed using Matlab™ software version 6.5. Screening for hearing level was done using Danplex DA64 or Maico Hearing Instruments Ltd MA32 audiometers.

**Procedure.** Prior to experimentation the participants were informed that the study was designed to test their ability to judge the order of tones, in two consecutive sessions, and they signed an informed consent. The study was done using a between-subjects design and the participants were randomly assigned to either performing the spectral or the spatial task. Both Spectral and Spatial TOJ tasks performance was preceded by a training phase.
Results

Figure 4 presents ISI thresholds for determining TOJ in the two sessions for the spectral and spatial TOJ tasks. Data were analyzed by a mixed factorial ANOVA with session as the within-subjects variables and task as the between-subjects variable. A significant effect was found for session ($F_{(1,37)} = 22.454, p < .001$), but not for task ($F_{(1,37)} = 2.520, \text{n.s.}$). Furthermore, a significant task X session interaction was found ($F_{(1,37)} = 24.896, p < .001$). Paired-samples t-test between sessions for each task separately revealed significant differences between session 1 and session 2 in the spectral TOJ ($t_{(8)} = 3.021, p < .05$), but not in the spatial TOJ task ($t_{(29)} = .444, \text{n.s.}$).

Discussion

In auditory temporal order studies, when presented with pairs of tones, participants are required to decide which one was presented first and which was presented second. Our expectations were that improvement would occur from session one to session two, as a result of greater familiarity with the task and stimuli. We certainly did not expect a change in threshold by a factor of five (from 45 to 9 msec, spectral TOJ). Moreover, we expected that if a similar learning process takes place when performing either the spatial or spectral TOJ task, the learning curves would be very similar for the two paradigms. The results of Experiment 3 show that while the spatial TOJ thresholds remain very similar for the two sessions, the performance on the spectrally-based TOJ task improved, not only significantly, but dramatically from the first session to the second. Similar findings were reported by Fink et al. (2005) who showed a significant reduction over three sessions in spectral TOJ thresholds, from 35 to 15 msec, while they
reported a small change from 55 to 50 msec in their spatial paradigm. These results show that training has very different effects on the spectral TOJ than on the spatial TOJ paradigm.

**General Discussion**

Spectral and spatial TOJ paradigms have been used to measure the ability to perceive the order of two brief tones for several decades. Recent studies that compared spectrally-based and spatially-based TOJ paradigms for the same participants have reported different estimates of TOJ thresholds for the two paradigms (Ben-Artzi et al., 2005; Fink et al., 2005, 2006a, 2006b; Szymaszek et al., 2006a, 2006b, 2009). Interestingly, the spatial TOJ threshold is consistently higher than the spectrally-based TOJ (Table 1). Fink et al. (2005) suggested that these threshold differences can be explained by differences in the stimuli used. Indeed, in some of the studies the stimuli to be judged for the spatially-based TOJ were clicks and for the spectrally-based TOJ were tones (Fink et al., 2005, 2006a, 2006b; Szymaszek et al., 2006, 2009). However, Ben-Artzi et al. (2005) used similar tone stimuli for both spectral and spatial TOJ paradigms and also found significant differences between the two paradigms.

The purpose of the current study was to compare the response patterns of spectrally-based and spatially-based TOJ paradigms. This was first addressed by comparing the performance level at a very short ISI (5 msec) in each of the paradigms; second, by comparing the predictive value of SOA for TOJ accuracy and response time on the spectral and spatial TOJ paradigms; and third, by testing the extent to which training improves performance on the two paradigms.
The findings of Experiment 1 showed that while only two out of sixty-eight participants were able to accurately judge (75%) the temporal order of the tones at a very short ISI (5 msec) when performing the spatial paradigm, 31 of the same participants (46%) succeeded in doing so while performing the spectral paradigm. In addition, the average accuracy level of the spectral TOJ at this short ISI was much higher (74%) than that of the spatial TOJ (55%). Furthermore no correlation was found between the paradigms for ISIs of 60 msec and less. The results of Experiment 2 indicated that stimulus onset asynchrony (SOA) predicted 86 to 97% of the accuracy variance in the spatial paradigm and 94 to 97% of the response time variance. However, zero to 33% of the variance of response accuracy and 36 to 62% of the response time variance was predicted by SOA for the spectral paradigm. Moreover, examination of the curve relating accuracy and response time to SOA indicates a gradual monotonic increase in performance from chance to an asymptotic level of close to 100% and a concomitant relatively gradual decrease in response time to an asymptotic level for the spatial paradigm, while for the spectral paradigm, a very steep increase in accuracy and decrease in response time over a very short range of SOA values was found. Finally, in Experiment 3 the results revealed a sharp improvement in the performance level of the spectral TOJ paradigm on the second session, while no change in the same participants’ performance on the second session was found while performing the spatial TOJ task.

Studies conducted almost half a decade ago had suggested that two different types of sequence perception are used by participants, depending on the task aim (e.g., Warren 1974a, 1974b). When the task involved discriminating between different sequences, accurate performance was achieved with durations as brief as 2msec (Green, 1971;
Repeating the order of the sequence required much larger durations (Broadbent and Ladefoged, 1959). The findings of the current study indeed show two different response patterns, each of which seems to dominate in the different TOJ paradigm. The data of the current study show that: 1) spatial TOJ cannot be performed without sufficient temporal separation between the elements (Experiment 1); 2) implicate SOA as the main predictor for accurate performance (Experiment 2); and 3) and that spatial TOJ threshold cannot be significantly improved within two sessions (Experiment 3). This response pattern might change with different methodology (for example different stimulus duration or quality), however, we suggest that spatially-based TOJ as was measured in the current study fits the response pattern most likely associated with the direct perception of the elements in the pair of tones and their sequence. In contrast, the results also show that: 1) spectral TOJ can be performed by almost half of the subjects with minimal separation between the elements in the tone pair (ISI= 5msec) (Experiment 1); 2) SOA is only a partial predictor of the variance in performance accuracy and response time (Experiment 2); 3) and that there is a very profound training effect while performing this task, such that the estimated group TOJ threshold decreases from 45 to 9 msec in the second session (Experiment 3).

The comparison of response patterns to the two TOJ paradigms leads us to suggest that spatial TOJ and spectral TOJ are successfully performed by different perceptual strategies. When taking into account the previous suggestion regarding two types of sequence perception (e.g., Warren 1974a, 1974b), the different response patterns for the two paradigms lead us to suggest that spatial TOJ can only be accurately performed by direct perception of the elements in the tone pairs and their temporal
sequence; while the spectral TOJ can also be accurately performed by the holistic perception of the different patterns created by the order of the tonal elements.

If this explanation is correct, we may be better able to understand the very dramatic improvement in the spectral TOJ threshold from the first to the second session. If spectral TOJ can be accurately performed either by correctly perceiving the sequence of the two elements in the tone pair or by holistic perception of the different patterns created by the order of the elements, the participant has two options for discrimination when confronted with a two tone sequence. The training process may, thus, also involve learning to use the cue provided by the different patterns created by the temporal order of the two elements. Training may then involve shifting from the use of the cue of temporality to using the cue of holistic-envelope perception. In contrast, the spatial TOJ paradigm, even with the shortest SOAs used in this and other studies (15-20 msec) does not provide any holistic cues of pattern differences (see e.g., Babkoff and Sutton, 1963). Consequently, the judgment of order in the spatial TOJ paradigm can only be resolved by reference to the separate elements (the tones at the two ears) and their temporal separation. Hence, with the spatial paradigm, training can only involve improved use of the temporality cue, rather than learning to use an additional cue.

**Conclusions**

The findings of the current paper show different response patterns between spectral and temporal TOJ. We suggest that spectrally-based TOJ may reflect both the direct judgment of temporality based on the perception of the temporal relationship of the elements and the holistic judgment of the pattern of the envelope formed by the order of
the elements. In contrast, spatially-based TOJ as measured in the current study reflects the direct judgment of temporality based on the perception of the temporal relationship of the elements. This may explain findings of large differences between the thresholds of spectrally-based and spatially-based TOJ paradigms (Ben-Artzi et al., 2005; Fink et al., 2005, 2006a, 2006b; Szymaszek et al., 2006, 2009). Whether these response patterns would be found when different methodologies are used requires further study.

The suggestion that different response patterns between spectrally-based and spatially-based TOJ might reflect different underlying perceptual mechanisms, has relevance when comparing the performance of different populations. For example, findings of TOJ deficit (e.g., in populations of dyslexic readers, aphasic patients, older adults, or sleep deprived young adults) when using the spectral TOJ paradigm may not necessarily reflect difficulty in discriminating temporality within the same temporal range as the control population, but rather a difficulty in using the "holistic-envelope" cue or even a difficulty in learning how to use the "holistic envelope" cue. Further research should examine the mechanisms underlying the holistic-envelope perception. For example, one might examine whether the frequencies of the tonal elements in a spectral TOJ paradigm and/or the difference in frequencies of the two tonal elements in the pair affect spectral TOJ threshold.
References


Fostick, L., Zukerman, G. (2010). Auditory temporal deficit among the elderly: Sleep deprivation as a model for normal aging. Presented at the International Conference on Adult Hearing Screening, Cernobbio (Como Lake), Italy.


### Table 1. TOJ thresholds reported in the literature

#### a. Spectral TOJ

<table>
<thead>
<tr>
<th>Study</th>
<th>Stimuli</th>
<th>Method</th>
<th>N</th>
<th>Age</th>
<th>ISI (in msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fink et al. (2005)</td>
<td>10ms; 800Hz &amp; 1.2kHz</td>
<td>Staircase</td>
<td>20</td>
<td>20-35</td>
<td>15</td>
</tr>
<tr>
<td>Fink et al. (2005)</td>
<td>10ms; 800Hz &amp; 1.2kHz</td>
<td>YAAP</td>
<td>20</td>
<td>20-35</td>
<td>20</td>
</tr>
<tr>
<td>Fink et al. (2006b)</td>
<td>10ms; 600Hz &amp; 1.2kHz</td>
<td>YAAP</td>
<td>45</td>
<td>21-50</td>
<td>21</td>
</tr>
<tr>
<td>Fostick et al. (2008)</td>
<td>15ms; 1 &amp; 1.8kHz</td>
<td>Constant Stimuli (Threshold criteria at 75%)</td>
<td>50</td>
<td>20-31</td>
<td>2</td>
</tr>
<tr>
<td>Kanabus et al. (2002)</td>
<td>15ms; 300Hz &amp; 3kHz</td>
<td>Constant Stimuli (Threshold criteria at 75%)</td>
<td>12</td>
<td>20-25</td>
<td>40</td>
</tr>
<tr>
<td>Stevens &amp; Weaver (2005)</td>
<td>20ms; 1 &amp; 4kHz</td>
<td>Adaptive staircase (Levitt, 1970)</td>
<td>11</td>
<td>31.3±7.9</td>
<td>19</td>
</tr>
<tr>
<td>Szymaszek et al. (2006, 2009)</td>
<td>10ms; 400Hz &amp; 3kHz</td>
<td>Adaptive Procedure</td>
<td>17</td>
<td>20-19</td>
<td>33</td>
</tr>
</tbody>
</table>
### b. Spatial TOJ

<table>
<thead>
<tr>
<th>Study</th>
<th>Stimuli</th>
<th>Method</th>
<th>N</th>
<th>Age</th>
<th>ISI Threshold (in msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babkoff et al. (2005)</td>
<td>10 ms; 1 and 1.5kHz tones</td>
<td>Constant stimuli</td>
<td>18</td>
<td>Mean age = 63</td>
<td>23.8</td>
</tr>
<tr>
<td>Ben-Artzi et al. (2005)</td>
<td>15 ms; 300Hz and 600Hz tones</td>
<td>Constant stimuli</td>
<td>26</td>
<td>20-26</td>
<td>49</td>
</tr>
<tr>
<td>Fink et al. (2005)</td>
<td>1 ms clicks</td>
<td>Adaptive staircase</td>
<td>20</td>
<td>20-35</td>
<td>50</td>
</tr>
<tr>
<td>Fink et al. (2005)</td>
<td>1 ms clicks</td>
<td>YAAP</td>
<td>20</td>
<td>20-35</td>
<td>50</td>
</tr>
<tr>
<td>Fink et al. (2006b)</td>
<td>1 ms clicks</td>
<td>YAAP</td>
<td>49</td>
<td>21-50</td>
<td>56.5</td>
</tr>
<tr>
<td>Fostick et al. (2010)</td>
<td>15 ms; 1.8kHz tones</td>
<td>Constant Stimuli</td>
<td>18</td>
<td>20-35</td>
<td>60</td>
</tr>
<tr>
<td>Kinsbourne et al. (1991)</td>
<td>1 ms clicks</td>
<td>Ascending method of limits</td>
<td>21</td>
<td>Mean age = 47</td>
<td>26.5</td>
</tr>
<tr>
<td>Kolodziejczyk &amp; Szelag (2008)</td>
<td>15 ms; 300 Hz tones</td>
<td>Constant Stimuli</td>
<td>17</td>
<td>19-25</td>
<td>37</td>
</tr>
<tr>
<td>Lotze et al. (1999)</td>
<td>1 ms clicks</td>
<td>Down-staging method</td>
<td>2</td>
<td>27-31</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>from 50 ms in 10 ms steps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lotze et al. (1999)</td>
<td>1 ms clicks</td>
<td>Down-staging method</td>
<td>5</td>
<td>24-29</td>
<td>21</td>
</tr>
<tr>
<td>Study</td>
<td>Presentation Duration</td>
<td>Method</td>
<td>N</td>
<td>Age Range</td>
<td>Mean Age</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------</td>
<td>---------------------------------</td>
<td>---</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Lotze et al. (1999)</td>
<td>1 ms clicks</td>
<td>ML-Bayes method</td>
<td>5</td>
<td>24-30</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Emmerson, 1986; Treutwein, 1997)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Szymaszek et al. (2006, 2009)</td>
<td>1 ms clicks</td>
<td>Adaptive Procedure</td>
<td>17</td>
<td>20-19</td>
<td>68</td>
</tr>
<tr>
<td>Von Steinbuchel et al. (1999)</td>
<td>1 ms clicks</td>
<td>Adaptive staircase</td>
<td>17</td>
<td>Mean age =</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Threshold criteria at 75%)</td>
<td></td>
<td>49.5</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. $R^2$ for Accuracy and response time data by SOA for spatial and spectral task in different fitting functions

<table>
<thead>
<tr>
<th>Fitting function</th>
<th>$R^2$ Spatial task</th>
<th>$R^2$ Spectral task</th>
<th>Fisher z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>.860</td>
<td>.331</td>
<td>3.47***</td>
</tr>
<tr>
<td>RT</td>
<td>.935</td>
<td>.362</td>
<td>4.76***</td>
</tr>
<tr>
<td>Polynomial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>.967</td>
<td>.536</td>
<td>5.11***</td>
</tr>
<tr>
<td>RT</td>
<td>.969</td>
<td>.620</td>
<td>4.76***</td>
</tr>
<tr>
<td>Gaussian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.968</td>
<td>0.000</td>
<td>8.52***</td>
</tr>
<tr>
<td>RT$^1$</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Logistic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.968</td>
<td>0.292</td>
<td>6.39***</td>
</tr>
<tr>
<td>RT$^1$</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

RT = response time

$^1$Fitting data is not available since RT trendline is negative.
Figure 1. The performance of spatially-based and spectrally-based TOJ paradigms. (a) Accuracy by ISI; (b) Percentage of participants that performed at accuracy level of at least 75%. Results for ISI=5msec are the focus of the current analysis and marked with frame.
Figure 2. Accuracy and RT by SOA for the spatially-based TOJ paradigm.
Figure 3. Accuracy and RT by SOA for the spectrally-based TOJ paradigm.
Figure 4. Spectral and Spatial TOJ thresholds in two consecutive sessions.